Handbook of Psychological Assessment in Primary Care Settings

The second edition *Handbook of Psychological Assessment in Primary Care Settings* offers an overview of the application of psychological screening and assessment instruments in primary care settings. This indispensable reference addresses current psychological assessment needs and practices in primary care settings to inform psychologists, behavioral health clinicians, and primary care providers of the clinical benefits that can result from utilizing psychological assessment and other behavioral healthcare services in primary care settings.

**Dr. Maruish** is the President of Southcross Consulting which offers consultation to test publishers, healthcare research organizations, and others regarding psychological assessment products, publications, and strategies. He has published 10 books and numerous book chapters and scientific papers on psychological testing and assessment. He is a Fellow of both the American Psychological Association and the American Academy of Assessment Psychology.

“The area of psychological assessment in primary care settings is certainly a hot topic, both within the assessment field, but also more generally within clinical psychology training and practice issues. The central focus of this handbook on integrated assessment within primary care settings [is] important and timely . . . Assessment is a critical component of integrative psychology services within primary care services. This handbook will provide an overview of this topic, detailed recommendations concerning specific assessment instruments, and in-depth case examples.”

For Olivia and Jack
# Contents

*Contributors*  
xii
*Preface*  
xvii

## PART I  
**General Considerations**  
1 Introduction  
*Mark E. Maruish*

2 The History of Fragmentation and the Promise of Integration: A Primer on Behavioral Health and Primary Care  
*Benjamin F. Miller and Samuel H. Hubley*

3 Uses of Psychological Assessment in Primary Care Settings  
*John H. Porcerelli and John R. Jones*

4 Necessary Training for Psychologists Working in Primary Care Settings  
Appendix A  
Appendix B  
*Jeffrey L. Goodie, Catherine M. Ware, and Christopher L. Hunter*

5 Selection of Psychological Measures and Associated Administration, Scoring, and Reporting Technology for Use in Primary Care Settings  
*Mark E. Maruish*

6 Screening for Psychiatric Disorders in Primary Care Settings  
*Leonard R. Derogatis*

7 Clinical and Program Monitoring and Outcomes Assessment for Behavioral Health in Primary Care  
*Kent A. Corso*

## PART II  
**Psychological Assessment Instruments and Procedures**  
8 The Clinical Interview in Integrated Primary Care  
*Alexa J. Connell and Amber Hewitt*

9 Screening for Depression  
*Paula Rhode Brantley and Phillip J. Brantley*
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Geriatric Depression Scale</td>
<td>277</td>
</tr>
<tr>
<td></td>
<td><em>Elizabeth A. DiNapoli and Forrest Scogin</em></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Assessment of Suicidal Risk</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td><em>Julia A. Harris, Erika M. Roberge, Kent D. Hinkson Jr, and Craig J. Bryan</em></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Assessment of Anxiety in Primary Care</td>
<td>319</td>
</tr>
<tr>
<td></td>
<td><em>Amber A. Martinson, Julia R. Craner, and Renn U. Sweeney</em></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Assessment of Stress</td>
<td>355</td>
</tr>
<tr>
<td></td>
<td><em>Phillip J. Brantley and Paula Rhode Brantley</em></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Screening for Cognitive Impairment</td>
<td>369</td>
</tr>
<tr>
<td></td>
<td><em>Michael D. Franzen</em></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Substance Abuse Screening and Assessment in Primary Care Settings</td>
<td>389</td>
</tr>
<tr>
<td></td>
<td><em>Santina Wheat, Candice Norcott, and Mary R. Talen</em></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Assessment of Pain in Primary Care Settings</td>
<td>411</td>
</tr>
<tr>
<td></td>
<td><em>Robert J. Gatchel, Richard C. Robinson, Andrew R. Block, and Natalie N. Benedetto</em></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Posttraumatic Stress Disorder in Integrated Primary Care: Strategies for Effective Implementation of Screening and Assessment</td>
<td>429</td>
</tr>
<tr>
<td></td>
<td>Appendix A</td>
<td>457</td>
</tr>
<tr>
<td></td>
<td>Appendix B</td>
<td>458</td>
</tr>
<tr>
<td></td>
<td>Appendix C</td>
<td>460</td>
</tr>
<tr>
<td></td>
<td><em>Jonathan P. Muther and Yajaira Johnson-Esparza</em></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Assessment of Sleep Disorders in Primary Care</td>
<td>467</td>
</tr>
<tr>
<td></td>
<td><em>Skye Ochsner Margolies, Bruce D. Rybarczyk, Allison Baylor, and Sarah Griffin</em></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Assessment of Eating Behavior in Primary Care</td>
<td>499</td>
</tr>
<tr>
<td></td>
<td><em>Karen E. Stewart, Autumn Lanoye, Laura Milliken, and Vanessa Milsom</em></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Assessment of Health Status and Health-Related Quality of Life</td>
<td>529</td>
</tr>
<tr>
<td></td>
<td><em>Mark E. Maruish</em></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Assessment of Disease Impact</td>
<td>565</td>
</tr>
<tr>
<td></td>
<td><em>Martha S. Bayliss, Mark Kosinski, and Jakob B. Bjorner</em></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>Symptom Checklist-90-Revised, Brief Symptom Inventory, and BSI-18</td>
<td>599</td>
</tr>
<tr>
<td></td>
<td><em>Leonard R. Derogatis</em></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Using the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) in Behavioral Medicine Settings</td>
<td>631</td>
</tr>
<tr>
<td></td>
<td><em>Ryan J. Marek and Yossef S. Ben-Porath</em></td>
<td></td>
</tr>
<tr>
<td><strong>PART III</strong></td>
<td>Examples of Integrated Healthcare Programs</td>
<td>663</td>
</tr>
<tr>
<td>24</td>
<td>Improving Care for a Primary Care Population: Persistent Pain as an Example</td>
<td>665</td>
</tr>
<tr>
<td></td>
<td><em>Patricia J. Robinson and David Bauman</em></td>
<td></td>
</tr>
</tbody>
</table>
25 Psychological Assessment in the Veterans Health Administration (VHA) Model of Integrated Primary Care
Lisa K. Kearney, Laura O. Wray, Katherine M. Dollar, and Andrew S. Pomerantz 709

26 The Virginia Commonwealth University Primary Care Psychology Training Network
Bruce D. Rybarczyk, Karen E. Stewart, Paul B. Perrin, and Zachary Radcliff 733

27 Implementation of the Primary Care Behavioral Health Model at a Federally Qualified Health Center
Neftali Serrano, Meghan Fondow, and Elizabeth Zeidler Schreiter 763

Index 783
Contributors

David Bauman
Central Washington Family Medicine Clinic
Yakima, WA

Martha S. Bayliss, MSc
QualityMetric, Inc.
Lincoln, RI

Allison Baylor, BA
Virginia Commonwealth University
Richmond, VA

Natalie N. Benedetto, MA
University of Texas Southwestern Medical Center
Dallas, TX

Yossef S. Ben-Porath, PhD
Kent State University
Kent, OH

Jakob B. Bjorner, MD, PhD
QualityMetric, Inc.
Lincoln, RI

Andrew R. Block, PhD
Texas Back Institute
Plano, TX

Paula Rhode Brantley, PhD
Tulane University School of Medicine
New Orleans, LA

Phillip J. Brantley, PhD
Pennington Biomedical Research Center
Baton Rouge, LA

Craig J. Bryan, PhD
National Center for Veterans Studies and the University of Utah
Salt Lake City, UT
Alexa J. Connell, PhD
University of Massachusetts Medical School
Worcester, MA

Kent A. Corso, PsyD
NCR Behavioral Health
Springfield, VA

Julia R. Craner, PhD
Spectrum Health
Grand Rapids, MI

Leonard R. Derogatis, PhD
Johns Hopkins University School of Medicine
Baltimore, MD

Elizabeth A. DiNapoli, PhD
University of Alabama
Tuscaloosa, AL

Katherine M. Dollar, PhD
VA Center for Integrated Healthcare
West Haven, CT

Meghan Fondow, PsyD
Access Community Health Centers
Madison, WI

Michael D. Franzen, PhD
Temple University School of Medicine
Philadelphia, PA

Robert J. Gatchel, PhD
University of Texas at Arlington
Arlington, VA

Jeffrey L. Goodie, PhD
Uniformed Services University of the Health Sciences
Bethesda, MD

Sarah Griffin, BA
Virginia Commonwealth University
Richmond, VA

Julia A. Harris, BS
National Center for Veterans Studies and the University of Utah
Salt Lake City, UT

Amber Hewitt, PsyD
University of Massachusetts Medical School
Worcester, MA
Kent D. Hinkson Jr, BS  
National Center for Veterans Studies and the University of Utah  
Salt Lake City, UT

Samuel H. Hubley, PhD  
University of Colorado School of Medicine  
Aurora, CO

Christopher L. Hunter, PhD  
Defense Health Agency  
Falls Church, VA

Yajaira Johnson-Esparza, PhD  
Salud Family Health Centers  
Commerce City, CO

John R. Jones, MA  
University of Detroit Mercy  
Detroit, MI

Lisa K. Kearney, PhD  
Veterans Health Administration, VA Central Office  
San Antonio, TX

Mark Kosinski, MA  
QualityMetric, Inc.  
Lincoln, RI

Autumn Lanoye, MS  
Virginia Commonwealth University  
Richmond, VA

Ryan J. Marek, MA  
Kent State University  
Kent, OH

Skye Ochsner Margolies, PhD  
Eastern Virginia Medical School  
Norfolk, VA

Amber A. Martinson, PhD  
VA Salt Lake City Health Care System  
Salt Lake City, UT

Mark E. Maruish, PhD  
Southcross Consulting  
Burnsville, MN

Benjamin F. Miller, PsyD  
University of Colorado School of Medicine  
Aurora, CO
Laura Milliken, PhD  
James A. Haley Veteran’s Hospital  
Tampa, FL

Vanessa Milsom, PhD  
James A. Haley Veteran’s Hospital  
Tampa, FL

Jonathan P. Muther, PhD  
Salud Family Health Centers  
Commerce City, CO

Candice Norcott, PhD  
University of Chicago  
Chicago, IL

Paul B. Perrin, PhD  
Virginia Commonwealth University  
Richmond, VA

Andrew S. Pomerantz, MD  
Veterans Health Administration, VA Central Office  
White River Junction, VT

John H. Porcerelli, PhD  
University of Detroit Mercy  
Detroit, MI

Zachary Radcliff, MS  
Virginia Commonwealth University  
Richmond, VA

Erika M. Roberge, MS  
National Center for Veterans Studies and the University of Utah  
Salt Lake City, UT

Patricia J. Robinson, PhD  
Mountainview Consulting Group, Inc.  
Portland, OR

Richard C. Robinson, PhD  
University of Texas Southwestern Medical Center  
Dallas, TX

Bruce D. Rybarczyk, PhD  
Virginia Commonwealth University  
Richmond, VA

Elizabeth Zeidler Schreiter, PsyD  
Access Community Health Centers  
Madison, WI

Forrest Scogin, PhD  
University of Alabama  
Tuscaloosa, AL
Neftali Serrano, PsyD
Center of Excellence for Integrated Care
Cary, NC

Karen E. Stewart, PhD
Virginia Commonwealth University School of Medicine
Richmond, VA

Renn U. Sweeney, PhD
VA Salt Lake City Health Care System
Salt Lake City, UT

Mary R. Talen, PhD
Northwestern University Family Medicine Residency
Chicago, IL

Catherine M. Ware, MA
Uniformed Services University of the Health Sciences
Bethesda, MD

Santina Wheat, MD, MPH
Erie Family Health Center
Chicago, IL

Laura O. Wray, PhD
VA Center for Integrated Healthcare
Buffalo, NY
Preface

At the close of the twentieth century we witnessed major advances in the evolution of health care. Probably the most notable of these is the emergence and proliferation of managed health care. Resulting (for the most part) from out-of-control healthcare costs, managed care has changed the way health care is both financed and delivered. The impact of managed care has been both positive and negative, depending on whether your perspective is that of patient, provider, employer, or payor. Regardless of one’s perception, managed care is destined to continue in one form or another well into the twenty-first century.

As part of the evolution of managed care, a relatively recent healthcare trend began and has been gaining momentum for more than two decades. This is the movement toward the integration of behavioral healthcare services in primary care settings. In a chapter appearing in the first edition of this book, Goldstein, Bershadsky, and Maruish (2000) very succinctly stated the factors supporting this movement:

1. A large portion of patients who seek services from primary care providers experience significant psychological distress or symptomatology.
2. Primary care providers, in general, are not sufficiently skilled to identify or provide appropriate treatment to these patients.
3. Consequently, patients with behavioral health problems consume a large portion of the available primary care resources.
4. Identifying and adequately treating the behavioral health problems of primary care patients in the primary care setting has been shown to result in significant cost savings.
5. Consultation, liaison, and educational services offered by behavioral health professionals can be instrumental in ensuring the success of these intervention efforts. (p. 735)

Even in the year 2000, it was unlikely that the interest and efforts in what Goldstein et al. refer to as “primary behavioral healthcare” was considered just a passing fad. It rose out of a realization that a large portion of behavioral health care—up to as much as 50%—is dispensed in primary care settings by physicians, nurses, physician assistants, and others charged with overall management and delivery of health care to their patients, but who are generally unequipped to provide appropriate, efficient, and cost-effective behavioral health services to these same patients. Consequently, the integration of behavioral health care and primary care has provided yet another opportunity for psychologists and other behavioral healthcare professionals to become part of the solution to the healthcare crisis.

Recognizing the potential that this opportunity held for psychologists, especially with their unique skills and expertise in the identification, assessment and monitoring of behavioral health disorders, I recruited a number of experts in the areas of primary care psychology and clinical assessment
to contribute relevant and useful chapters to the first edition of this book. The book was intended not only for behavioral healthcare professionals, but also for primary care providers who may have wanted to integrate (with additional, supervised training, as necessary) instruments and/or other information discussed in the book into their practices. With the passing of more than 16 years since the first edition's publication, much has changed. The interest in and implementation of integrated primary and behavioral health care has grown tremendously, particularly in the wake of the enactment of the Patient Protection and Affordable Care Act in 2010. The changes that have taken place in the area have necessitated an updating and expansion of the information presented in 2000 and have served as the impetus for the development of this second edition of the original publication.

As with the first edition, I have organized this edition such that it will facilitate the readers' understanding of the subject matter. Part I contains seven chapters that deal with general topics and issues that provide a context for the information contained in the subsequent chapters. Chapter 1 discusses the prevalence of behavioral health disorders in primary care settings, their costs, and the concomitant need for better identification and treatment of mental health and substance use problems in those settings. Integrated primary and behavioral health care is identified as a solution to the existing problems—a solution that may take any of several forms, depending on the particular needs of the individual primary care setting. Chapter 1 also provides an overview of the potential contributions of various types of psychological assessment instruments for screening patients, planning an appropriate course of treatment, and assessing the outcomes of treatment.

The remainder of Part I consists of chapters that address various topics worthy of consideration for anyone who currently provides or plans to provide psychological assessment services in primary care settings. Along with a chapter that serves as a primer for integrated care, the topics addressed include the training necessary for individuals to successfully function as psychologists in primary care settings, an overview of the uses of assessment in these settings, and important criteria to consider when selecting psychological test instruments and modes of test administration and scoring to be used with those instruments in integrated settings. Rounding out Part I are separate chapters on screening for psychiatric disorders and on monitoring treatment progress and outcomes assessment with psychological test instruments.

Part II presents discussions of approaches and psychological test instruments appropriate to assist in the assessment of various types of psychological problems that commonly present themselves in primary care patients. These problems include depression, anxiety, stress, substance abuse, cognitive impairment, pain, suicidal risk, posttraumatic and other types of stress, disordered sleep and eating behaviors, impaired health-related quality of life, and the negative impact of chronic physical diseases. However, this section begins with a discussion of conducting what I and many others consider the key element to any psychological assessment—the clinical interview—in a primary care setting. Finally, Part II concludes with chapters describing in great detail each of a few specific instruments that have demonstrated their utility in assessing psychological and psychiatric problems in primary care, medical, and/or behavioral medicine patients for decades. These are the current version of the MMPI, the MMPI-2-RF, and the SCL-90-R and other members of the Derogatis “family of instruments.”

The third and final part of the book was developed to provide the reader with actual examples of how psychological assessment can be integrated into primary care settings. In addition, Part III includes a chapter describing one US graduate program specifically devoted to training students to work in integrated primary and behavioral healthcare programs. This program description serves to highlight how standard clinical training programs do not adequately prepare psychologists to function and serve patients and primary care staff in the manner or at the level required for them to be considered integral members of the integrated care team.
For more than the past two decades, the enterprise of integrated primary and behavioral healthcare service delivery has garnered progressively more interest, attention, and recognition for its potential contributions to the improvement of the US healthcare delivery system. As a result, the body of literature addressing the relevant issues has grown significantly, and quite probably it will continue to grow as the benefits of integrated medical and mental health/substance use services become more evident to patients, providers, and payers. I hope that psychologists, other behavioral health clinicians and primary care providers who currently are involved or are contemplating involvement in integrated primary behavioral healthcare services, as well as graduate students, interns and postdoctoral fellows training to provide such services, find the chapters that follow to be helpful in maximizing the potential benefits that can result from the provision of such services.

Mark E. Maruish
Burnsville, MN
PART I

General Considerations
CHAPTER 1

Introduction
Mark E. Maruish

The historically inefficient, costly, and unsustainable system of health care in the United States is going through dramatic changes. Spurred by the passage of the Patient Protection and Affordable Care Act (ACA, 2010) and focused on the linked goals of the Triple Aim—improved population health, improved quality or experience of health care by the individual patient, and reduction in healthcare costs (Berwick, Nolan, & Whittington, 2008; Laderman, 2015; Lewis, 2015; McDaniel & deGruy, 2014; Peek, Cohen, & deGruy, 2014)—efforts are underway to transform the nation’s healthcare delivery system into one that is more efficient, effective, and cost-effective. One way that this is being achieved is through a shift in the view of and the approach to health and disease from a biomedical model, which addresses mental disorders only if they can be explained biologically, to a biopsychosocial model (Engel, 1977) which also takes into account psychological and social factors and their interactions. This shift can be seen in the delivery of health care in primary care settings, and it may be viewed as one impetus for the ongoing movement toward more collaboration among medical and behavioral healthcare providers and varying degrees of the integration of physical and behavioral health care. Although integration can be seen in several types of medical specialty care settings such as obstetrics/gynecology (Poleshuck & Woods, 2014) and oncology (Kazak & Noll, 2015), its appearance and impact is perhaps most visible in primary care settings.

Accompanying the trend to integrate behavioral healthcare services in primary and specialty care settings are opportunities for psychologists to capitalize on their clinical, research, and educational training and skills to contribute to the improvement in health care during this time of transformation. Among the clinical skills that can have the most impact on primary care patients is the psychologist’s assessment skills. The clinical psychologist’s training and expertise in psychological testing-based assessment distinguishes him or her from other behavioral healthcare professions more than anything else. For this reason, the further development of psychological assessment skills as they can be utilized in and contribute to the delivery of quality, effective health care in primary care settings is the focus of this book.

The purpose of this chapter is to provide the reader with a brief overview of the impetus for, current interest in, and efforts toward the integration of behavioral health care in primary medical care settings, with particular emphasis on the role that psychological assessment can play in integrated physical and mental healthcare programs. The intent is not to present a comprehensive exposition of endeavors in this area; rather, it is hoped that the information contained herein will provide a context that facilitates an understanding of the detailed information presented in the chapters that follow. As a point of clarification, when used in this chapter, the term primary care will be meant to describe
the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practicing in the context of family and community.

(Institute of Medicine [IOM], 1994, p. 15)

HEALTHCARE COSTS

That the cost of health care in the United States is staggering and has continued to rise since the publication of the first edition of this book (Maruish, 2000a) should be of no surprise to anyone. And as will be shown, there are indications that these costs—for health care in general and behavioral health care—will continue to rise in the foreseeable future.

Overall Healthcare Costs

In the introductory chapter of the first edition of this book (Maruish, 2000b) it was noted that Ray (1996) reported the cost of health care in 1995 was about $1 trillion, or 14.9% of the gross domestic product (GDP), and was expected to rise by 20% by the year 2000. Based on Centers for Medicare and Medicaid Services (CMS) data reported by Martin, Hartman, Benson, Catlin, and the National Health Expenditure Accounts Team (2016), national health expenditure was $2.80 trillion, or 17.3% of the GDP in 2012; $2.88 trillion, or 17.3% of the GDP in 2013; and $3.03 trillion, or 17.5% of the GDP in 2014. Spending increased 3.8%, 2.9%, and 5.3% in those same three years, respectively. Per capita expenditure for each of the three years was $8,927, $9,115, and $9,523, respectively, representing a per capita growth increases of 3.0%, 2.1%, and 4.5%, respectively, over the previous year. Martin et al. also reported that for 2014, Medicare spending was $618.7 billion, Medicaid spending $495.8 billion, and private health insurance spending was $991.0 billion.

Nordal (2012) indicated that by 2020, healthcare costs are expected to rise to $4.7 trillion, a figure that will represent 19.8% of the GDP. This increase can in part be attributed to projected growth in Medicare enrollees and the expanded healthcare coverage mandated by the ACA. Also, CMS (2015b) projected average health spending to grow an average of 5.8% from 2014 to 2024 annually and to represent 19.6% of the GDP by 2024.

Behavioral Healthcare Costs

The projected costs for mental and substance use disorders in the future present a somewhat more positive picture than those for general health care. The Substance Abuse and Mental Health Services Administration (SAMHSA; 2014b) projects direct spending costs for behavioral disorder treatment (excluding comorbid health costs and indirect costs for things such as lost productivity and wages) to rise to $280.5 billion in 2020. This estimate includes the costs resulting from the implementation of the ACA and compares to reported behavioral health treatment costs of $171.7 billion in 2009. SAMHSA had projected a behavioral healthcare average annual cost increase of 4.6% from 2009 and 2020, which is less than the projected annual cost increase of 5.8% for all health care. One notable healthcare cost is prescription medications for behavioral healthcare patients. A report from the IMS Institute for Healthcare Informatics (2014) revealed that spending for mental health medications (antipsychotics and antidepressants) in 2013 was $23.9 billion. This represented a 5.2% decrease in spending from 2012.
There are indications that positive changes related to the delivery of behavioral health care—changes that are likely to impact behavioral healthcare costs—are taking place. Legislation has been enacted to eliminate the lack of parity that exists between general medical and behavioral healthcare benefits in health plans. This includes the 2008 Mental Health Parity and Addiction Equity Act and the 2010 ACA.

Perhaps as important is the realization of the patient care and financial benefits that can accrue from the integration of primary medical and behavioral health care. This has resulted not only in professional- and academic-level discussions and investigations, but also in the implementation of an increasing number of integrated programs in primary care settings (for example, see Hunter, Goodie, Oordt, & Dobmeyer, 2009; Kolbasovsky, Reich, Romano, & Jaramillo, 2005). Currently, the degree to which integrated service delivery is present in participating practices varies as a function of a number of factors, such as available funding, third-party reimbursement criteria, staff interest and commitment to the program, availability of resources, and office space limitations. Regardless of the extent to which these services are merged, efforts toward attaining this goal attest to the belief that any steps toward integrating behavioral healthcare services in primary care settings represent an improvement over the more traditional model of segregated service delivery.

PREVALENCE OF BEHAVIORAL HEALTH DISORDERS

Mental health and substance abuse disorders, that is, *behavioral health disorders*, have a significant presence in the US population. The demands of those suffering from these disorders can have a substantial impact on healthcare resources—both behavioral and medical—and thus merit the attention of those who are charged with their care as well as those attempting to control the associated costs.

Prevalence in the US General Population

A commonly cited statistic in the literature is that approximately 25% of the US adult population have a mental illness (e.g., see Reeves et al., 2011; Kessler & Wang, 2008). However, SAMHSA (2015) recently reported what might be the most current estimates of the prevalence of mental health and substance abuse disorder. These data come from the 2014 National Survey on Drug Use and Health (NSDUH) that was administered to 67,901 individuals ages 12 or older, of which 17,046 were aged 12 to 17. Data were analyzed separately for those aged 12 to 17 (adolescents), those aged 18 to 25 (younger adults), and those aged 26 years and older (older adults). Results indicated that 9.4% of adolescents, 22.0% of younger adults, and 8.3% of older adults reported the use of illicit drugs (marijuana/hashish, cocaine, inhalants, hallucinogens, heroin, and nonmedical use of prescription-type drugs) during the past month. Heavy alcohol use (five or more drinks on the same occasion on five or more days during the past 30 days) was reported for 1.0% of adolescents, 10.8% of younger adults, and 6.0% of older adults. Based on the American Psychiatric Association’s dependence and abuse past-year criteria from the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; American Psychiatric Association, 2000), 5.0% of adolescents, 16.3% of younger adults, and 7.1% of older adults had a substance use disorder during the past year. The percentages meeting the criteria for alcohol use disorders in the past year for the three age groups were 2.7%, 12.3%, and 5.9%, respectively.

As for mental health issues, the reported NSDUH results revealed that 10.7% of adolescents experienced a DSM-IV criteria-based major depressive disorder episode during the past year (SAMHSA, 2015). For all adults, 18.1% experienced a mental, behavioral, or emotional disorder (excluding substance use or developmental disorders) based on DSM-IV criteria during the past year. Serious mental
illness (i.e., one that substantially interfered with at least one major life activity) was found for 4.1% of the adult population during the past year.

The 2014 NSDUH also investigated the prevalence of co-occurring mental health and substance use disorders based on the available data. Thus, SAMHSA (2015) found that among all adults with a serious mental illness in the past year, 23.3% had a co-occurring substance use disorder while 18.2% for those with any mental illness had a co-occurring substance use disorder. In considering the entire US adult general population, it was found that 3.3% had both a mental illness and a substance use disorder while 1.0% of US adults had co-occurring serious mental illness and substance use disorders.

Among adolescents (aged 12 to 17), 28.4% who had a past-year major depressive episode (MDE) had a co-occurring substance use disorder while 10.5% of those with a past-year substance use disorder had a co-occurring major depressive episode (SAMHSA, 2015). The prevalence rate for co-occurring past-year substance use disorder and major depressive episode was 1.4% in the US adolescent general population; when considering only those with MDE with severe impairment, the prevalence dropped to 1.1%.

In considering comorbid disorder patients, it is also important to be mindful of how many of them receive needed treatment. SAMHSA (n.d.) reports that approximately 8.9 million individuals have such disorders. Importantly, only 7.4% receive treatment for both disorders while 55.8% do not receive any treatment.

Unfortunately, with the exception of major depressive disorder, the reported NSDUH results do not include prevalence estimates for specific mental disorders. However, disorder-specific prevalence estimates for the US general population derived from earlier surveys, surveillance systems, and studies are available. Some of these data are presented in Table 1.1. One will note that some of the reported prevalence estimates for a given disorder may vary from study to study. Differences in samples, the time at which the survey was conducted, and other factors likely play into the variations in findings. Regardless, the data indicate that mental and substance use disorders present a significant problem in the US general population.

Table 1.1 A Sample of Estimated Prevalences in the US General Population for Common Behavioral Health Disorders and Conditions

<table>
<thead>
<tr>
<th>Source</th>
<th>Survey/Study</th>
<th>Year</th>
<th>Disorder/Condition</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Obsessive-compulsive disorder</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PTSD</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Social phobia</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Major depressive disorder</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dysthymia</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Bipolar I-II disorders</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alcohol dependence</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Drug dependence</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Any disorder</td>
<td>26.2</td>
</tr>
<tr>
<td>Centers for Disease Control and Prevention (2011)</td>
<td>Behavioral Risk Factor Surveillance System (BRFSS)b</td>
<td>2006b</td>
<td>Current depression</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2008b Current depression</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2006b Major depressive disorder</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2008b Major depressive disorder</td>
<td>3.0</td>
</tr>
</tbody>
</table>
### Prevalence in Primary Care Settings

The literature is replete with data that attests to the high prevalence of patients with clinically significant behavioral health problems that are seen in primary care settings. Table 1.2 presents only a small sample of the data that demonstrate the frequency with which primary care providers encounter these patients.

<table>
<thead>
<tr>
<th>Source</th>
<th>Survey/Study</th>
<th>Year</th>
<th>Disorder/Condition</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2007&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Serious psychological distress</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2009&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Serious psychological distress</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2009&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Mentally unhealthy days</td>
<td>3.5</td>
</tr>
<tr>
<td>National Health and Nutrition Examination Survey (NHANES)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2005–2008&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Current depression</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>National Health Interview Survey (NHIS)</td>
<td>2009&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Serious psychological distress</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>National Nursing Home Survey (NNHS)</td>
<td>2004</td>
<td>Mental illness as primary diagnosis</td>
<td>18.7–23.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Major depression</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Dysthymia</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Any anxiety disorder</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Generalized anxiety disorder</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Social phobia</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Specific phobia</td>
<td>7.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Any substance use disorder</td>
<td>9.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Any substance use disorder with a mood disorder</td>
<td>19.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Any substance use disorder with an anxiety disorder</td>
<td>17.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mood disorder with any substance use disorder</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Anxiety disorder with any substance use disorder</td>
<td>15.0</td>
</tr>
</tbody>
</table>

<sup>Note</sup>. Some listed prevalence rate data were extracted from a secondary source.

<sup>a</sup> Twelve-month rates
<sup>b</sup> Two-week rates
<sup>c</sup> Thirty-day rates
It is especially important to note the extent to which comorbidity of behavioral health disorders with one another can occur in primary care settings. Johnson et al. (1995) found that among the 5% of the 1,000 primary care patients from the PRIME-MD 1000 Study (Spitzer et al., 1995) identified with alcohol abuse and dependence (AAD) using the Primary Care Evaluation of Mental Disorders (PRIME-MD), 47.1% met the criteria for one or more other mental disorders. Thus, among those identified with AAD, 33.3% were also PRIME-MD-diagnosed with a mood disorder, 21.6% with an anxiety disorder, 5.9% with an eating disorder, and 13.7% with a somatoform disorder (see Table 1.2). Among the non-AAD patients, 35.2% were found to have one or more PRIME-MD-identified disorders, including 25.5% with a mood disorder, 17.5% with an anxiety disorder, 3.1% with an eating disorder, and 13.9% with a somatoform disorder. The rate of comorbidity of these disorders with one or more other PRIME-MD-identified disorders was 65% for mood disorders, 82% for anxiety disorders, 84% for eating disorders, and 73% for somatoform disorders.

Table 1.2 A Sample of Reported Prevalence Rates for Behavioral Health Disorders in Primary Care Settings

<table>
<thead>
<tr>
<th>Study/Source</th>
<th>Disorder</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perez-Stable et al. (1990)</td>
<td>Major depression</td>
<td>6–16</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Any</td>
<td>10–30</td>
</tr>
<tr>
<td>Jenkins (1997)</td>
<td>Psychosocial problems (UK)</td>
<td>Approximately 33</td>
</tr>
<tr>
<td>Johnson et al. (1995)</td>
<td>Alcohol abuse and dependence</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Mood disorder</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Anxiety disorder</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Eating disorder</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Somatoform disorder</td>
<td>14</td>
</tr>
<tr>
<td>Ciarcia (1997)</td>
<td>Major depression</td>
<td>6</td>
</tr>
<tr>
<td>Katon &amp; Schulberg (1992)</td>
<td>Depression (outpatient)</td>
<td>5–10</td>
</tr>
<tr>
<td></td>
<td>Depression (inpatient)</td>
<td>6–14</td>
</tr>
<tr>
<td>Locke &amp; Larsson (1997)</td>
<td>Somatization</td>
<td>50+</td>
</tr>
<tr>
<td>Institute for International Research (1997)</td>
<td>Anxiety disorders</td>
<td>20</td>
</tr>
<tr>
<td>Spitzer et al. (1994)*</td>
<td>Any</td>
<td>30–52</td>
</tr>
<tr>
<td></td>
<td>Any mood disorder</td>
<td>19–35</td>
</tr>
<tr>
<td></td>
<td>Major depression</td>
<td>7–19</td>
</tr>
<tr>
<td></td>
<td>Dysthymia</td>
<td>5–15</td>
</tr>
<tr>
<td></td>
<td>Any anxiety disorder</td>
<td>10–25</td>
</tr>
<tr>
<td></td>
<td>Generalized anxiety disorder</td>
<td>2–13</td>
</tr>
<tr>
<td></td>
<td>Any somatoform disorder</td>
<td>9–29</td>
</tr>
<tr>
<td></td>
<td>Probable alcohol abuse/dependence</td>
<td>3–7</td>
</tr>
<tr>
<td></td>
<td>Any eating disorder</td>
<td>1–7</td>
</tr>
<tr>
<td>Johnson et al. (1995)*</td>
<td>Probable alcohol abuse/dependence with</td>
<td>47.1</td>
</tr>
<tr>
<td></td>
<td>one or more other disorders</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probable alcohol abuse/dependence with</td>
<td>33.3</td>
</tr>
<tr>
<td></td>
<td>any mood disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probable alcohol abuse/dependence with</td>
<td>21.6</td>
</tr>
<tr>
<td></td>
<td>comorbid anxiety disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probable alcohol abuse/dependence with</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>comorbid eating disorder</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Probable alcohol abuse/dependence with</td>
<td>13.7</td>
</tr>
<tr>
<td></td>
<td>comorbid somatoform disorder</td>
<td></td>
</tr>
</tbody>
</table>

Note. Some listed prevalence rate data were extracted from secondary sources. *Data represent findings from the PRIME-MD-1000 study (Spitzer et al., 1994).
CHRONIC DISEASES AND BEHAVIORAL HEALTH DISORDERS

Chronic diseases present a huge burden on the US population in terms of both morbidity and mortality, as well as their costs—financial and otherwise—to the healthcare system responsible for treating individuals with chronic conditions. According to Loeppke (2008), more than 133 million people in the United States have one or more chronic conditions. Related to chronic conditions at the time were 70% of all deaths and 75% of the then $2 trillion in healthcare expenditures. Frequently compounding the burden of chronic illnesses is the presence of a comorbid behavioral health condition.

Chronic Conditions and Their Costs

Chronic disease presents a huge financial burden in the United States. Estimates are that from 70% (Newman & Steed, 2004) to 78% (Bodenheimer, Chen, & Bennett, 2009; Stanton & Rutherford, 2005) of healthcare costs can be attributed to people with chronic conditions. Moreover, the cost for individuals with multiple chronic conditions can be as much as seven times that for those with just one chronic condition (Morgan, 2012). According to the CDC (2009), the major chronic diseases that affect the US population are heart disease and stroke, cancer, diabetes, arthritis, obesity, respiratory diseases, and oral conditions (tooth decay, oral cancers, and periodontal disease). For example, data from the 2007 National Health Interview Survey (Pleis & Lucas, 2009) indicated chronic condition prevalence rates to be 8% for diabetes, 24% for hypertension, 7% for asthma, 6% for chronic pulmonary disease, 21% for arthritis, 6% for coronary heart disease, and 26% for obesity.

DeVol and Bedroussian (2007) identified the seven most common chronic diseases as being cancer, diabetes, hypertension, stroke, heart disease, pulmonary conditions, and mental disorders. They reported treatment costs for noninstitutionalized patients suffering from one or more of these conditions to be $277 billion in 2003, jumping to just over $1 trillion when indirect costs are also considered. The researchers projected that by 2023, the costs for these seven chronic conditions would grow to $4.2 trillion. DeVol and Bedroussian also projected substantial increases in the prevalence for from 2003 to 2023. For each of the seven studied conditions, these projected percentage prevalence increases are 62% for cancers, 53% for diabetes, 39% for hypertension, 31% for pulmonary conditions, 41% for heart disease, 54% for mental disorders, and 29% for stroke. In addition to these findings, the IOM (2011) reported that over 100 million American adults suffer from chronic pain, with an estimated annual cost of up to $635 billion in medical treatment and lost productivity.

Comorbidity of Chronic Illnesses and Behavioral Health Disorders

Healthcare consumers present themselves to medical healthcare providers for any number of reasons. Lipsitt (as cited in Locke & Larsson, 1997) reported that over 80 million physician visits account for eight common complaints: fatigue, back pain, headaches, dizziness, chest pain, dyspnea, abdominal pain, and anxiety. Locke and Larsson’s review of the literature suggested that the most common of the somatic complaints addressed by physicians could be placed into one of five symptom groupings—gastrointestinal, neurological, autonomic, cardiovascular, and musculoskeletal—and that many of these are commonly seen in anxious and depressed patients. The comorbidity of mental health disorders and medical conditions thus is not surprising. For example, data from the 2001–2003 National Comorbidity Survey Replication (NCS-R) found that about 25% of the adult population had a mental
disorder within a given year while about 58% had one or more general medical conditions (Druss & Walker, 2011). It also revealed that 68% of adults with a mental disorder indicated they also had one or more general medical disorders while 29% of those reporting a medical disorder also reported having a comorbid mental health disorder (Alegría, Jackson, Kessler, & Takeuchi, 2003; Druss & Walker, 2011; Kessler et al., 2004). The literature contains numerous examples of how comorbid mental disorders can exacerbate chronic conditions (for example, see American Psychological Association [APA], 2014c).

That physical complaints often belie the presence of an underlying behavioral health problem is most clearly illustrated in a 1991 national study of the prevalence of depressive symptoms by Zung, Broadhead, and Roth (1993). In this study, 765 family physicians from across the United States provided data from the administration of the Zung Self-rating Depression Scale (SDS) on nearly 76,000 adult patients, along with information related to the reason for their medical visit. The five most frequently reported reasons for seeing these physicians were for checkups, upper respiratory infections, hypertension, throat symptoms, and problems with skin/nails/hair. Among the top 25 most common reasons for seeing their physician, depression ranked 24th (1.2% of the sample) and anxiety ranked 25th (1.1% of the sample). At the same time, using a pre-established cutoff of the Zung score (SDS ≥ 55), 20.9% of the total sample were found to be experiencing clinically significant depressive symptoms. Note that Anderson and Grundy (2016) report that up to 40% of those being treated for a chronic disease have comorbid depression.

Patients suffering from a chronic condition are more likely than those without such a condition to also suffer from a psychological disorder (Wells, Golding, & Burnam, 1988), and the influence of a mental illness and chronic disease on each other is evident. For example, the CDC (2009) reported that chronic conditions can worsen depressive symptoms while depressive disorders can actually lead to chronic conditions. Data on the prevalence of the comorbidity of behavioral health disorders and other illnesses, diseases, or disorders commonly seen in primary care settings provide an even clearer picture of the situation. For example, Ciarcia (1997) reported significant comorbidity rates of depression with cancer (18–39%), myocardial infarction (15–19%), rheumatoid arthritis (13%), Parkinson’s disease (10–37%), stroke (22–50%), and diabetes (5–11%). In a 1988 study of a community sample of over 2,500 people, Wells, Golding, Burnam, et al. (cited in Academy of Psychosomatic Medicine, 1997) found that the six-month, risk-adjusted prevalence rate of mental disorder was 24.7% and 17.5% for those with and without a chronic medical condition, respectively. The lifetime mental disorder prevalence rates were 42.2% and 33%, respectively. As for specific disorders, Wells et al. found prevalence rates of comorbid mental disorders to be 37.5% for neurologic disorders, 34.5% for heart disease, 30.9% for chronic lung disease, 30.3% for cancer, 25.3% for arthritis, 22.7% for diabetes, and 22.3% for hypertension.

The comorbidity of serious mental illness (SMI) and chronic medical conditions is particularly notable. In the United States, people with an SMI (major depression, bipolar disorder, or schizophrenia) represent 4% of the adult population (Scharf et al., 2014). In 2012, this equated to approximately 9.6 million adults (SAMHSA, 2014a). On average, people with an SMI die 25 years earlier than the general population largely as a result of preventable medical conditions (Morgan, 2012).

Based on trended costs from Medicare, Medicaid, and commercial insurance enrollees in 2010, the 2012 costs for treating those suffering from comorbid chronic medical and mental health/substance use disorder conditions was estimated to be two to three times higher than those patients with chronic medical conditions without a comorbid behavioral health problem (Melek, Norris, & Paulus, 2014). This same study also revealed the effect of having a comorbid behavioral health disorder on the cost of treating chronic diseases. For example, commercial insurance enrollees’ per-member-per-month (PMPM) costs for what the CDC (2009) might consider the major medical conditions affecting
the US population, without and with the presence of a comorbid non-SMI mental health condition, respectively, are as follows: Congestive heart failure, $1,274 and $1,955; ischemic heart disease, $1,443 and $2,319; stroke, $1,673 and $2,590; cancer, $1,360 and $2,338; diabetes, $811 and $1,353; arthritis, $814 and $1,586; and chronic obstructive pulmonary disease, $992 and $2,088.

Moreover, as noted in the Hogg Foundation (2008) summary of findings from the literature,

When psychiatric disorders are not addressed in people with chronic medical illness, they have worse outcomes. Patients with chronic medical conditions are less able to take care of their illnesses or follow prescribed treatment (Katon & Ciechanowski, 2002). These patients feel and function worse than patients with the same medical illness who do not have depression (Katon, Von Korff, Lin, & Simon, 2001; Unützer et al., 2000; Wells et al., 1989). They are more likely to die from their illnesses than those without depression (Frasure-Smith, Lesperance, & Talajic, 1993). These patients also have higher medical costs (Simon, Von Korff, & Barlow, 1995; Unützer et al., 1997).

(p. 9)

The only positive aspect of the comorbidity of mental health and chronic medical disorders may be the benefit derived from mental health treatment on both conditions. As Runyan (2011) pointed out,

There is . . . a sizable and growing body of literature demonstrating the clinical and economic benefits of treating mental health conditions associated with chronic medical conditions with a variety of pharmacological and nonpharmacological interventions. These studies consistently demonstrated positive effects on clinical outcomes, cost reduction, and decreased fragmentation of care, which often reduces redundancy in services that drive up health care costs.

(p. 62)

Similar conclusions were asserted by the CDC (n.d.).

PRIMARY CARE: THE DE FACTO BEHAVIORAL HEALTHCARE SYSTEM

Given the of prevalence of behavioral health disorders in the United States, the frequency of their comorbidity with chronic physical disorders and the rate at which patients suffering from both types of disorders seek treatment from primary care physicians, it seems only logical that primary care has been considered the “de facto behavioral healthcare system” for more several years now. This term, commonly attributed to Regier et al. (1993), appears to be as apt a description of the current system of behavioral healthcare delivery as it was more than two decades ago. Numerous studies and reports providing evidence to this effect and often contained in bullet-point lists of findings (for example, see Gray, Brody, & Johnson, 2005; Gunn & Blount, 2009; Kessler, 2009; Kuramoto, 2014) attest to claims that primary care settings and providers are the major points of provision of behavioral healthcare services.

A review of the literature reporting that primary care is the main behavioral healthcare system in the United States reveals several types of factors and conditions that support this contention. One condition has to do with the presentation of psychological symptomatology in the primary care provider’s office. Hunter et al. (2009) indicate 80% of those with symptoms that cannot be explained medically receive psychosocial treatment from a primary care physician. More generally, mental health issues ranks third among the reasons for people seeking treatment at healthcare facilities,
surpassing the rates for some common chronic conditions such as asthma, diabetes, and heart conditions (Holden et al., 2014). Moreover, Strosahl (1997) noted that the direct or indirect management of 80% of patients with psychological disorders is provided by primary care.

Particularly notable is the treatment of depression in primary care settings. For example, Olfson et al. (2002) reported that from 1987 to 1997, the number of patients per thousand being treated for depression increased from 7.3 to 23.3. During this time span, the percentage of patients being prescribed SSRIs also increased, from 37.3% to 74.5%. According to Coyne, Thompson, Klinkman and Nease (2002), studies show that in most cases depression is treated in primary care, with 66% to 75% of cases of depression being treated by primary care providers (PCPs) rather than by specialists in mental health.

A more general condition has to do with the percentage of primary care office visits involve complaints that have a psychological component. Hunter et al. (2009) and Gunn and Blount (2009) reported that up to 70% of primary care visits can be attributed to psychological problems.

Another factor involves the prescription of psychotropic medications in primary care settings. Kuramoto (2014) reported that 67% of psychoactive drugs and 80% of antidepressants are prescribed by primary care physicians. Moreover, 48% of all appointments for all psychoactive medications are with a PCP. Coyne et al. (2002) indicated that patients are increasingly more likely to go to and receive antidepressant medications from a PCP than from a psychiatrist. As Gray, Brody, and Johnson (2005) noted, In the space of about 10 years, the mental health system reinvented itself, changing from a specialist-driven system to a PCP-driven system. New medications replaced psychotherapy as the first line treatments for a broad range of depression and anxiety disorders as the majority of patients eagerly embraced pharmaceutically mediated solutions to their problems. The advent of SSRIs also promoted a new understanding of mental health disorders in society. They were akin to general medical disorders, less problems of the mind, more disorders of the brain. The new drugs also challenged psychotherapy’s—some would say—puritanical ethos that prized the hard work of psychotherapy over medication.

(p. 124)

A fourth factor involves how frequently chronic medical conditions are accompanied and exacerbated by comorbid psychological symptoms and disorders. Katon and Ciechanowski (2002) reported that high rates of behavioral health problems can be found with people with common medical disorders. The most common of these are cardiovascular disease, diabetes, hypertension, arthritis, and digestive disorders (Miller, Druss, Dombrowski, & Rosenheck, 2003). As summarized by Kessler (2009), “many patients attending primary care with a medical presentation are masking a psychological disorder, or have a psychological disorder that is affecting their acute or chronic problem” (p. 255). A more detailed discussion of the comorbidity of chronic medical conditions and behavioral health conditions was presented earlier in this chapter.

An interesting factor is the use of healthcare services by those with psychological problems, including a demonstrated preference for these patients to be treated by in a primary care setting rather than in a specialty mental health setting. Kessler and Stafford (2008) indicated that primary care provides more psychological services to people than the specialty mental healthcare system. Olfson et al. (2002) noted that while the percentage of patients being treated for depression involved in psychotherapy decreased from 71.1% in 1987 to 60.2% in 1997, the percentage of patients being treated for depression by primary care providers rose from 68.9% to 87.3% during that same period of time. Kuramoto (2014) reported that the percentage of patients refusing referral for mental health services ranges from 33% to 50% while Miranda, Hohnmann, and Attikison (1994) reported that of the patients with a mental health condition, less than one-third ever meet with a mental health clinician. The stigma associated with seeing a behavioral health practitioner (see Hogg Foundation, 2008; Holden et al.,
2014; Kuramoto, 2014; Laderman, 2015; Miller, Mendenhall, & Malik, 2009) contributes to the reluctance to seek mental health treatment with the system of mental health care. As noted in the Peterson, Miller, Payne-Murphy, and Phillips' (2014) analysis of the pooled data of 109,593 adults, primary care continues to be the sole source of treatment for most patients with mental health conditions.

Arean and Gum (2013) attribute the trend for the delivery of mental health services to move away from mental health specialty settings to general medical settings to several factors. These include the comorbidity of behavioral and physical health conditions, the greater familiarity with and preference for medical system over mental health settings, barriers to accessing mental health treatment, and the increasingly higher rate at which new generation psychotropic medication are being prescribed by non-psychiatric physicians. Overall, Arean and Gum view this set of circumstances as leading to less access to psychotherapy as a treatment option and lower quality of care offered in general medical settings.

In all, Laderman (2015) makes a strong case for the integration of primary and behavioral health care. In doing so, she points to the higher costs and poorer outcomes that come from the treatment of comorbid chronic medical and behavioral health conditions, attributing these results to fragmented care and no financial incentives to treat the whole person rather than discrete problems. She concludes that

> Integrating behavioral health and primary care makes sense, as primary care is already seeing many patients with behavioral health conditions; approximately 50% of visits for behavioral health conditions take place in primary care (Kessler et al., 2005), and estimates suggest that 20% to 40% of primary care patients have some behavioral health need (Prince et al., 2007).

(p. 1)

**TREATMENT OF BEHAVIORAL HEALTHCARE PROBLEMS BY PRIMARY CARE PROVIDERS**

As has been shown, the situation that exists in the United States today is one in which a) behavioral health problems of varying degrees of severity exist in significant number; b) more than half of the people with these problems seek treatment from their family physician or other primary care provider; and c) a significant proportion of patients with a chronic medical condition have a comorbid behavioral health problem, which can both exacerbate the condition and increase costs. The question one must ask is: How is the primary care medical system performing as the de facto behavioral healthcare system? The answer is clear: Not well. The inadequacy of primary care providers in dealing effectively with patients with behavioral health problems has been recognized numerous times in the literature. This inadequacy exhibits itself in two general areas of service delivery: problem detection and appropriate treatment.

**Inadequate Detection and Treatment of Behavioral Health Disorders**

Detection and treatment of mental health and substance abuse/dependence symptoms and disorders is problematic for the primary care provider. Selden’s (1997) and Burns’ (1997) reviews of the literature suggested that anywhere between one-third to one-half of patients with behavioral health disorders seen in primary care settings go undetected. Higgins (1994) found the rate of unrecognized mental illness in primary care settings to be 33% to 79% for adults and 44% to 83% for children, based on the studies employing DSM-III/DSM-III-R-based structured interviews.
The seriousness of these problems was highlighted by Barracough, Bunch, Nelson, et al. (as cited in Cole, Raju, & Barrett, 1997), who indicated that approximately 15% of patients with severe depression lasting one month or longer commit suicide, and that half of these patients see their physicians sometime during the month before their death. The American Psychological Association (APA; 2014b) reported that about 90% of those who committed suicide had a mental disorder, 40% of which saw their PCP within a month of their demise. Ahmedani et al. (2014) found that 83% of 5,894 individuals who committed suicide between 2000 and 2010 received health care when only 24% of these patients had a mental health diagnosis within the four-week period preceding their death. In another study, Ahmedani et al. (2015) found that about 95% of 22,287 individuals from 10 health systems who attempted suicide made a healthcare visit within the preceding year, with over 38% having made such a visit within a week before the attempt. Primary care and general medical outpatient settings were the most common sites for these visits, and racial/ethnic variations were noted in these percentages. As Unützer, Harbin, Schoenbaum, & Druss (2013) summarized:

When behavioral health problems are not effectively treated, they can impair self-care and adherence to medical and mental health treatments and they are associated with poor health outcomes and increased mortality. They are also associated with decreased work productivity (Katon, 2009; Wang, Simon & Kessler, 2008) and substantial increases in overall health costs.

(p. 2)

Numerous other examples of the PCPs’ inadequacies in the identification and treatment of behavioral health disorders are reported by several investigators throughout the literature. Coyne et al. (2002) indicated that there is a tendency for PCPs to overprescribe psychotropic medications, thus increasing the cost of the treatment and the risk of the patient experiencing side effects. Cited in Gray et al. (2005), Brody reported that PCPs fail to detect depression in 30% to 50% of the patients they treat, fail to follow up or adjust treatment once a mental disorder is diagnosed, and often do not prescribe adequate dosages and for an adequate duration. Gray et al. also cited a “Beyond Diagnosis” article indicating that 55% of patients taking antidepressants stopped taking their medication, almost half reported side effects, and 17% missed doses. Moreover, only about 30% of PCPs treating patients were noted to have asked patients with chronic depression about their depression on an ongoing basis. Data from a Milliman, Inc. study (cited in DeAngelis, 2010) indicated that up to two-thirds of those needing treatment for a behavioral health disorder are first seen in a general medical setting but only 12.7% receive treatment considered minimally adequate. Addressing the needs of healthcare delivery for women, Poleshuck & Woods (2014) cite several studies indicating OB/GYN providers experience problems in detecting and providing appropriate treatment to pregnant women with depression.

Kessler (2009) reported improvement in physicians’ ability to detect psychological issues in their patients, but their identification of patients with comorbid psychological disorder remains limited and nonsystematic. Even if behavioral health problems are detected, the treatment that is provided to these patients frequently may be inadequate.

Factors Related to Inadequate Identification and Treatment of Behavioral Health Disorders in Primary Care Settings

Gray et al. (2005) identified five problems with the PCP-driven de facto mental health system: (a) difficulties in making referrals to mental health professionals; (b) limited training in detection and management of mental disorders; (c) limited time to attend to these disorders; (d) the possibility that
care for mental health problems may not be compensated; and (e) the fact that patients may present with somatic symptoms, are reluctant to report psychological symptoms, and/or refuse to accept that their symptoms may be psychosocial in origin.

The inadequacies in the care provided to patients with behavioral disorders patients in primary care settings are the result of multiple factors. First, there are those associated with the medical professions or the providers themselves. Primary care providers generally practice within what Cole and Raju (1996) term the “biomedical model of illness,” which is grounded in the philosophies of dualism (the distinction of mind and body) and reductionism (understanding health problems at a molecular level of physiology). This model promotes the idea that the only problems worth attending to are physical problems, thus perpetuating the stigma of mental illness. Adoption of this model is reflected in the type of training medical providers have undergone and skills that they have (not) developed. This compares to the previously mentioned biopsychosocial model to understanding and treating the human condition.

Somewhat related to the dualistic philosophy is the theory of competing demands (Klinkman, 1997). According to this theory, primary care encounters are comprised of many demands for the PCP’s attention, such as the treatment of illness, patient requests, and providing preventative services. As there is not enough time to address each demand, the PCP must choose which to attend to. In making this choice, three interrelated domains—the clinician, the patient, and the “practice ecosystem”—influence the PCP. In turn, these domains are indirectly influenced by the general policy environment. In the primary care environment where physical symptoms are often given greater priority, there may be a negative relationship between severity of physical symptoms and the delivery of mental health services. At the same time, as Ledoux, Barnett, Garcini, and Baker (2009) note, the competing demands theory would predict that attention to psychological problems may become a higher priority as the severity of the related distress grows worse. Moreover, Laderman (2015) contends that there is fragmentation of the care provided by different caregivers. Contributing to this fragmentation is a lack of communication, information sharing, and collaboration in the development of a shared treatment plan. In addition, there is no financial incentive to treat the whole person instead of treating a series of individual diseases.

The problem of inadequate treatment is often exacerbated by the fact that PCPs neither have the skills nor the time to adequately assess or otherwise deal with a patient’s psychological problems (Gunn & Blount, 2009; Hogg Foundation, 2008). On average, PCPs have 12–15 minutes to deal with all of a patient’s health problems, including those of a physical, emotional, and/or social nature (Coyne et al., 2002; Rollman et al., 2002).

Patient-related factors also play into the picture of inadequate services. Here, probably the most prominent factor is the stigma of mental illness (Hogg Foundation, 2008; Holden et al., 2014; Kuramoto, 2014; Laderman, 2015; Runyan, 2011). The fear or shame commonly related to being diagnosed with a mental health disorder remains today in American society. A 2000 survey of over 3,200 adults conducted by the National Mental Health Association (as cited in Runyan, 2011) indicated that among those with a mental health condition, 42% reported embarrassment or shame over their symptoms. Of those without a mental health condition, 32% said they would seek treatment from their primary care provider for help with mental health issues while only 4% would specifically seek help from a mental health professional. Findings of what percentage of patients referred for appropriate, effective treatment that the PCP cannot provide vary considerably. Percentages from several different sources (as cited in APA, 2014b; DeAngelis, 2010; Hogg Foundation, 2008; Laderman, 2015; Miller et al., 2009) indicate that anywhere from 10% to 50% do not follow up with a mental health specialist upon referral from a PCP.

It is important to recognize that the stigma attached to having a mental illness is not the only reason that an appropriate referral of a patient to a mental health specialist for treatment often
does not always result in effective treatment. Trude and Stoddard's (2003) survey of over 6,500 primary care physicians indicated that almost 54% encountered problems in obtaining referrals for high-quality outpatient mental health services when medically necessary. Also, 53% encountered problems in obtaining referrals for high-quality inpatient mental health services when medically necessary. Type of practice (e.g., staff model HMO vs small group practice), availability of specialists to refer to, and gatekeeping responsibilities are among some of reasons offered by the authors as reasons for the problem. Other problems discussed in the literature include health plan barriers, cultural beliefs and attitudes, lack of coordination of efforts between the PCP and mental health providers, physical proximity of the patient to a mental health clinic, and inadequate insurance coverage (see Hogg Foundation, 2008; Laderman, 2015; Runyan, 2011).

Another reason for the inadequacies has to do with the healthcare system in which primary and behavioral healthcare providers must operate. There are actually two sets of circumstances operating: one related to patients who have been identified as having behavioral health problems, and another in which such problems have yet to be identified by the patient or a healthcare professional. Mitchell and Haber (1997) noted that in many instances the patient's co-payment for behavioral health services frequently is more than for primary care visits. This encourages the patient to seek either no treatment, or treatment that is less costly but provided by a less qualified healthcare professional (i.e., the PCP).

**Solutions to the Problem**

Approximately half of the behavioral healthcare services in the United States today are delivered in primary care settings by professionals limited in their ability to accurately identify and appropriately treat individuals suffering from mental health or substance abuse problems. And the fact is, patients with behavioral healthcare problems will continue to turn to their primary care provider rather than to a behavioral healthcare specialist for treatment, in spite of these well-intentioned professionals' frequent inability to provide the highest possible quality of service. Given this state of affairs, one must ask what can be done to improve the chances that patients needing behavioral healthcare services will receive appropriate, high-quality care.

As reported by the Hogg Foundation (2008), several approaches have been taken to improve the treatment of common mental disorders in primary care settings. Noted here were standardized screening, training PCPs on behavioral health matters and using practice guidelines, and efforts to increase PCP referrals to mental health specialists. All were determined to be necessary but not sufficient to improve behavioral health outcomes in primary care settings. Co-location also was identified as being ineffective unless implemented in settings with an infrastructure that promotes shared treatment and collaboration. However, Gunn and Blount (2009) identify co-location as being the most important step toward collaboration.

Similarly, Unützer et al. (2013) identified attempts to improve the detection and treatment of mental disorders in primary care settings. These included screening for mental disorders, improving the primary care provider’s level of knowledge and skills in the area of behavioral health care, developing treatment guidelines, referring patients to specialty mental health care, co-locating mental health specialists in primary care settings, and the use of call centers to support the care provided in primary care settings. According to Unützer and his colleagues, the literature indicates that none of these approaches, either alone or along with another approach, resulted in improvement in patient outcomes. What they did find in the literature was more than 70 randomized controlled trials supporting a **collaborative care** approach to treatment. Taking this approach requires a team consisting
of a PCP, care management staff (e.g., psychologist, clinical social worker, or nurse) and a psychiatric consultant, and employing measurement-based care, target-to-treatment, stepped care, and other components of a model of care for chronic illnesses.

An important consideration in understanding these findings is understanding what is meant by the term “collaboration” or, more specifically, “collaborative care.” This term is variously defined and is sometimes used interchangeably with the term “integrated care.” Numerous definitions for each term are available throughout the literature (for example, see Hogg Foundation, 2008; Kuramoto, 2014; Laderman, 2015; McDaniel et al. 2014; Miller et al., 2009). The results of Boon, Mior, Barnsley, Ashbury, and Haig’s (2009) interview study of 16 healthcare and related professionals indicated that most of the respondents differentiated between “collaboration” and “integration.” Here, collaboration was viewed as reflecting circumstances or situations in which professionals work together in care delivery; however, unlike integration, there is not as single, formal framework of organization that subsumes the collaborating professionals. This distinction generally seems to hold among the definitions one is likely to find in the literature. As Boon et al. noted, “whereas integration requires collaboration as a precondition, collaboration does not require integration” (p. 720).

For the purpose of this chapter, collaborative care is defined according to the general definition provided by Boon et al. (2009, p. 720), that being “a model of team care that [enables] health care practitioners to maintain their autonomy while working together in the absence of formal structures or processes to deliver optimal patient care.” In this case, “health care practitioners” means PCPs and behavioral healthcare providers.

Using the definition provided by the Association for Behavioral Health and Wellness (Croze, 2015), the term integrated care is used in this chapter to mean

whole person care that focuses on overall health; creates partnerships across all aspects of health; and is facilitated by a variety of clinical, structural, and financial arrangements and community supports that remove barriers between physical and behavioral healthcare.

(p. 1)

For further clarification of the term, one is directed to definition of integrated care provided by Laderman (2015, p. 1), that being

proactive care provided by an interdisciplinary team of primary care and behavioral health workers (e.g., psychiatrist, psychologist, social worker, nurse) who share responsibility and comanage patients. The behavioral team is a consistent presence where medical care is provided and has a strong, collaborative relationship with the medical team. This team [as noted by Peek (2013)] ‘may address mental health, substance abuse conditions, health behaviors (including their contribution to chronic medical illness), life stressors and crises, stress-related physical symptoms, and ineffective patterns of health care utilization.’

It is important to note that the degree of both collaboration and integration can vary, as is discussed in detail in Doherty, McDaniel, and Baird’s (1996) five levels of systematic collaboration and Heath, Wise Romero, and Reynolds’ (2013) six levels of collaboration/integration (see later).

McDaniel et al. (2014) note that integrated physical and mental health care in primary care settings is becoming the foundation for the changes that are taking place and will be taking place in the US healthcare system under the ACA and other forces. It is the system of care that psychologists will increasingly find themselves applying their assessment, treatment, and care management
skills—skills that enable important contributions to the overall treatment and well-being of individual patients—in the future. A brief overview of integrated care thus will be presented before addressing issues related to contributions of psychological assessments in integrated settings. A more detailed discussion of the integrated primary and behavioral health care is presented in Chapter 2 of this book.

INTEGRATION OF PRIMARY AND BEHAVIORAL HEALTHCARE SERVICES

The integration of primary medical care and behavioral health care has existed in various forms for many years. However, only recently has it gained the attention and prominence of both the medical and behavioral healthcare professions as a potentially effective and efficient means of attending to a significant portion of those patients suffering from behavioral health problems. The reasons for the growing interest in integrated primary and behavioral health care are as varied as the qualities that define it and the various options by which it can be provided.

Impetus for Integrating Primary and Behavioral Health Care

In 2008, the World Health Organization (WHO; 2008) presented their “seven good reasons for integrating mental health care into primary care.” Its logic was:

1. The burden of mental disorders is great.
2. Mental and physical health problems are interwoven.
3. The treatment gap for mental disorders is enormous.
4. Primary care for mental health enhances access.
5. Primary care for mental health promotes respect for human rights.
6. Primary care for mental health is affordable and cost effective.
7. Primary care for mental health generates good health outcomes. (pp. 21–22)

At the time of the publication of the first edition of this book, Maruish (2000b) identified several forces driving for the integration of services. These included practical considerations (e.g., recognition that primary care settings were and would continue to be key sources of behavioral healthcare service delivery, recognition of the value associated with the identification and treatment of behavioral health disorders), financial factors (e.g., the cost of primary care providers treating behavioral health problems in their practices), and external forces (e.g., providing behavioral healthcare services through a carve-out company to less costly “one-stop shopping;” growing demands for accountability information pertaining to behavioral healthcare services, such as those from accreditation organizations like the National Committee for Quality Assurance [NCQA]). While these pressures continue today, recently occurring circumstances are stimulating the move toward integration of services even more.

In addition, several circumstances related to health care have occurred over the years since the publication of the first edition of this book. Drum and Sekel (2013) provide a detailed summary of the evolution of health care in the United States from the standpoint of various approaches to healthcare services and payment practices (e.g., managed care, fee-for-service, carve outs) and pieces of legislation that have been enacted over the past several decades (e.g., the 2008 Paul Wellstone and Peter Domenici Mental Health Parity and Addiction Equity Act). Gray et al. (2005) also present a tabular summary of changes in the delivery of mental health services. However, as alluded to at the beginning of this chapter, perhaps one of the most important factors in the way that healthcare services are delivered and paid for now and will be in the foreseeable future is the passage of the 2010 ACA.
Although integrated care models were being promoted prior to the passage of the ACA, Arean and Gum (2013) anticipated that the dissemination of these models would increase in the coming years. The ACA provides for the financing to test new models of healthcare service delivery designed to improve patient outcomes (Melchert, 2015). Rozensky (2011, 2012) notes the ACA’s recognition of the importance of interprofessional care with regard to cost savings and quality of care. In doing so, he points to Section 3502 of the ACA that indicates the Secretary of Health and Human Services shall establish means to provide grants or contracts to support primary care through the establishment of community-based interdisciplinary teams of healthcare providers, which may include psychologists and other mental health providers. Also, Rozensky indicates that Section 935 of the ACA discusses interprofessional models of health care, including integrated physical and mental health care (also see Melchert, 2015). Moreover, Clay (2015) reported that the ACA requires insurers to cover services such as screening and counseling for problems such as depression, obesity, and intimate partner violence—all of which fall within a trained psychologist’s skill set.

The ACA and other efforts seek to transform the healthcare environment, with the goal of this transformation being the Triple Aim. The Triple Aim framework was developed in 2007 by the Institute for Healthcare Improvement (IHI; Lewis, 2015). Laderman (2015) notes that an integrated behavioral health strategy across the continuum of care is needed by those organizations wishing to achieve the Triple Aim. In order to do so, McDaniel and deGruy (2014) indicate that primary care systems are undergoing innovations, such as trying new models of financing, redesigning practices, and forming new teams and partnerships.

Qualities of Integrated Care

For primary and behavioral health care to be considered “integrated,” experts would argue that several conditions must be present. Most would probably agree that integrated systems would have to include at least two key components: co-location and collaboration. According to definitions provided by McDaniel et al. (2014), co-located care refers to “behavioral health (BH) and primary care providers (i.e., physicians, nurse practitioners) delivering care in the same practice but without a common framework or practice to integrate that care” (p. 411, citing Peek & the National Integration Academy Council, 2013). Gunn and Blount (2009) view co-locating as the first step toward health psychology moving from being a specialty service to being a primary care service. Collaborative care was defined earlier in this chapter. While collaboration is viewed as members of different professions working closely in care delivery, it does not imply they fall under the same organizational framework (Kuramoto, 2014). Miller et al. (2009) view co-location as a primary attribute to collaboration and integration while Boon et al. (2009, as cited in Kuramoto, 2014) consider collaboration as a precondition to integration.

Going beyond co-location and collaboration, Strosahl (2011) identified six key dimensions or components of primary and behavioral healthcare integration. These include integration of: mission, or the degree to which the two healthcare systems are aiming for the same objectives, goals, and strategies; clinical services, or the extent to which the two healthcare systems seamlessly engage in assessment, intervention, and follow-up; physical work space, or the extent to which the two healthcare systems enable instantaneous access to care by having providers work in the same space (i.e., co-location); operations, or the degree to which primary and behavioral healthcare providers share operations processes (e.g., support staff, reception, scheduling, quality improvement activities); information, or the extent to which primary and behavioral healthcare providers can access all patient care information in real time; and finances, or the extent to which primary and behavioral health services are funded as a basic form of health care. Strosahl’s key components are consistent with Oss’s (2016) view that consumers want their healthcare experience to be seamless, that is, “easy access with no wrong entry point (either vertical or physical), a single shared consumer record with consumer
external access, and shared standards of service” (p. 2). Other specific characteristics of integrated care are apparent in the discussion of models of integration that follows.

# Models of Collaborative and Integrated Service Delivery

The actual characteristics of any integrated program of primary behavioral healthcare service delivery will vary according to the organization’s characteristics, patient population, and goals it strives for from the integration (Institute for Healthcare Improvement [IHI]; 2014). The degree to which the healthcare organization provides support—financial, infrastructure, etc.—for establishing and maintaining the program is also a consideration. Several authors have provided their conceptualizations of how cooperative delivery of medical/physical and behavioral healthcare services may be delivered. Following are three ways in which relationships between primary and behavioral healthcare services can be modeled. Others views may exist, but they are likely to reflect variations to those presented later.

## Five-Level Integration Model

Doherty et al. (1996) provide a long-established model for various degrees of collaboration and integration of primary and behavioral healthcare services. The hierarchical structure of this five-level system depends on both the physical proximity of behavioral health care and other healthcare providers, and the degree to which collaboration, communication, and a shared sense of vision and treatment paradigms exist. The higher the level of a given service delivery organization or program, the more integrated are the cultures and services of its medical and behavioral healthcare providers and, consequently, the better equipped it is to handle more demanding cases.

## SAMHSA-HSRA Center for Integrated Health Solutions Model

This model (Heath, Wise Romero, & Reynolds, 2013) is essentially a modification and extension of Doherty et al.’s (1996) five-level system, consisting of six levels of patient care along a continuum of progressively increasing collaboration and integration and under an overarching framework of Blount’s (2003) three categories of care: coordinated, co-located, and integrated. It reflects experience that has accumulated and brings valuable aspects that have evolved since the publication of the Doherty et al. article. The model incorporates enhancements that (a) through its comprehensiveness, enable it to serve as a national standard of integrated care; (b) provide recognized benchmarks for organizations to assess their levels of integration; and (c) serve as a means of comparing outcomes data between levels. Table 1.3 provides a summary of this model, including a brief description of each level of collaboration/integration. Note that organizations using this model as a guide do not necessarily need to move through each of the six levels in order to achieve their goals, and achieving the highest level of integration may not necessarily be appropriate or even feasible for all organizations (IHI, 2014).

### Table 1.3 SAMHSA’s Six Levels of Collaboration and Integration

<table>
<thead>
<tr>
<th>Overarching Category of Care</th>
<th>Level</th>
<th>Work Space for Primary, Behavioral Health, and Other Healthcare Providers</th>
<th>Core Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinated</td>
<td>1</td>
<td>Separate facilities</td>
<td>Have separate systems; communicate about cases only rarely, usually under compelling circumstances and driven by provider need; may never meet in person; have limited understanding of each other’s roles</td>
</tr>
<tr>
<td>Overarching Category of Care</td>
<td>Level</td>
<td>Work Space for Primary, Behavioral Health, and Other Healthcare Providers</td>
<td>Core Description</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------</td>
<td>-------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Coordinated</td>
<td>2</td>
<td>Separate facilities</td>
<td>Have separate systems; communicate periodically about shared patients, usually driven by specific patient issues; may meet as part of a larger community; appreciate each other’s roles as resources.</td>
</tr>
<tr>
<td>Co-Located</td>
<td>3</td>
<td>Same facility, not necessarily the same offices</td>
<td>Have separate systems; communicate regularly about shared patients, by phone or email; collaborate, driven by need for each other’s services and more reliable referral; meet occasionally to discuss cases due to close proximity; feel part of a larger yet nonformal team.</td>
</tr>
<tr>
<td>Co-Located</td>
<td>4</td>
<td>Same space within same facility</td>
<td>Share some systems; communicate in person as needed; collaborate, driven by need for consultation and coordinated plans for difficult patients; have regular face-to-face interactions about some patients; have a basic understanding of roles and culture.</td>
</tr>
<tr>
<td>Integrated</td>
<td>5</td>
<td>Same space within same facility with some shared space</td>
<td>Actively seek system solutions together or develop work-arounds; communicate frequently in person; collaborate, driven by desire to be a member of the care team; have regular team meetings to discuss overall patient care and specific patient issues; have an in-depth understanding of roles and culture.</td>
</tr>
<tr>
<td>Integrated</td>
<td>6</td>
<td>Same space within same facility, sharing all practice space</td>
<td>Have resolved most or all system issues, functioning as one integrated system; communicate consistently at the system, team, and individual levels; collaborate, driven by shared concept of team care; have formal and informal meetings to support integrated model of care; have roles and cultures that blur or blend.</td>
</tr>
</tbody>
</table>


**Four Quadrant Clinical Integration Model**

This model represents a different approach from the Doherty et al. (1996) and the Heath et al. (2013) models of collaboration and integration. It is a conceptual framework for designing integrated programs that is based on notion that the severity of a patient’s behavioral and physical health needs determine what type of care he or she requires and where it should be delivered (Hogg Foundation, 2008). Accompanying this assumption is a recognition that a variety of primary and specialty care services need to be appropriate and available because the needs of individual patients vary.

Each of the four quadrants comprising this population-focused model outlines the types of services and system elements that would be needed to meet the needs of those patients whose behavioral and physical conditions meet the quadrant’s criteria (Mauer, Parks, Pollack, & Bartels, n.d., cited in Mauer, 2006). The four mixed conditions are as follows:

**Quadrant I:** Low to moderate risk/complexity for both behavioral and physical health issues.

**Quadrant II:** High behavioral health risk/complexity and low to moderate physical health risk/complexity.
Quadrant III: High physical health risk/complexity and low to moderate behavioral health risk/complexity.

Quadrant IV: High risk/complexity for both behavioral and physical health issues.

As examples of differences in system requirements for each quadrant, Quadrant I requires a PCP with standard screening tools and behavioral health practice guidelines along with a PCP-based behavioral health consultant or care manager. In contrast, Quadrant IV requires a PCP with standard screening tools and behavioral health practice guidelines, a PCP-based behavioral health consultant or care manager, specialty medical/surgical, psychiatric consultation, emergency department, medical/surgical inpatient care, nursing home/home-based care, and other community supports. Those interested in further information about the Four Quadrant Model are referred to Mauer (2003, 2006, 2009) and her colleagues (Mauer, Parks, Pollack, and Bartels, n.d., cited in Mauer, 2006) who have described and offered revisions to incorporate person-centered healthcare home concepts to the model.

Examples of Integrated Care

There are several types of healthcare organizations that function as integrated care service providers. Among the more common are patient-centered medical homes, accountable care organizations, and federally qualified health centers.

Patient-Centered Medical Home (PCMH)

Perhaps the best-known model of integrated care is that of the PCMH (Melchert, 2015). Referring to joint principles identified by four physician groups, the Patient-Centered Primary Care Collaborative (cited in Runyan, 2011; also see Holden et al., 2014; Nash, Khatri, Cubic, & Baird, 2013; Nielsen, Olayiwola, Grundy, & Grumbach, 2014; Nordal, 2012) reported several characteristics of PCMHs. These include a physician-directed medical practice with a personal physician point-of-contact for each patient; a whole-person orientation with coordinated, integrated care across all elements of the patient’s healthcare system and community; enhanced access to care that meets high quality and safety standards; and a system of payment that recognizes the value-added aspect of the PCMH. Runyan noted that two of these characteristics—integrated service delivery and whole-person orientation—are the most central points to the inclusion of behavioral health into PCMH services. Other characteristics include comprehensiveness and a team-based approach to care that is biopsychosocial in nature (McDaniel & deGruy, 2014; Melchert, 2015).

Nielsen et al. (2014) reviewed studies involving PCMH initiatives that were released between August 2012 and December 2013 and that contribute to its evidence base. They found demonstrations of reductions in the cost of care and the use of avoidable or unnecessary services, increases in preventive services, and improvements in population health indicators, access to care, and patient satisfaction with services. A need for studies regarding clinician satisfaction with PCMHs was identified. PCMHs were also recognized for their foundational role in other delivery models, such as accountable care organizations (see later) and medical neighborhoods. Nielsen et al. also noted that significant payment reforms in public and commercial health plans where PCMHs are being incorporated.

The NCQA views PCMHs as the foundation for a medical system that achieves the Triple Aim (NCQA, 2014). To assist in achieving this goal, NCQA has established a recognition program for PCMHs that demonstrate their ability to meet six standards. These standards include: patient-centered access to care, team-based care, population health management, care management and support, care coordination and care transitions, and performance measurement and quality improvement. Noteworthy is NCQA’s acknowledgment of behavioral health care as being key for better integration of care.
Accountable Care Organization (ACO)
Like PCMHs, ACOs are examples of an attempt to grow the use of integrated care models in the United States (Kelly & Coons, 2012). According to the Centers for Medicare and Medicaid Services (CMS; 2015a), ACOs are groups of doctors, hospitals, and other health care providers, who come together voluntarily to give coordinated high quality care to their . . . patients. The goal of coordinated care is to ensure that patients, especially the chronically ill, get the right care at the right time, while avoiding unnecessary duplication of services and preventing medical errors.

As suggested by CMS, Nordal (2012) notes that at the minimum, ACOs must have at least one hospital, primary care physicians, and specialty care providers. Their mission is to manage the entire continuum of care and to assume responsibility for costs and quality of care provided to the covered population of patients under collaborative agreements (Croft & Parish, 2013; Rittenhouse, Shortell, & Fisher, 2009). Moreover, the success of an ACO will be dependent on a “strong foundation of high-performing primary care” (Rittenhouse et al., 2009, p. 2302). In fact, NCQA (2014) indicates that ACOs build upon a PCMH foundation.

Like PCMHs, ACOs are designed to provide improved access to comprehensive, patient-centered primary care; facilitate partnerships between providers and patients; involve the patient’s family as appropriate; ensure that coordination of services across the continuum of care on-site or through referrals is seamless; assess quality of care and health outcomes; increase patient satisfaction; and reduce costs (Kelly & Coons, 2012). Unlike PCMHs which focus on a strong primary care foundation for its system of health care, ACOs emphasize the alignment of incentive provider accountability across all care provided within the system (Rittenhouse et al., 2009).

Federally Qualified Health Center (FQHC)
Funded by the Health Center Consolidation Act, FQHCs are nonprofit, public, community-based organizations providing comprehensive primary and preventative services in designated underserved communities (Open Minds, 2015a, 2015b). There are four general types of FQHCs: community health centers, migrant health centers, health care for the homeless programs, and public housing primary care programs. In addition to primary and preventative services, FQHCs must provide six other types of services, including mental health, substance abuse, dental, hospital, specialty care, and transportation needed for adequate care. All of the required services may be provided on-site or by arrangement through another provider. Funding comes from reimbursement from Medicaid, Medicare, and out-of-pocket sliding scale fees; Section 330 Public Health Service Act grants; and state and private grants. Auxier, Hirsh, & Warman (2013) note that FQHCs are not restricted to serving only vulnerable populations, and that over half of the populations served are minorities, over 50% live in rural areas, and approximately 70% live below the poverty line.

Benefits of Collaborative and Integrated Care
The numerous benefits of collaborative and integrated care have been identified in the literature. Among the most commonly cited benefits are improved experiences for both patients and providers (Laderman, 2015; Melchert, 2015; Nielsen, Langer, Zema, Hacker, & Grundy, 2012; Ray-Sannerud et al., 2012); the effectiveness and efficacy of interventions which is sometimes found to be superior to “usual” care (Auxier, Farley, & Seifert, 2011; Bryan et al., 2012; Kessler, 2009; Ray-Sannerud et al., 2012;
Runyan, 2011; Unützer et al., 2013; Unützer et al., 2009) and with improvement sometimes occurring rapidly (Bryan, Morrow, & Appolonio, 2009; Kessler, 2009); reduced stigma associated with seeking mental health treatment (Kearney, Post, Pomerantz, & Zeiss, 2014; Kelly & Coons, 2012; Kuramoto, 2014; Runyan, 2011); reduced emergency room visits and inpatient admissions (Grumbach & Grundy, 2010; Nielsen et al., 2012; Rozensky, 2014); improved disease-specific quality of care (IHI, 2014; Rozensky, 2014; Runyan, 2011; Scharf et al., 2014); improved quality of (Peak et al., 2014; Rozensky, 2014) and access to care (Felleman, Athenour, Ta, & Stewart, 2013; Nielsen et al.; 2014); and a positive effect on outcomes (Butler et al., 2008; IHI, 2014; Kuramoto, 2014; Nielsen et al., 2012; WHO, 2008), including long-term outcomes (Ray-Sannerud et al., 2012) and physical health outcomes (IHI, 2014).

Perhaps one of the most important benefits—and one of the major arguments in support of the integration of primary and behavioral healthcare services—is related to the reports of overall medical cost savings that occur when patients with mental health and substance abuse problems, who are seen in primary care settings, are identified early on and receive appropriate treatment. Following is a sample of the data cited in the literature supporting this contention:

- Noting that the medical costs for individuals with a chronic medical illness and a comorbid behavioral health condition can be two to three times that for those without a comorbid behavioral health condition, Melek et al. (2014) analyzed medical claims data from published literature and from proprietary and purchased data sources to determine how much of the cost of care could be reduced if these patients were treated under a model of collaborative care. The assumption is that this type of care model would likely identify previously unidentified behavioral health disorders, thus allowing the opportunity for treatment. Data analyses were used to extrapolate national cost estimates for Medicare, Medicaid, and commercial health plan populations. Based on these analyses, Melek et al. estimated that a total cost savings of $26.3–$48.3 billion through effective integration of care. Note that most of the increased costs for those with comorbid conditions were attributed to medical rather than behavioral healthcare services.

- Graham (2011) reported that the Salt Lake City-based Intermountain Health plan yielded $667 in savings per person for those who were treated in integrated care conditions versus those who were not. Graham also indicated that according to a 2006 report, integrated care provided to 370 Medicaid patients with chronic diseases and mental illnesses resulted in $754,800 in annual savings.

- In a similar investigation, Katon et al. (2008) compared the five-year medical costs of depressed diabetics with comorbid major depression or dysthymia treated by either a collaborative care intervention or usual PCP care. Even with adjustments, the mean medical costs for the collaborative care intervention patients were $3,907 lower than those for the usual care group. Improved depression outcomes were also noted for the collaborative care group.

- Blue Cross Blue Shield of Michigan reported a reduction in emergency room visits for both adult (10%) and child (13.5%) patients treated through a PCMH compared to adults (6.5%) and children (9%) treated through non-PCMHs (Sammer, 2011). Reduction in use of high-technology radiology was also greater for the PCMH group (5% vs 3%). Moreover, a 22% reduction in adult ambulatory care was noted for the PCMH adult patients.

- Another source of potential cost savings derived from depression treatment was investigated by Rost, Smith, and Dickinson (2004) using changes in work productivity and absenteeism and their associated cost savings as dependent variables. Rost et al. found that over two years, both consistently and inconsistently employed patients who received enhanced care (i.e., increased antidepressant months and specialty care counseling) versus usual care for their depression through their primary care physician or their care managers resulted in a 6.1% increase in productivity,
estimated to be worth $1,491 annually per depressed full-time employee (FTE). Moreover, absenteeism was reduced by 22.8% (i.e., 10.6 days), worth $539 annually per depressed FTE, for the enhanced treatment sample. These intervention effects were seen more in patients that were consistently employed (i.e., reported being employed full- or part-time at every six-month reassessment over 2 years). Interestingly, the intervention had little or no impact on depression severity or emotional role functioning.

In an effort to more thoroughly assess the economic efficiency of collaborative care versus usual care for the management of depressive disorders in primary care settings, Jacob et al. (2012) conducted a systematic review of a total of 30 randomized clinical trials and other studies published from 1980 to 2009. In summary, (a) only four of the seven studies that look at economic benefits of collaborative care in avoiding health care or loss of productivity found positive effects resulting from intervention while the other three studies found either minimal or no incremental benefits; (b) three of the five studies investigating both cost and benefits found lower costs for collaborative care due to increased productivity or reduced healthcare utilization; and (c) five of six cost-utility studies found collaborative care to be cost effective. Based on these and other published findings, Jacob et al. review concluded that collaborative care for the management of depression is a good economic value.

Taking a different approach, Chiles, Lambert, and Hatch (1999) performed a meta-analysis of 91 studies investigating the impact of psychological interventions on medical cost-offset. The interventions included different forms of behavioral medicine, psychotherapy, and psychiatric consultation. Based on weighted and unweighted analyses, Chiles et al. found effect size values of 0.34 and 0.34, respectively, for 40 treatment-comparison groups articles, and 0.16 and 0.24, respectively, for 17 experimental pretest/posttest only articles. Moreover, analyses of the hospital length of stay (LOS) data indicated findings of (a) 90% of the studies reporting a decrease in medical utilization following a psychological intervention, (b) an average of 15.7% utilization reduction for treatment groups in contrast to an average of 12.3% utilization increase for control groups, (c) dollar savings of $2,205 per person based on an average of reduced LOS of 2.52 days, (d) 20%–30% savings across the cost-offset articles, and (e) 31% of the 28 articles reporting dollar savings, continuing to report savings even after mental health intervention costs were subtracted from the original savings. Moreover, differences in the amount of offset were found, with greater savings being noted for females, children 18 years or younger, and adults 66 years or older.

As for others who have examined the relevant literature, Levant, House, May, and Smith (2006) concluded that the data demonstrates that the cost-offset hypothesis is robust. They also cite the work of Cummings (1999, 2002) and his colleagues (Cummings, Cummings, & Johnson, 1997) which identified factors important to the production of this the offset effect. These include the co-location of and specialized training for providers, targeted interventions, evaluation of outcomes, pervasiveness and system-wide acceptance, and targeting of the highest users of medical services. And according Blount et al.’s (2007) review of the literature,

It appears that the better targeted the behavioral health intervention is to the needs of patients with specific medical conditions (by means of behavioral medicine, care management, or behavioral health integrated care), the more medical cost savings are realized. The more generic the behavioral health intervention (outpatient psychotherapy) is, the less medical cost savings are realized.

(p. 292)

In sum, the literature is replete with examples of cost savings or cost-offset that can result from collaborative or integrated primary and behavioral healthcare services. In some cases, however, this
has not been demonstrated or the elements of integration by which this occurs may not be fully understood (see Butler et al., 2008; also see Levant et al., 2006 for criticisms of the cost-offset effect). For those interested in examining this matter further, several tabular summaries of numerous studies and projects examining the effects of integrated care on the costs and other aspects of healthcare provision are provided by Butler et al. (2008), Grumbach and Grundy (2010), Kuramoto (2014), Mauer (2006), Mauer and Jarvis (2010), Nielsen et al. (2012), and Nielsen et al. (2014).

Roles, Responsibilities, and Opportunities for Psychologists

The integrated primary and behavioral healthcare practice provides an excellent setting for clinical psychologists, properly trained to function within that type of setting, to function, flourish, and contribute to the improvement of healthcare service delivery. The psychologist’s skills and training make him or her uniquely qualified over other behavioral health professionals to assume many roles and responsibilities within the de facto setting for behavioral health service provision.

Guideline 8 of the American Psychological Association’s Guidelines for Psychological Practice in Health Care Delivery Systems (2014b) states that “Psychologists strive to provide collaborative services to the broadest range of health care settings, and to apply their expertise to the full spectrum of health issues” (p. 5). The guideline points out that in these types of settings psychologists’ roles include, but are not limited to: helping patients and families adjust to diagnoses of acute, chronic and/or life-threatening medical conditions, including assisting them with complex treatment decisions; preparing patients for invasive medical procedures; assisting patients with adherence to difficult treatment or rehabilitation regimens; assisting patients and providers with pain management; evaluating candidates for surgery or transplantation; promoting positive health behaviors, such as smoking cessation and adherence with diet and exercise regimens; providing consultation to physicians and other providers regarding effective ways to communicate with patients and families; providing developmental, psychological, or neuropsychological assessment; providing or overseeing intervention for developmental or behavioral problems; evaluating or treating psychiatric conditions; helping patients and families with access to resources; and consulting with providers in systems other than health care delivery systems. . . . Additionally, some psychologists collaborate with medical providers regarding medications or prescribe medications themselves.

(p. 5)

Although also applicable to other healthcare settings, this excerpt essentially provides a listing of potential roles, responsibilities, and opportunities that present themselves to psychologists who become members of an integrated primary and behavioral healthcare team. The types of activities that integrated care psychologists commonly engage in perhaps can be placed in one of five broad roles outlined by DiTomasso, Golden, Cahn, and Gradwell (2013) but also identified by many others in the published literature.

Clinician

Perhaps the most obvious role for psychologists in integrated care settings is that of a direct service behavioral health clinician offering both assessment and behavioral intervention services to primary care patients (APA, 2014a, 2014b; Arean & Gum, 2013; Fisher & Dickinson, 2014; McDaniel et al., 2014; Nash et al., 2013). Assessment can involve many components, including screening, clarification of a consultation question, case review with the referring clinician, clinical interview, and conceptualization, recommendations, and communication of findings (Nash et al., 2013). It is important
to recognize that the primary distinguishing ability of psychologists over other behavioral health professionals is their training and expertise in psychological test-based assessment, a skill that is particularly relevant in conducting screening and more extensive psychological evaluations, and in establishing outcomes assessment programs.

A psychologist in an integrated care setting may be called upon to provide any number of different types of direct interventions, including brief interventions (Nash et al., 2013); diagnosis and treatment of mental and behavioral problems (APA, 2014a); behavioral medicine interventions for issues such as sleep disturbance, pain management, and medication adherence (Arean & Gum, 2013); promotion of wellness and healthy behaviors (APA, 2014a, 2014b; McDaniel et al., 2014); and assisting chronic disease patients with comorbid mental health conditions (Fisher & Dickinson, 2014).

Consultant
DiTomasso et al. (2013) considers consultation as a major means by which primary and behavioral healthcare integration is achieved. As a consultant, the psychologist’s role will vary, depending on the particular primary care setting he or she finds themselves (Hunter et al., 2009). Regardless, consultation may involve patients, programs, or the practice itself (Nash et al., 2013). With regard to clinical questions, it requires that the psychologist be able to respond in a manner that is “collaborative, team oriented, actionable, and promotes peak scope of practice service delivery” (McDaniel et al. 2014, p. 426). Important to the consultation is ensuring recommendations are appropriate to the clinicians and the setting and providing timely and meaningful feedback (Nash et al., 2013).

Teacher/Supervisor
Nordal (2012) indicated that healthcare reform has made it imperative that psychologists think of their roles as being much more than direct service delivery. This is especially true for those working in integrated care settings. Here, the psychologist may be called upon to provide formal or informal training to staff on behavioral health evidence-based practices and processes (DeLeon et al., 2013), supervise pre- or postdoctoral behavioral health students and medical residents (McDaniel et al., 2014), or instruct patients and their families about mind-body interactions and expectations regarding health and illness (McDaniel & Fogarty, 2009).

Researcher
Research is an area in which psychologists are likely to exhibit greater training and expertise than other behavioral health professionals (Arean & Gum, 2013). This allows psychologists to further contribute to the overall advancement of the primary care practice through activities such as evaluations of programs and interventions, and quality improvement assessments (McDaniel & deGruy, 2014; McDaniel et al., 2014; Stancin & Perrin, 2014). Their research skills, along with their expertise in psychometrics, are especially valuable in those settings which plan to initiate ongoing treatment outcomes assessment programs.

Administrator/Leader
Through their training, psychologists have acquired a unique skill set that provides them with perspectives and expertise that enable to perform administration and management duties in both psychological and other health services tasks (APA, 2013; DiTomasso et al., 2013). In a leadership/administrative role, psychologists may promote effective communication among staff at all organizational levels and work to improve patient care systems (McDaniel et al., 2014); supervise/oversee psychological and related services and quality improvement activities, as well as serve on organizational committees (e.g., peer review, credentialing, institutional review board [APA, 2014b]); contribute to
programs aimed at practice transformation (Fisher & Dickinson, 2014); and generally help move the practice or a program therein forward (McDaniel & deGruy, 2014).

In commenting on psychologists interested in working within the evolving healthcare system, Nordal (2012) makes the following recommendation:

It is imperative that we [psychologists] think of our roles as much broader than direct service delivery, to include such things as administrative roles, program design and evaluation, population-based intervention design and evaluation, development of appropriate outcomes measures, and supervision of subdoctoral behavioral healthcare providers. We must be able to demonstrate our ability to contribute to improving the healthcare system in ways that produce more efficient, evidence-based care that results in better health outcomes.

(p. 541)

Based on the previous discussion, it is clear that Nordal’s recommendations are consistent with what psychologists wanting to work in an integrated primary and behavioral healthcare setting should prepare themselves for in order to make themselves an invaluable member of the primary care team. In addition, it is important to remember that in integrated settings, all of these roles and activities described above are carried out by the psychologist as a member of the primary care team.

More extensive discussions of the roles, responsibilities, and opportunities for psychologists in primary care settings are presented in Chapters 2 and 4 of this book. In addition, excellent discussions of core competencies of integrated care providers can be found in Cohen, Davis, Hall, Gilchrist, and Miller (2015); Hoge, Morris, Laraia, Pomerantz, and Farley (2014); and Keniman, Gilchrist, Payne-Murphy, and Miller (2015).

**Barriers and Challenges to Integration**

Implementation of any of the previously described models for primary and behavioral healthcare integration will likely require overcoming one or more barriers or challenges. The potential barriers and challenges are numerous but can be classified into one of three general types: *clinical, operational,* and *financial.*

**Clinical Barriers**

There are several clinical barriers that can be present when attempts to integrate primary and behavioral healthcare services. Many of these can be attributed to *cultural differences* between medical and behavioral health professions (Fisher & Dickinson, 2014; Gunn & Blount, 2009; Hogg Foundation, 2008; Laderman & Mate, 2014; Levant et al., 2006). Among these are differences in treatment philosophies. Where PCPs view themselves as responsible for their patients and provide treatment that tends to be action-oriented, behavioral health providers tend to be more interested in the treatment process and see patients as being at least partly responsible for change (Gunn & Blount, 2009; Fisher & Dickinson, 2014). They also view health and illness from different perspectives (Fisher & Dickinson, 2014). Each provider has undergone training separate from the other (Gunn & Blount, 2009) and speaks a “language” which the other may not fully understand (Gunn & Blount, 2009, Hogg Foundation, 2008). The manner in which confidentiality of patient information between providers is handled can also be an issue (Kuramoto, 2014; Levant et al., 2006). Moreover, despite behavioral health services being delivered in a medical setting, for some the stigma of seeking mental health care may still remain (Hogg Foundation, 2008; Levant et al., 2006).
Operational Barriers
Operational barriers to integration are numerous. Perhaps the foremost barrier of this type is the difference in how primary care and behavioral health practices function. Primary care practices tend to be hectic, high-volume, time-focused settings (Fisher & Dickinson, 2014) geared to the treatment of acute problems (Hogg Foundation, 2008). PCPs and behavioral health clinicians are most likely unfamiliar with each profession’s disciplines, hierarchy, roles, training, and treatment philosophy (Gunn & Blount, 2009), as well as with the clinical routines that take place when functioning in separated practices (Blount et al., 2007) with different workplace cultures (Kuramoto, 2014). In addition, the introduction of behavioral healthcare services and professionals into the primary care settings can present many challenges to often resistant staff members, including the changing of staff roles, expertise and comfort in dealing with behavioral health problems, disrupting the practice’s current workflow; hiring and training new staff (Butler et al., 2008; Kuramoto, 2014; Scharf et al., 2014); and even more mundane issues such as the allocation space (Gunn & Blount, 2009; Fisher & Dickinson, 2014) and modification of the practice’s electronic health record software to accommodate behavioral health information (Laderman & Mate, 2014).

Financial Barriers
The most frequently cited barrier to the establishment of integrated care is the issue of how it will be paid for (Hogg Foundation, 2008). It is also the most difficult barrier to overcome (Kuramoto, 2014). According to Butler et al. (2008), “Financial barriers are a major impediment, primarily because many activities associated with integrated care, such as many care management functions, consultations and other communication activities between providers, and telephone consultation with patients, are not traditionally reimbursed under typical fee-for-service care” (p. 3). Similar issues were reported by Gray et al. (2005), Kuramoto (2014), and Scharf et al. (2014).

There are many other aspects to the problem of financial barriers. These include other, generally nonreimbursable activities such up-front costs for hiring and training new staff members, retraining existing staff, and addressing important operational and clinical issues (IHI, 2014; Levant et al., 2006); lack of parity of reimbursement for PCP and behavioral health clinician services (Gray et al., 2005); and billing and payment system issues (Fisher & Dickinson, 2014; Gray et al., 2005; Miller, Petterson, Burke, Phillips, & Green, 2014). Patients covered under carve-out mental healthcare plans present additional challenges (Butler et al., 2008; Gray et al., 2005; Kuramoto, 2014). All in all, the financial barriers must be overcome if integration is to be sustainable (Butler et al., 2008).

Solutions to Integration Barriers
A discussion of potential solutions to overcoming the various types of barriers to integration that may arise in a particular primary care setting is beyond the scope of this chapter. Discussions in Chapters 2 and 4 of this book will address issues related to this topic. However, worth mentioning briefly here are a few actions that can be initiated that will help facilitate the integration process in most primary care practice settings. One is having someone within the organization taking a leadership role in promoting and supporting integration effort (IHI, 2014; Kuramoto, 2014). Operationalizing new workflow processes to accommodate both the medical and behavioral health components of the new enterprise (Laderman & Mate, 2014) and providing to the staffs of both care components any training needed to maximize successful integration (IHI, 2014; Kuramoto, 2014) will also be helpful. Finally, it will be important for members of both staffs to understand the roles, responsibilities, and perspectives of all team members (Fisher & Dickinson, 2014).
THE ROLE OF PSYCHOLOGICAL ASSESSMENT IN PRIMARY CARE SETTINGS

During the past several years, psychological assessment has come to be recognized for more than just its usefulness at the beginning of treatment. Its utility has been extended beyond being a mere tool for describing an individual’s current state, to a means of facilitating the treatment and understanding behavioral healthcare problems throughout and beyond an episode of care. Generally speaking, several available commercial and public-domain psychological tests can be employed as tools a) to assist in clinical decision-making activities, including screening, treatment planning, and treatment monitoring; b) for outcomes assessment for the purpose of measuring and monitoring the effects of treatment, and outcomes management; and c) more directly, used as treatment tools or techniques in and of themselves. These instruments are useful not only in mental health and substance abuse treatment settings, but also in other settings in which the need to identify and provide services for behavioral health problems exists. One such setting is the primary care setting.

As discussed in the previous section, psychologists and other trained behavioral healthcare professionals can uniquely contribute to efforts to fully integrate their services in primary care settings and other healthcare settings. Levant et al. (2001) indicate that

Psychologists’ core skills in assessment and treatment can be integrated into roles in supervision, administration, program design, program evaluation, and research. As a consequence, psychologists are uniquely positioned to assume a greater role in the management of health and disease. . . . Further, psychologists’ strong research background—a unique qualification among health care professionals—prepares psychologists to play key roles in the design, implementation, and evaluation of prevention and intervention programs at the individual, system, and community level.

Levant et al.’s (2001) opinion is consistent with this author’s opinion that one of the most significant contributions that psychologists can make to the integration of medical and behavioral health care is through the establishment and use of test-based psychological assessment services. Information obtained from psychometrically sound self-report tests and other-report instruments (e.g., clinician rating scales, parent-completed instruments) can assist the primary care provider in several types of clinical decision-making activities, including screening for the presence of mental health or substance abuse problems, planning a course of treatment, and monitoring patient progress. Testing can also be used to measure the outcomes of the treatment that has been provided to patients with mental health or substance abuse problems, thus assisting in determining what works for whom.

Following is an overview of considerations related to psychological assessment procedures and instruments and the potential roles that psychological assessment can play in primary care settings. Note that how and for what purpose psychological testing is applied in primary care settings is not necessarily dependent on the type or degree of behavioral healthcare integration that exists therein. Nor is its use necessarily dependent on who among the provider team administers the test or applies the results. With proper training or under the supervision of a psychologist or other qualified professional formally trained and experienced in the use of these instruments, primary care providers can appropriately administer most psychological test instruments and employ the results in their offices and facilities.
General Considerations

The introduction of psychological testing in primary care settings, either alone or as part of a package of behavioral healthcare services, can greatly enhance the diagnostic and treatment options offered by primary care providers. However, there are a few very important considerations to be mindful of if testing is to be fully accepted as a part of the services offered by primary care providers or provider organizations.

Psychological Assessment versus Psychological Testing

The focus of this book is on psychological assessment. This term is frequently and mistakenly used interchangeably with the term psychological testing when, in fact, the two terms mean different things. According to the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999), while psychological testing is defined as “any procedure that involves the use of tests or inventories to assess particular psychological characteristics of an individual” (p. 180), psychological assessment serves a different purpose and involves much more. Thus, it is defined as

A comprehensive examination of psychological functioning that involves the collecting, evaluating, and integrating test results and collateral information, and reporting information about an individual. Various methods may be used to acquire information during the psychological assessment: administering, scoring and interpreting tests and inventories; client and third-party interviews; analysis of prior educational, occupational, medical, and psychological records.

(p. 180)

The distinction between these two types of evaluation is not an unimportant one. Not only does it draw attention to the different levels of training, skills, and required effort that are involved in these two clinical services, but it also demonstrates that psychological tests results should be viewed as just one source of patient-related information to be integrated with information gathered from other information sources in order to formulate a good understanding of the patient and the complexity of his or her problems. It is this ability that most distinguishes psychology from other behavioral health professions (Krishnamurthy et al., 2004).

The Biopsychosocial Approach to Psychological Assessment

The approach to assessing an individual patient can vary according to any of several factors. As integrated care has adopted a biopsychosocial view of health and illness (McDaniel et al., 2014), it follows that assessment within these primary care settings is conducted within this framework. Addressing assessment within the biopsychosocial model, Sweet, Tovian, Breting, and Suchy (2013) state that this approach to assessment requires going beyond the framework of the mind-body duality that characterizes the biomedical model to one in which psychosocial factors (e.g., culture, environment, behaviors, beliefs) are taken into consideration along with the patient’s presenting symptomatology. Thus, this approach attempts to assess the interaction of the type of data (e.g., physiological, cognitive, behavioral, affective) with the source of that data, whether it be the patient, his or her environment, or both. In doing so, it helps to provide a broader understanding of the patient.
Assessment Skills and Competencies Psychologists in Primary Care Settings Must Have

First and foremost, all psychologists involved in assessment, including primary care psychologists, should have a solid grounding in all important aspects of psychological assessment. Eight essential core competencies related to psychological assessment for all health service psychologist practitioners were identified by the psychological assessment work group at the 2002 Competencies Conference: Future Directions in Education and Credentialing in Professional Psychology, reported by Krishnamurthy et al. (2004). These included basic psychometric theory, various bases of psychological assessment (e.g., scientific, contextual), assessment skills and techniques, outcomes assessment skills, the ability to evaluate the roles and contexts within which psychologists and patients function and how they affect the assessment activity, an understanding of collaborative professional relationships, an understanding of the relationship between assessment and intervention, and technical skills (e.g., case conceptualization, integration of information).

In 2014, the American Psychologist published an article outlining the outcomes of an American Psychological Association presidential initiative to identify those competencies determined to be necessary for the practice of primary care psychology (McDaniel et al., 2014). These are competencies that the Interorganizational Work Group on Competencies for Primary Care Psychology Practice indicated all primary care psychologists should have expertise in or at least be familiar with. Among the competencies identified under the Application competency cluster were those related to assessment, including:

1. Selecting and implementing screening methods (using evidence-based assessment measures) to identify patients who are either at risk or requiring specialized services.
2. Ensuring the maintenance of test integrity whenever psychological assessments are used.
3. Using assessment instruments focused on the patient’s current functioning while incorporating physical, behavioral, and psychological aspects of health and well-being.
4. Rapidly determining the patient’s needs and how soon he or she needs to be scheduled for an appointment.
5. Identifying behavioral risk factors pertinent to the patient.
6. Incorporating input from significant others into the assessment process.
7. Evaluating and using the strengths of the patient’s family, community, and others to better understand the patient’s needs and to promote their health.
8. Monitoring the patient’s status over time to determine changes in the presenting problem and the effectiveness of the prescribed treatments.

Runyan (2011) also identified some of these same assessment-related core competencies as being helpful for becoming an effective practitioner in primary care settings.

Nash et al. (2013) presented a list of essential competencies for psychologists working in integrated primary care settings. These eight competency groupings, adapted from the Cube Model (Rodolfa et al. 2005) and the Competency Benchmarks Document (Fouad et al., 2009), includes an Assessment subcategory. Under this subcategory competencies essential for psychologists are identified in the following areas: (a) screening for mental illness, (b) clarification of the consultation referral question from the PCP, (c) performing a clinical case review, (d) conducting a clinical interview with the patient, and (e) conceptualizing the problem, formulating recommendations, and communicating the findings to the patient and the primary care team.

Finally, as part of the NCQA’s 2014 standards and guidelines for Patient-Centered Medical Home Recognition (2014), PCMHs are required to collect and update comprehensive health assessments
on their patients. Among the required information in the health assessment are behaviors affecting health, patient and family mental health and substance abuse histories, both depression and developmental screening using a standardized instrument, and the assessment of health literacy. Because of their training, psychologists can play a key role in ensuring that these requirements are met by those PCMHs seeking NCQA recognition.

Comprehensive versus Brief Assessment
Doing what it takes to function effectively in fast-paced primary care settings is what some would consider antithetical to how most healthcare psychologists have been trained, especially when it comes to assessing patients in those settings. The 15 to 30 minutes often allotted for evaluation/assessment (Gatchel & Oordt, 2003; Kelly & Coons, 2012; Nash et al., 2013) is frequently not enough time to perform what most psychologists would consider a comprehensive assessment of a patient’s problems, much less provide for some form of intervention. As noted by Hunter et al. (2009), sometimes it may not be practical to conduct a thorough assessment; however, administering a screener or conducting another type of brief assessment procedure may be sufficient for adequate disposition of the case at hand.

When time is issue, Gatchel and Oordt (2003) recommend conducting an assessment that is focused on the referral question or primary problem. They recommend using broad, closed-end questions covering general areas, followed up with more in-depth questioning in those areas yielding significant findings. Questions with the greatest predictive power should be used, and other areas need not undergo comprehensive assessment or screening unless the patient’s condition or circumstances places them at risk for some untoward event or condition (e.g., risk of harm to self or others). As Gatchel and Oordt noted, “For most patients . . . a focused assessment will be sufficient to identify target problems and make genuinely helpful recommendations, or will be adequate to recognize that more definitive care is indicated” (p. 13). Referral to a higher level of care thus can be made when indicated.

Instrumentation and Procedures for Behavioral Healthcare Assessment
The psychological test instrumentation and procedures required for any assessment—whether in primary medical care, behavioral health care, or other settings—will depend on (a) the general purpose for which the assessment is being conducted and (b) the level of information that is required for that purpose. Generally, most psychological test instrumentation and procedures that would serve the purpose of psychological/behavioral healthcare assessment in a primary care setting falls into one of several general categories.

Clinical Interview
At the core of any psychological assessment should be the clinical interview. In primary care or any other treatment setting, its primary purpose is to obtain information relevant to the problem(s) for which the patient is seeking treatment (Aklin & Turner, 2006). Findings from psychological testing, a review of medical and other pertinent records of historical value that are available, collateral contacts, and other sources of information about the patient all are important and can help clinicians understand patients and their problems. However, nothing can substitute for the type of information that can only be obtained through face-to-face contact with the patient. As Groth-Marnat (2009) stated,

Probably the single most important means of data collection during psychological evaluation is the assessment interview. Without interview data, most psychological test results are meaningless [emphasis added].
The interview also provides potentially valuable information that may be otherwise unobtainable, such as behavioral observations, idiosyncratic features of the client, and the person’s reaction to his or her current life situation. In addition, interviews are the primary means for developing rapport.

(p. 65)

The clinical interview in integrated care settings is discussed in detail in Chapter 8 of this book.

Psychological/Psychiatric Symptom Measures

Probably the most frequently used instrumentation for screening and treatment planning, monitoring, and outcomes assessment are measures of psychopathological symptomatology. These measures also are the type of instruments on which the majority of the clinical psychologist’s psychological assessment training was likely focused. These instruments were developed to assess behavioral health problems that typically prompt people to seek treatment.

There are several subtypes of these measures of psychological/psychiatric symptomatology. The first is the comprehensive multidimensional measure. This is typically a lengthy, multiscale, standardized instrument that measures and provides a graphical profile of the patient on several psychopathological symptom domains (e.g., anxiety, depression) or disorders (schizophrenia, antisocial personality). Also, summary indices sometimes are available to provide a more global picture of the individual with regard to his or her overall psychological status or level of distress. Probably the most widely used and/or recognized of these multidimensional measures are the restandardized version of the Minnesota Multiphasic Personality Inventory, the MMPI-2 (Butcher et al., 2001) and the most recent revision of the MMPI, the Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008; Tellegen & Ben-Porath, 2008).

Multidimensional instruments can serve a variety of purposes that facilitate therapeutic interventions in primary and behavioral healthcare settings. They may be used upon initial patient contact to screen for the need for service and, at the same time, yield information that is useful for treatment planning. These instruments might also be useful in identifying specific problems that may be unrelated to the patient’s chief complaints (e.g., low self-esteem). In addition, they generally can be administered numerous times during the course of treatment to monitor the patient’s progress toward achieving established goals and to assist in determining what adjustments (if any) must be made to the clinician’s approach. Moreover, use of such instruments pre- and posttreatment can provide individual treatment outcomes data and, at the same time, can be analyzed with the results of other patients to evaluate the effectiveness of an individual provider, a particular therapeutic approach or program, or an organization.

Abbreviated multidimensional measures are quite similar to the MMPI-2/MMPI-2-RF and other comprehensive multidimensional measures in many respects. First, by definition, they contain multiple scales for measuring a variety of symptom domains and disorders. They also may allow for the derivation of an index that can indicate the patient’s general level of psychopathology or distress. In addition, they may be used for screening, treatment planning and monitoring, and outcomes assessment purposes just like the more comprehensive instruments. The distinguishing feature of the abbreviated instrument is, of course, its length: by definition, these instruments are relatively short and easy to administer and score. Their brevity does not allow for an in-depth assessment of the patient and his or her problems, but this is not what these instruments were designed to do. Probably the most widely used of these brief instruments are Derogatis’ family of symptom checklists. These include the revision of the original Symptom Checklist-90, the SCL-90-R (Derogatis, 1994); the shortened version of the SCL-90-R, the Brief Screening Inventory (BSI; Derogatis, 1993); and the shortened version of the BSI, the BSI 18 (Derogatis, 2001). These instruments contain checklists of 90, 53, and 18
psychological symptoms, respectively. The SCL-90-R and BSI score on each of the instruments’ nine symptom domain scales. Each instrument also has three summary indices. The BSI 18 scores on three of the SCL-90-R/BSI symptom domain scales and has one summary index. The SCL family of instruments is discussed in detail in Chapter 22.

The major strength of the abbreviated multiscale instruments is their ability to quickly and broadly survey psychological symptom domains and disorders. Their value is most clearly evident in settings—such as primary care settings—where both the time and dollars available for assessment services are limited. These instruments provide a lot of information quickly and are much more likely to be completed by patients than their lengthier counterparts. This last point is particularly important if one is interested in monitoring treatment or assessing outcomes, both of which require at least two or more assessments of a patient to obtain the necessary information.

Finally, there are the disorder-specific measures which are designed to measure one specific disorder or family of disorders (e.g., anxiety, depression, substance abuse, suicidality). These instruments are usually brief, requiring less than five minutes to complete. Examples include the Beck Depression Inventory (2nd ed.; BDI-II) and the Beck Anxiety Inventory (BAI) which are discussed in detail in Chapters 9 and 12, respectively.

Measures of Neuropsychological Functioning

Neuropsychological tests are designed to provide information related to the presence and/or degree of cognitive impairment resulting from brain disease, disorder, or trauma. According to the APA (2012), “Neuropsychological evaluation and testing remain the most effect differential diagnostic methods in discriminating pathophysiological dementia from age-related cognitive decline, cognitive difficulties that are depression-related, and other related disorders” (p. 2). The results of these tests can also be used to draw inferences about the extent to which this impairment interferes with the patient’s daily functioning. There are many psychometrically sound tests of neuropsychological functioning. Some instruments assess only specific areas of functioning (e.g., immediate verbal memory); others assess broader areas of functioning (e.g., a battery of memorial measures, such as the Wechsler Memory Scale (4th ed.; WMS-IV), assessing immediate, intermediate, and remote verbal and nonverbal memory); and others are part of a battery of measures that provide a more “comprehensive” assessment of neuropsychological functioning (e.g., batteries that include tests of memory, language, academic skills, abstract thinking, nonverbal auditory perception, sensorimotor skills, etc.).

Werthman (1995) suggests that neuropsychological testing can make a significant contribution in the primary care setting by assisting primary care providers in accurately assessing or ruling out neuropsychological impairments (e.g., dementias), particularly if there is a comorbid DSM-5 disorder (American Psychiatric Association, 2013). Accordingly, Werthman specifically points to the utility of specific tests or parts of tests, rather than entire test batteries, as being the most appropriate for meeting the needs of managed care organizations. (This also holds true for primary care practices.) Werthman also suggests that the role of the psychologist should only be to “determine . . . which of the neuropsychological tests is appropriate . . . to assist the primary provider in managing the case” (p. 15). However, it is this author’s opinion that the use of neuropsychological instruments—whether they be ability-specific tests or complete batteries of tests—in primary care and other settings requires training and experience that necessitates psychologist involvement that goes well beyond offering advice.

Measures of Health Status

During the past decade, there has been an increasing interest in patient self-assessment of health status, role functioning, and health-related quality of life (HRQOL) in medical and behavioral healthcare
delivery systems. Measures of health status and physical functioning can be classified into one of two groups: generic and condition-specific. Whereas generic measures of health status typically measure aspects of health, well-being and functioning in general, condition-specific measures evaluate aspects of functioning as they relate to or are affected by a specific disease or condition. Arguably the most widely used and respected generic health status measure is the 36-item SF-36v2 Health Survey (SF-36v2; see Maruish, 2011). The SF-36v2 assesses eight dimensions of health—four addressing mental health-related constructs and four addressing physical health-related constructs—that reflect the WHO concept of “health.” An abbreviated, 12-item version of the SF-36v2 also has been developed. The SF-12v2 Health Survey (SF-12v2; see Maruish, 2012b) was developed for use in large-scale, population-based research where broadly monitoring health status is all that is required. Important data that has been gathered in support of the use of both of the SF instruments are presented in Chapter 20.

Condition-specific health status and functioning measures have been utilized for a number of years. Most have been developed for use with medical rather than behavioral health disorders, diseases, or conditions. For example, the Asthma Control Test (ACT; Kosinski, Bayliss, Turner-Bowker, & Fortin, 2004) is a condition-specific health status instrument that asks the respondent about how often their asthma interferes with their sleep and daily functioning as well as how well they think their asthma is being controlled. However, condition-specific measures of behavioral health status and functioning now are beginning to appear.

Disease Impact Measures
As the category’s name implies, disease impact measures are designed to measure the impact of specific diseases, disorders, or health conditions on an individual. Aspects of a person’s life that may be measured may include such things as emotional and social functioning, as well as role functioning at work or at home. For example, the Pain Impact Questionnaire-Revised (PIQ-R; see Maruish, 2012a) assesses the impact of pain on a patient’s work, leisure activities, and emotional well-being, as well as the patient’s reported pain severity. Another example is the Asthma Impact Scale (Kosinski, Turner-Bowker, Bayliss, & Fortin, 2003).

Service Satisfaction Measures
Given the expanding interest in assessing the treatment outcomes for the patient, it is not surprising to see an accompanying interest in assessing the patient’s (and in some instances, the patient’s family’s) satisfaction with the services received. Satisfaction should be considered a measure of the overall treatment process, encompassing the patient’s (and at times, others’) view of how the service was delivered, the capabilities and attentiveness of the service provider, the perceived benefits of the service (if any), and any of a number of other selected aspects of the service the patient received. Patient satisfaction surveys don’t answer the question, “What was the result of the treatment rendered to the patient?”; they do answer the question, “How did the patient feel about the treatment he or she received?” Thus, they serve an important program evaluation/improvement function.

Psychological Assessment as a Tool for Screening
Among the most significant ways in which the integration of behavioral health services into a primary care setting can contribute to an effective healthcare delivery system is through the enabling of quick identification of individuals in need of behavioral health or substance abuse treatment. In the field of psychology, the most efficient and thoroughly investigated screening procedures involve the use of self-report psychological test instruments. The power or utility of a psychological screening test or procedure lies in its ability to determine, with a high level of confidence, whether or not the
respondent likely has a particular disorder or condition and the degree of its severity, and/or whether or not he or she is a member of a population with clearly defined characteristics. The most commonly used psychological screeners in daily clinical practice are those designed to identify the presence of some specific type of psychological disturbance or some specific aspect of psychological dysfunction, or to provide a broad overview of the respondent’s point-in-time mental status or level of distress. The 9-item Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001), the 7-item Generalized Anxiety Disorder Scale (GAD-7; Spitzer, Kroenke, Williams, & Lowe, 2006), and the Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993) are among the most commonly employed behavioral health screeners in primary care settings.

Implementation of Screeners into the Daily Flow of Service Delivery

The utility of a screening instrument is only as good as the degree to which it can be integrated into a primary care setting’s daily regimen of service delivery. This, in turn, depends on a number of factors. The first is the ease and speed of administering and scoring of the screener, and the amount of time required to train the provider’s staff to successfully incorporate the screener into their daily work flow. As Pollard, Margolis, Niemiec, Salas, and Aatre (2013) noted, PCPs are likely to be unwilling to incorporate any procedures that will disrupt the flow of their practice.

The second factor relates to the instrument’s use. Generally, screeners are developed to assist in determining the likelihood that the patient does or does not have the specific condition or characteristic the instrument is designed to identify. Use for any other purpose (e.g., assigning a diagnosis based solely on screener results, determining the likelihood of the presence of other conditions or characteristics) only serves to undermine the integrity of the instrument in the eyes of staff, payers, and other stakeholders with a vested interest in the screening process.

The third factor has to do with the ability of the provider to act on the information obtained from the screener. In order for the screening procedure to be effective, a system of further assessment to arrive at an accurate diagnosis, means of providing an appropriate intervention (either in-house or by referral to a specialty behavioral health professional) when called for, and subsequent follow-up must be in place (for example, see Deneke, Schultz, & Fluent, 2015; Gray et al., 2005; Hogg Foundation, 2008; Holden et al., 2014; Kessler, 2009; Narayana & Wong, 2015; Poleshuck & Woods, 2014; Stepleman et al., 2014; US Preventive Services Task Force [USPSTF], 2009a). It also must be clear to the clinician how he or she should proceed based on the information available.

Several factors may be taken into consideration in determining whether screening can be determined to be “effective.” Narayana and Wong (2015) considers screening to be effective if the following conditions are present:

1. The problem for which the patient is being screened should be significantly burdensome in the population.
2. A screening test or procedure with high sensitivity and specificity to the condition must be available.
3. The problem for which the patient is being screened should be identified at a stage in which it is treatable or when it is more treatable than at a later stage.
4. The screening and ensuing treatment must have benefits that are clinically meaningful, outweigh potential harm to the patient, and can be provided at an acceptable cost. Here, potential harms include the stigma, psychological effects, and the unnecessary treatment and medication side effects that might accompany false positive screening results.
5. The final factor is staff acceptance and commitment to the screening process. This comes only with a clear understanding of the importance of the screening, the usefulness of the obtained
information, and how the screening process is to be incorporated into the organization’s daily business flow. In general, this should not be a problem. Nash et al. (2013) indicate that PCMH team members who value objective data often appreciate quantified information yielded by screeners and other behavioral health assessment instruments.

Somewhat similar to the sequence described by Derogatis and Culpepper (2004), positive findings would lead to a second level of testing. Here, another screener that meets the same requirements as those for the first screener and that affirms or rules out a diagnosis would be administered. Positive findings would lead to confirmation of screener findings by a qualified psychologist or physician. Auxier et al. (2011) provide an example of how they have incorporated this procedure in their organization. Also, they recommend that groups of patients (e.g., grouped by symptoms or medical diagnosis) be prioritized so the psychologist or responsible staff knows which patients should be screened first when it is not possible to screen every patient.

Other Considerations for Screening
Screening that involves the use of disorder-specific instruments is an acceptable approach in settings where screening instruments are administered only to those who are suspected of having a specific disorder (e.g., depression), or if effort is being made to identify only those with one or only a few types of disorders. However, in both these and other situations, there are others that recommend a different approach. As Sperry, Brill, Howard, and Grissom (1996) pointed out,

A serious limitation of all of these [disorder-specific] scales is their focus on only one type of pathology. Studies have demonstrated high comorbidity among psychiatric disorders, and this also suggests that a focus on any single disorder is inappropriately narrow. Screening programs for psychiatric disorders in primary care settings should be broad-based, and not limited to any single disorder.

(p. 206)

Similarly, in addressing the issue of screening for depression in primary care settings, Zimmerman et al. (1994) noted that

Coexistence of depression with other forms of pathology seems to be the rule, not the exception. We would predict the same to be true of other disorders as well. . . . Consequently, even if the clinical evaluation that follows the screening questionnaire is limited to those disorders that are positive on the questionnaire, it will nevertheless often need to cover multiple psychiatric disorders. Proper case finding in primary care . . . requires attention to a range of illnesses.

(p. 394)

Narayana and Wong (2015) have noted that the comorbidity of some psychiatric disorders, such as depression and anxiety, supports the screening for both disorders at the same time, or for the screening of one disorder if the other disorder is diagnosed.

Another important consideration related to the use of psychological screeners is their level of sensitivity, or the degree to which those taking the test are identified as having the problem(s) the instrument was designed to detect. The more sensitive the instrument is, the greater the number of test takers who will be mistakenly identified as having the assessed problem. These false positives will lead to unnecessary work and costs for the provider (Locke & Larsson, 1997). At the same time, instruments with lower sensitivity (and thus, higher specificity) will yield more false negatives that may result in the type of increased work and medical costs that some of the literature has associated with
undetected behavioral health problems. Thus, unless one incorporates something like the two-staged screening described earlier and in Chapter 6 of this book, one must decide which type of identification error is more acceptable, and then carefully evaluate potential screening instruments for the desired psychometric characteristics.

It is important to recognize some important aspects regarding screening for depression, one of the most common behavioral health disorders presenting in primary care settings. The USPSTF (2009; as cited in Stepleman et al., 2014) indicated that the majority of those with a positive screen for depression will not meet the criteria for a major depressive disorder. Individuals with positive screening results who are later determined to have a diagnosis of major depressive disorder can range from 12% to 50%. At the same time, Williams, Pignone, Ramirez, and Stellato (2002) found in their review of the literature of 16 case-finding depression instruments (comprised of 1 to 30 items) that they had a mean sensitivity of 85% and mean specificity of 74% for major depression, and an overall sensitivity and specificity of 79% and 75%, respectively, for major depression or dysthymia. These somewhat contradictory findings suggest that in employing screeners—in this case, depression screeners—one must ensure that the instrument selected meets the needs of the particular setting one practices in, in terms of the percentages of false positives and false negatives that are acceptable.

Given the associated cost and the potential stigma that could accompany screening, as well as the potential for misdiagnosis and the negative consequences that may ensue (see Narayana & Wong, 2015), one must ask which of the patients being seen in an integrated practice setting should be screened for behavioral health problems. The opinions on this matter are varied. For example, some would argue that maybe all patients should be screened for depression while others recommend screening only when adequate, staff-assisted depression care follow-up supports and effective treatment are available (see Miller et al. [2009] and USPSTF [2009b] for discussions of this). In addition, Narayana and Wong (2015) point out that there is a high rate of depression associated with certain medical comorbidities (e.g., cardiovascular) and social risk factors (e.g., being a veteran), with similar associations also being found with anxiety. They also note the common comorbidity of anxiety with depression and vice versa. Thus, as alluded to earlier, they suggest that screening for depression and/or anxiety may be appropriate when one of these medical or psychiatric comorbid conditions is present. In the end, individual practices will have to determine the optimal solution to the issue of how screening for behavioral health problems should be conducted based on a number of factors (e.g., cost, available staff resources).

**Psychological Assessment as a Tool for Treatment Planning**

Problem identification through the use of screening instruments is only one way in which psychological assessment can facilitate the treatment of behavioral health problems in primary care settings. When employed by a trained clinician, psychological assessment also can provide information that can greatly facilitate and enhance the planning of a specific therapeutic intervention for the individual patient. It is through the implementation of a tailored treatment plan that the patient’s chances of problem resolution are maximized.

The role that psychological assessment can play in planning a course of treatment for behavioral healthcare problems can be significant. Butcher (1990) indicated that information available from instruments such as the MMPI-2 not only can assist in identifying problems and in establishing communication with the patient, it also can help ensure that the plan for treatment is consistent with the patient’s personality and external resources. In addition, when using certain multidimensional tests, the assessment may reveal potential obstacles to therapy, areas of potential growth, and problems that the patient may not be consciously aware of. Moreover, both Butcher and Appelbaum (1990)
viewed testing as a means of quickly obtaining a second opinion. Other benefits of psychological assessment identified by Appelbaum include assistance in identifying patient strengths, weaknesses and the complexity of the patient’s personality, and in establishing a reference point during the therapeutic episode.

There are several ways in which psychological assessment can assist in the planning of behavioral healthcare treatment for patients seen in primary care settings. The more common and evident contributions can be organized around problem identification and clarification and identification of important patient characteristics.

Problem Identification and Clarification
Perhaps the most common use of psychological assessment in the service of treatment planning is for problem identification. The value of psychological testing becomes apparent in those cases where the patient is hesitant or unable to identify the nature of his or her problems—a common occurrence in primary care settings. With a motivated and engaged patient who responds honestly to items on a reliable and well-validated test, the process of identifying what led the patient to seek treatment from a primary care provider may be greatly facilitated. Cooperation shown during testing may be attributable to the nonthreatening nature of questions presented on a paper form or a computer monitor (as opposed to those posed by another human being); the subtle, indirect qualities of the questions themselves (compared to those asked by the clinician); or a combination of these reasons. In addition, the nature of some of the more commonly used psychological test instruments allows for the identification of secondary, but potentially significant, problems that might otherwise be overlooked, even with the most forthcoming of patients.

Note that the type of problem identification and clarification described here is different from that conducted during screening (see earlier). Whereas screening is more focused on determining the presence or absence of one or more specific problems, problem identification and clarification generally take a more in-depth approach to specifying the problem(s) the patient is experiencing. At the same time, there also is an attempt to determine, to a greater degree of certainty, problem severity and complexity as well as the extent to which the problem area(s) affect the patient’s ability to function. Information gained from testing can enhance both the patient’s and clinician’s understanding of the problem and lead to the development of the most appropriate treatment plan.

Identification of Important Patient Characteristics
The identification and clarification of the patient’s problems is of key importance in planning a course of treatment. However, there are numerous other types of patient information not specific to the identified problem(s) that may be useful in planning treatment and are easily identified through the use of psychological tests (particularly tests that are multidimensional in nature). Treatment plans may be developed or modified with consideration of at least some of these nonpathological characteristics, including such things as readiness to change, ego strength, and social comfort. The exceptions are generally found with clinicians or programs that take a “one size fits all” approach to treatment.

Psychological Assessment as a Tool for Treatment Monitoring
Patient assessment is important throughout the course of behavioral health treatment (Hunter et al., 2009) although this has not been the norm for behavioral health patients being seen in primary care settings. As primary care practices have traditionally been designed to treat and manage acute problems and not monitor treatment and adjust care over time when needed, those patients with longer term problems would have a tendency to “fall through the cracks” (Hogg Foundation, 2008).
However, the increasing integration of behavioral health services into primary care settings has been accompanied by an increasing interest in monitoring the treatment progress of behavioral health patients being seen in these settings (Bryan et al., 2014; also see Unützer et al., 2013). This allows the clinician an opportunity to determine if the treatment being provided is working or requires some form of adjustment, thus enhancing behavioral health outcomes.

The general process of treatment monitoring begins by assessing the patient at the beginning of treatment. The screening process may serve as the source of the baseline data against which data obtained at other points during treatment can be compared. (This assumes, of course, that the screening measure is what one wants to use to monitor treatment progress.) Once baseline data is obtained, those organizations with the necessary information and data analytic resources in place can generate an expected recovery curve for the patient. This curve, based on the progress of other patients with similar characteristics, will enable the clinician to determine if the patient is on the expected track of recovery through the episode of care. Deviations noted on remeasurement should lead the clinician to consider modifying this treatment strategy. Providers without the required technical resources will need to rely on their own clinical experience and judgment to determine if recovery is progressing as expected.

**Monitoring Treatment Progress**

Information from repeated testing during the treatment process can help the clinician determine if the treatment plan is appropriate for the patient at a given point in time. Thus, primary care providers can use psychological assessment information to determine whether their patients are showing the expected improvement as treatment progresses. If not, adjustments can be made. These adjustments may reflect the need for: (a) more intensive or aggressive treatment (e.g., increased medication dosage, referral for inpatient treatment); (b) less intensive treatment (e.g., reduction or discontinuation of medication, transfer from inpatient to outpatient care; Poleshuck & Woods, 2014); or (c) a different therapeutic approach (e.g., changing from a humanistic approach to a cognitive-behavioral approach). Regardless, any modifications in treatment require later reassessment to determine if the treatment revisions have impacted patient progress in the expected direction. This process may be repeated any number of times during treatment, ultimately providing information important for determining when to appropriately terminate treatment.

The work of Lambert and his colleagues (for example see Lambert, 2010; Lambert, Harmon, Slade, Whipple, & Hawkins, 2005; Lambert, Whipple, Smart, et al., 2001; Lambert, Whipple, Vermeesch, et al., 2001; Shimokawa, Lambert, & Smart, D. W., 2010) lends support for benefits accruing from the use of assessment-based feedback provided to clinicians during treatment. Summarizing a review of four controlled studies examining the effect of informing clinicians on patients’ treatment response on treatment outcomes, Lambert et al. (2005, p. 165) stated that “the collective results suggest that measuring, monitoring, and predicting treatment failure (feedback) enhance treatment outcomes for patients who have a negative response.” At the same time, Lambert, Hansen, and Finch (2001) point out that in order to be effective, the feedback needs to be timely and provide information that is action-oriented. Also, provision of clinical support tools (e.g., a diagnostic decision tree, a list possible interventions, supplemental measures, a tracking form) with the feedback can increase its effectiveness (Whipple et al., 2003).

**Monitoring Change**

Methods for determining if statistically and clinically significant change has occurred from one point in time to another have been developed and can be used for treatment monitoring. Many of these methods are the same as those that can be used for outcomes assessment. In addition, the reader is
also referred to an excellent discussion of analyzing individual and group change data in Newman and Dakof (1999) and Newman and Tejeda (1999).

Whether psychological test data obtained from treatment monitoring are used as fodder for generating complex statistical predictions or for simple point-in-time comparisons, the work of Howard, Lambert, and their colleagues demonstrate that psychological test data obtained for treatment monitoring can provide an empirically based means of determining the effectiveness of mental health and substance abuse treatment during an episode of care. Their value lies in their ability to support ongoing treatment decisions that must be made using objective means. Consequently, they can help improve patient care while supporting efforts to demonstrate accountability to the patient and interested third parties.

A more detailed discussion of treatment monitoring is presented in Chapter 7 of this book.

Psychological Assessment as a Tool for Outcomes Management

In the 1990s, the behavioral healthcare field began to witness accelerating growth in the level of interest and development of outcomes assessment programs (see Bobbitt, Cate, Beardsley, Azocar, & McCulloch, 2012; Maruish, 2013a). Now, much of the accountability that is built into the ACA involves tracking clinical outcomes (Rozensky, 2012). As this author has noted previously, the interest in and necessity for outcomes measurement and accountability in this era of healthcare delivery transformation provides a unique opportunity for psychologists to use their training and skills in assessment (Maruish, 2002, 2004). However, the extent to which psychologists and other trained professionals become a key and successful contributor to an organization’s outcomes initiative (whatever that might be) will depend on their understanding what “outcomes” and their measurement and applications are all about.

Before discussing outcomes, it is important to have a clear understanding of what is meant by the term. Experience has shown that its meaning varies depending on the source. From the most basic perspective, outcomes can be defined as “the intermediate or final result of either the natural course of a disorder (that is, how an individual would do in the absence of any treatment) or the processes of medical or other interventions” (McGlynn, 1996, p. 20).

However, outcomes are probably best understood as just one component of quality of care. Donabedian (1985) has identified three dimensions of quality of care. The first is structure. This refers to various aspects of the organization providing the care, including how the organization is “organized,” the physical facilities and equipment, and the number and professional qualifications of its staff. Process refers to the specific types of services that are provided to a given patient (or group of patients) during a specific episode of care. These might include various tests and assessments (e.g., psychological tests, lab tests, magnetic resonance imaging), therapeutic interventions (e.g., group psychotherapy, medication), and discharge planning activities. Processes that address treatment complications (e.g., drug reactions) also are included here. Outcomes, on the other hand, refer to the results of the specific treatment that was rendered. These results can include any number of relevant variables to stakeholders in the patient’s care. For example, Seid, Varni, and Jacobs (2000) have identified three general categories of outcomes: clinical (e.g., signs, symptoms, indicators of disorders), financial (e.g., costs saved), and patient-based (e.g., patient perceptions of quality of life).

The outcomes, or results, of treatment should not imply a change in only a single aspect of functioning. This is why the plural of the term is used throughout this chapter. Treatment may impact
multiple facets of a patient’s life. Stewart and Ware (1992) have identified five broad aspects of general health status: physical health, mental health, social functioning, role functioning, and general health perception. Treatment may affect each of these aspects of health in different ways, depending on the disease or disorder being treated and the effectiveness of the treatment rendered. Some specific aspects of functioning related to these five areas of general health status that are commonly measured include feelings of well-being, psychological symptom status, use of alcohol and other drugs, functioning on the job or at school, marital/family relationships, utilization of healthcare services, and ability to cope. Cella and Stone (2015) point to the importance of HRQOL concepts (e.g., symptoms, functional status, perception of well-being) as both primary and secondary outcomes measures in the evaluation of new medical treatments and devices, as being an indication of their relevancy across health care. (Aspects of HRQOL and its measurement are discussed in Chapter 20 of this book.)

In considering the types of behavioral health outcomes that might be assessed in healthcare settings, a substantial number of clinicians probably would identify symptomatic change in psychological status as being the most important. However important change in symptom status may have been considered in the past, psychologists and other behavioral healthcare providers have come to realize that change in many other aspects of functioning are equally important indicators of treatment effectiveness. As Sederer, Dickey, and Hermann (1996) explained,

Outcome for patients, families, employers, and payers is not simply confined to symptomatic change. Equally important to those affected by the care rendered is the patient’s capacity to function within a family, community, or work environment or to exist independently, without undue burden to the family and social welfare system. Also important is the patient’s ability to show improvement in any concurrent medical and psychiatric disorder. . . . Finally, not only do patients seek symptomatic improvement, but also they want to experience a subjective sense of health and well being.

(p. 2)

Maruish (2002, 2013b) acknowledged that it is not always easy to determine what exactly should be measured to support a specific outcomes measurement endeavor. He recommended that asking questions such as why the patient sought treatment and what does he or she hope to gain from treatment, and what are the patient’s and the clinician’s criteria for successful treatment should be helpful. Also important is defining “measurable outcomes” so as to permit more objective treatment monitoring as the patient works toward those outcomes (Nash et al., 2013). Thus, “outcomes” holds a different meaning for each of the different parties who have a stake in behavioral healthcare delivery; what is measured generally depends on the purpose(s) for which outcomes assessment is undertaken. As was explained, these can vary greatly.

Notable is the fact that in both behavioral and medical health care, there has been an increased focus on patient-reported outcomes (PRO), that is, outcomes that are based on patient self-report information (Bobbitt et al., 2012). However, the growing interest in PROs is probably the greatest in the medical arena. Psychologists have always found information obtained directly from the patient’s own report as generally valid and certainly valuable in understanding and treating behavioral health problems. It is within the medical arena that seeking out and listening to information reported by the patient—particularly as it pertains to intrapsychic phenomena—is now gaining acceptance as a legitimate means of obtaining information important in diagnosing and treating patients. Moreover, among the many advantages of PROs is that they recognize the patient as a consumer of healthcare services who should be an active participant in their treatment.
CONCLUSION

The US healthcare industry has undergone dramatic changes during the past few decades. What once was a loosely monitored system of care with skyrocketing costs was changed to one with tight controls providing only limited services and choice of providers. These and other efforts to keep costs down (e.g., carve-out behavioral health benefits, lack of parity between medical and behavioral health benefits) were particularly detrimental to those seeking help for mental health and substance abuse problems. For these and other reasons (e.g., stigma associated with mental illness), a considerable number of individuals with behavioral health problems began turning to their primary care providers for help—help that is often not provided or does not adequately meet the patient’s needs.

Fortunately, the winds of change are now blowing in the favor of behavioral healthcare proponents. The enactment of the 2010 ACA, the industry’s realization of the benefits of one-stop health care, NCQA accreditation standards, the growing belief that potential long-term healthcare cost savings can result from the appropriate treatment of behavioral health disorders, and other circumstances bode well for greater access to more and/or better behavioral healthcare services. They also serve as the impetus for a more pervasive integration of primary and behavioral healthcare services throughout the United States.

The alliance of primary and behavioral healthcare providers is not a new phenomenon; it has existed in one form or another for decades. This is partially due to the fact that as many as 70% of primary care visits can be attributed to psychological problems and as many as 50% of primary care patients have a diagnosable behavioral health disorder. Many such problems go undetected and/or are inadequately treated by PCPs. Further, some data suggest that patients with undetected or inadequately treated disorders utilize a disproportionate amount of services and drive costs up, and that early identification and treatment of these individuals can result in cost savings related to long-term medical care, lost income, and lowered productivity in the workplace. Yet, other data do not support the medical cost-offset conclusions. The reality of the situation probably lies somewhere in between.

Regardless, the value that psychologists and other behavioral healthcare professionals bring to the primary care setting—either as an offsite consultant or as an on-site, integrated member of the primary care team—is attested to daily in primary care settings throughout the country. Moreover, there is every indication that the picture of interdisciplinary cooperation in the primary care setting will become more commonplace as the move to integrate behavioral and primary care continues to gain momentum. The extent to which these two services will become integrated will, of course, depend on any number of factors (e.g., funding, available office space, staff interest and motivation) that will vary from setting to setting.

Clinical psychologists and other trained behavioral healthcare professionals can uniquely contribute to efforts to fully integrate their services in primary care settings through the establishment and use of psychological assessment services. Information obtained from psychometrically sound self-report tests and other means (e.g., patient and collateral interviews, clinician rating scales, parent-completed instruments, medical record reviews) can assist the PCP in several types of clinical decision-making activities, including screening for the presence of mental health or substance abuse problems, planning a course of treatment, and monitoring patient progress. Testing can also be used to assess the outcome of treatment that has been provided to patients with mental health or substance abuse problems, thus assisting in determining what works for whom.

The degree to which psychological testing and assessment services become part of the package of primary behavioral healthcare services will depend on the value they bring to the integrated service delivery system. The key to the success of this endeavor will be in the behavioral healthcare
professional’s ability to educate and demonstrate to PCPs and their staffs how test-based assessment can be a cost-effective means of serving the needs of a significant portion of their patient population. It is hoped that this and the other chapters in this volume will assist in these efforts.

REFERENCES

Introduction


The United States (US) healthcare story is grounded in a series of decisions that have created an increasingly fragmented delivery system that costs more and gets less compared to other industrialized nations. Many decisions were commonsense reactions aimed toward improving suboptimal physical and mental health care. The downstream, unintended consequences of these decisions were difficult to anticipate yet functioned to further segregate the healthcare delivery system into isolated silos. When the first wave of national reform began changing the healthcare landscape in the mid-1800s, the evidence base for treating physical and mental illnesses was in its infancy, so the ability to create a coordinated “system” in and of itself was an inconceivable proposition.

In attempts to address some of the challenges of the fragmentation between the mind and the body, efforts throughout healthcare’s history have aimed to better align clinical, operational, and financial structures to support whole-person health care. Over time, creative clinicians and health-care leaders have come to realize the importance of integrating care for physical and mental illnesses. Local innovation and communities of solution throughout the country have arisen to offer more guidance to care for the health of our community than what has been institutionalized in practice. Said differently, despite our history, we are witnessing a massive surge in grassroots innovations that aim to integrate care for physical and mental illness in order to meaningfully improve the health of the population, decrease per capita spending, and enhance the overall experience for the person.

The Triple Aim (Berwick et al., 2008) is currently the dominant health policy in the United States and orients healthcare innovations to focus on improving clinical outcomes, decreasing cost, and leaving patients satisfied with their experience. Sadly, much of what has occurred in health care has not helped achieve these goals. In fact, 17.4% of the federal budget in the United States, which amounts to $2.9 trillion, was allocated to healthcare spending in 2013 (Center for Medicaid/Medicare Services, 2015). Despite being second in the world in appropriating federal dollars to health care—only the country of Tuvalu spends more (World Bank, 2015)—the United States ranks fifth in care quality, ninth in access, last in efficiency, last in equity, last in healthy lives, and last overall (Davis, Stremikis, Squires, & Schoen, 2014). All innovations in health care, whether they are focused on mental health or not, should strive to fulfill these goals. That being said, the research is increasingly clear that ignoring mental health through such a redesign does not produce the desired results.

In 2013, estimates of annual health spending were calculated for medical conditions including mental health. These data indicated that spending for mental health topped out at $201 billion, far costlier than any other health condition (Roehrig, 2016). Other studies have broken down the cost of chronic disease and how these costs increase exponentially when a mental health condition is added (Petterson et al., 2008). One study completed by Milliman (2015), an international actuarial firm, examined the per-member-per-month (PMPM) cost of commercially insured patients with medical conditions who presented with and without a comorbid mental health disorder. They compared these costs
to other patients who had no comorbid mental health need. Findings indicated striking cost differences between these two groups. For example, an individual with arthritis costs $814 PMPM with no mental health diagnosis, $2,065 with Severe and Persistent Mental Illness (SPMI), $1,586 with nonsevere mental health diagnosis, and $1,827 with a substance use disorder (Melek, Norris, & Paulus, 2015).

Milliman took these analyses a step further and explained how most of the surplus healthcare costs from these mental health comorbidities were mainly due to additional medical costs, not mental healthcare costs, which stayed true regardless of the severity of the mental health condition (Melek & Norris, 2008). This finding is not unique, and builds off decades worth of data demonstrating that failure to effectively treat mental health problems increases medical treatment and all associated costs (Cummings, O’Donohue, & Cummings, 2009; Lurie, Manheim, & Dunlop. 2009; Roehrig, 2016). For example, the Milliman study highlights how a person within a commercial health plan who has co-occurring mental health significantly uses more inpatient hospital and emergency room services due to multiple lifestyle factors such as medication treatment nonadherence and unhealthy lifestyles (Melek, Norris, & Paulus, 2015). Additionally, Milliman analyses show similar cost savings potential for the Medicare and Medicaid populations (Melek & Norris, 2008; Melek, Norris, & Paulus, 2015). The authors conclude that cost savings opportunities across all payers is about $300 billion annually (Melek & Norris, 2008; Melek, Norris, & Paulus, 2015).

If such cost savings are accounted for by better mental health treatment, how are various stakeholders responding? Stakeholders across the board ranging from patients to clinicians to policy leaders are advocating for one of the most promising models to systematically identify, treat, and improve health outcomes—a more proactive approach to addressing mental health within primary care settings. And this should come as no surprise considering the linchpin role primary care plays throughout the healthcare system, and the myriad barriers to accessing effective mental health care that has effectively been relegated to an isolated silo.

One needs look no further than how our nation has evolved to treat mental health to recognize how mental health treatment centers have operated largely independent of the larger healthcare system. Across the country, innovative frontline practices and healthcare clinicians are leading the way toward establishing integrated care as the new gold standard. If fragmentation is the enemy of treating the whole person in healthcare settings, then integration is the solution. In fact, the promise of integration is such that it has become a buzzword in policy circles. Scientific evidence no longer allows for naivety that the mind and the body are separate and should be seen as distinct and independent entities in need of specialty care for their parts. No, with a call for more comprehensive health care that places the person first, integrating care is the challenge and the solution. The journey for healthcare transformation begins with a recognition of its history—what has worked, and what has not worked.

This chapter will build off of history and describe new and exciting ways the mental health world has slowly converged with the medical world, with a special emphasis on the integration of mental health with primary care. This chapter is laid out in such a way that the reader can see throughout history where decisions were made that kept mental health as a specialty service and made it challenging to be considered in other delivery settings like primary care. In addition, it is important to understand the development of primary care and what role this setting has played throughout history in taking care of people. Together, primary care and mental health offer a potent combo of comprehensiveness and continuity that is unmatched anywhere else in health care.

**DEFINITIONS**

It is important in any discussion that terms are clearly defined. For the purpose of this chapter, we use definitions from the Agency for Healthcare Research and Quality (AHRQ) Lexicon for Behavioral
Health and Primary Care Integration (Peek, 2013). These definitions are provided below to ensure consistency for the reader throughout the chapter.

Integrated Behavioral Health

The care that results from a practice team of primary care and behavioral health clinicians, working together with patients and families, using a systematic and cost-effective approach to provide patient-centered care for a defined population. This care may address mental health, substance abuse conditions, health behaviors (including their contribution to chronic medical illnesses), life stressors and crises, stress-related physical symptoms, and ineffective patterns of healthcare utilization.

Mental Health Care

Broad array of services and treatments to help people with mental illnesses and those at particular risk of developing them—to suffer less emotional pain and disability and live healthier, longer, more productive lives. Although often defined separately, substance abuse services sometimes are regarded as part of mental health care.

Chemical Dependency/Substance Use Care

Services, treatments, and support to help people with addictions and substance abuse problems of all kinds suffer less emotional pain, family and vocational disturbance, and physical risks and live healthier, longer, more productive lives. Sometimes included under “mental health care.”

Behavioral Health Care

A very broad category often used as an umbrella term for care that addresses behavioral problems bearing on health, including patient activation and health behaviors, mental health conditions, substance use, and other behaviors that relate to health. In this sense, behavioral health care is the job of all kinds of care settings, and is done by clinicians and health coaches of various disciplines or training, including but not limited to mental health professionals. It is a competency of clinics, not only of individuals.

HUMBLE BEGINNINGS: HISTORY AND ORIGINS OF FRAGMENTATION AND THE TURNING POINT TO INTEGRATION

This history of mental health care in the United States (see Figure 2.1) conforms to a recurrent pattern: optimistic policy reform that promises meaningful reductions in the burden of mental illness, followed by pessimism and abandonment as well-intentioned plans for building a better future crumble beneath the weight of the realities and complexities of attempting to provide comprehensive mental health care (Goldman & Morrissey, 1985). However, like most evolving efforts, there are points throughout that signal change, points that lead the effort in a different direction. For mental health and primary care, while there is a history of fragmentation, there is an equally exciting future of integration.

Born out of spirited criticisms of the status quo, each of the three major institutional mental health reforms in the United States aspired to improve the quality of mental health care and the location in which it was delivered. At each juncture, opportunities were missed for aligning mental health care with physical health care in terms of importance. First, the moral treatment movement in the mid-1800s opposed treating people with mental illness like criminals by locking them up in prisons.
Rather, federally sponsored psychiatric hospitals became the locus of care for the mentally ill. Second, the mental hygiene movement in the early 1900s decried the conditions of and treatment rendered in psychiatric hospitals and ushered in a national emphasis on education and prevention. Finally, the community health movement in the mid-1900s acknowledged the need for a more robust acute care delivery system for treatment of less severe psychopathology (i.e., depressive and anxiety disorders) and catalyzed a network of federally sponsored community mental health centers.

**Wave I: Moral Treatment**

The moral treatment movement rested on the belief that residential, hospital-based care was the most effective treatment for people with mental illness. Most notably, Dorothea Dix (1848) acknowledged the prevailing notion of a duty to protect the public from dangerous people, who at the time were people deemed to have mental illness; however, she advocated loudly and effectively with admonishment that “this [duty to public protection] does not justify the public in any state or community, under any circumstances or conditions, in committing the insane to prisons” (p. 2). She petitioned successfully for federally sponsored state psychiatric hospitals that assumed responsibility for caring for the mentally ill. In this model, moral treatment referred to society’s ethical obligation to provide mental health treatment, and advocates persuasively argued that treatment in confined hospitals offered the best prognosis for people with mental illness.

At the time, the advent of the psychiatric hospital represented a substantial social innovation that reflected civilized society’s conviction to treat people with dignity, even those with serious mental illness (Morrissey & Goldman, 1984). This was especially true for communities and families who felt at a loss in how to best care for members and relatives with mental illness (Knapp et al., 2011). In many ways, the promise of the moral treatment movement inspired optimism that people with serious mental illness would no longer be ignored, and would receive the appropriate treatment they needed and deserved. As planned, many people received mental health care who had previously never received treatment. Treatment focused on rehabilitating patients with a combination of positive attention from well-trained attendants and occult practices such as bloodletting, forced vomiting, and extremely cold baths (Sundararaman, 2009). Establishing the psychiatric hospital as the location for treatment of the mentally ill also had the effect of confirming a public perception that mental illness was a serious, yet generally rare condition that affected a small portion of the population who needed intensive specialized care. It was therefore common sense to create a special system, separate from the medical care system, to care for this unique slice of the population. This, like many other decisions, had downstream unintended consequences for the future of mental health delivery.

Soon after, however, several key publications emerged in the literature exposing these less humane aspects of moral treatment and threatened the viability of this approach to mental health care. Penned by people who received treatment in psychiatric hospitals, these reports shaped the public perception that these institutions were characterized less by ethically sound and humane treatment, and more by squalor and basic human rights violations. Nelly Bly, a late nineteenth-century journalist, posed as a woman with mental illness and received treatment at a psychiatric hospital that she denounced in *Ten Days in a Mad-House* (1880). Clifford Beers (1917) similarly wrote a compelling memoir based on his personal experience with bipolar disorder and the concomitant treatment in three different psychiatric hospitals. The images of people with mental illness chained to walls in crowded rooms conjured by these important autobiographical accounts persisted in the public’s perception of mental health care.
In addition to these damning reports, it also became increasingly clear that moral treatment in traditional psychiatric hospitals was not, on the whole, preventing conversion of mental illness into long-term disability (Sundaraman, 2009). As a result, the populations of chronically disabled patients steadily increased, leading to overcrowded, understaffed hospitals that provided fewer resources for both the patients who needed long-term, rehabilitative care and those who less severe forms of mental illness who could potentially recover high levels of functioning. The tension was further exacerbated by the approach to funding these psychiatric hospitals—state funds paid for the construction and renovation of the physical space while local communities maintained responsibility for staffing the hospitals and providing clinical care; this created an unfortunate disincentive for communities to provide long-term care to the chronically mentally ill, and instead incentivized lower cost options such as placement in prisons or denial of treatment (Grob, 1992).

Wave II: Mental Hygiene

In response to the inadequacies of the psychiatric hospitals to humanely and effectively care for people with mental illness, the mental hygiene movement took hold in the early 1900s as the second major reform to mental health services delivery. This reform sought to address mental illness by prioritizing prevention through public health education. With semblance to the social hygiene movement that focused on education and research to reduce juvenile delinquency, crime and sexual deviance, mental hygiene advocates worked to shift the locale of mental health treatment from asylums to public sectors like schools and the media.

Based on the success of his book and the acclaim it received from circles of mental health advocates, Beers collaborated with Adolf Meyer and the preeminent American psychologist William James to launch the mental hygiene movement. Meyer articulated a “biopsychosocial” approach that emphasized the importance of biological, personality, environmental, and social factors in the etiology of mental illness (Wallace, 2007). Further, he emphasized the limitations of the psychiatric hospital and endorsed the idea that, “the problems of mental health and prevention of misfits and diseases must be attacked beyond the walls of the hospitals which today deal with mental defect and disease” (Meyer, 1917, p. 632). With the support of Meyer and other leading mental health experts, Beers successfully created the Connecticut Society for Mental Hygiene, which became the National Committee for Mental Hygiene one year later (Mental Health America, 2016).

The mental hygiene movement did succeed in reframing the treatment of mental illness as a public health concern and emphasized the importance of prevention. In the way of concrete advances, the mental hygiene movement is credited with constructing a database of patients, clinicians, and institutions to inform mental health policy, published the first epidemiological report of mental illness, developed a diagnostic nomenclature to improve communication, and advanced public and professional education related to mental illness (Ridenour, 1961). Public educational campaigns on appropriate childhood and sexual behavior maintained the spotlight as the centerpieces of the mental hygiene movement and suggested that the burden of mental illness could be meaningfully reduced through the development of a more educated public.

Despite these important advances, it remained clear that many people with both severe and more moderate forms of mental illness continued to suffer, and the psychiatric hospital, jointly funded by the state and local communities, would not suffice as the sole provisioner of mental health treatment. A more robust system was needed to provide acute and long-term care to people with mental illness; a new vision was needed for mental health.
Wave III: Community Mental Health

The third reform, the Community Mental Health Movement, recognized the limitations of relying purely on a public health education approach and advocated for the need to provide timely care for people experiencing mental illness in service of preventing the transition from acute suffering to chronic disability. Beginning in the 1940s on the heels of World War II, psychiatrists had been developing short-term treatments to deliver in the field and returned home from the war with firsthand experience of how effective these treatments of shorter duration could be. Simultaneously in the 1950s, the anesthetic properties of phenothiazine and chlorpromazine led to the discovery of anti-psychotic properties of these medications and their use expanded rapidly for the next 20 years (Shen, 1999). Chlorpromazine (generic name for Thorazine) was approved by the FDA in 1954 (Whitaker, 2005) and “initiated a revolution in psychiatry, comparable to the introduction of penicillin in general medicine” (Shorter & Marshall, 1997, p. 255). These two important innovations aligned well with the intensity of treatment to be delivered in the community and helped catalyze this third wave of reform.

Community mental health activists also began developing close ties with the civil liberties reformers as data accumulated on the correlations between low economic class, minority membership, and mental illness (Hollingshead, 1958). Thus, collaborative relationships between community mental health centers (CMHCs) and several other public sectors began to emerge. For example, the Human Relations Service in Massachusetts (Grob, 1992) functioned as a consultative institution to community agencies to provide psychoeducation and guidance on how to address mental illness. Later, the CMHCs developed outpatient services and began coordinating with inpatient hospitals, court agencies, prisons, schools, religious organizations, and other community agencies (Goldman & Morrissey, 1985).

The CMHC movement is sometimes coupled with “deinstitutionalization,” which refers to large-scale efforts to discharge patients from psychiatric hospitals. Although calls for deinstitutionalization date back to the mid-1930s (Grimes, 1934), advocacy turned to exodus in the 1950s (Anthony & Liberman, 1986) as thousands of patients were released from psychiatric hospitals into the community. To accommodate the influx of patients from psychiatric hospitals into the community, President Kennedy (1963) signed the Community Mental Health Act to provide federal funding for the construction of community-based preventive care and treatment facilities, and declared,

this approach relies primarily upon the new knowledge and new drugs acquired and developed in recent years which make it possible for most of the mentally ill to be successfully and quickly treated in their own communities and returned to a useful place in society.

(p. 680)

The long-term effects of the CMHC movement are mixed and continue to be felt today. Positively, people with severe mental illnesses escaped the inhumane conditions of many poorly run psychiatric hospitals and people with less severe mental illnesses were not inappropriately treated in psychiatric hospitals. On the other hand, the population need for mental health services overwhelmed the capacity of the CMHC and the same problems that characterized the psychiatric hospitals—under-resourcing spawning suboptimal treatment—were translocated to the community. By the late 1970s, increased rates of homeless, and to some extent violence, increased among people with more severe mental illnesses, prompting President Jimmy Carter’s Commission on Mental Health (1977) and the Mental Health Systems Act (1980), which further sought to strengthen the community’s capacity within the CMHCs. One year later, President Reagan’s Omnibus Budget Reconciliation Act repealed Carter’s community health legislation, substantially reducing the federal
government’s role in providing mental health care with federal spending decreasing by 30% over the next 10 years (Grob, 1992).

**Summary**

Throughout each wave of reform, opportunities to integrate mental health care within the larger healthcare context were missed; in essence, it was preserved as a specialty function. First, the moral treatment movement successfully increased attention paid to mental health problems with the establishment of the psychiatric hospital. In so doing, it created a powerful precedent that mental illness requires specialized (and often bizarre by today’s standards) care in specialized institutions. Second, the mental hygiene movement sought to improve the conditions of psychiatric hospitals and focus the nation on prevention of disabling mental illness through education and the promotion of healthy behaviors. The ripe opportunity presented to frame mental health as a public health concern, worthy of the same consideration given to physical health, was missed and mental illness remained a severe illness treated separately from physical health problems. Finally, the CMHC movement increased the nation’s capacity to provide mental health and substance use care for a broad spectrum of people that included those with less severe mental illness (i.e., major depression, anxiety disorders, substance use disorders). Once again, however, the reforms were reactive and did not include plans to address mental health issues earlier on in settings like primary care. Thus, not integrating mental health care with medical care helped preserve the majority of mental health provision as a siloed specialty service.

One present day consequence of deinstitutionalization and the CMHC movement is an overwhelming need for acute inpatient psychiatric care because there are literally “not enough beds.” In recent years, states have greatly decreased the number of psychiatric beds due to budget cuts, and an increasing demand on ensuring hospitals maintain their role as profit centers (e.g., Geller & Biebel, 2006). Individuals who are often in need of acute psychiatric services are left looking for a place to go to receive care. Because of the backlog of beds, many individuals are now being housed in prisons (see Figure 2.1). Not only is the prevalence of poor mental health staggering in these settings, but very few, if any, receive the type of care they may need. Another setting that has been shown, time and time again, to see more mental health need is that of primary care.

![Figure 2.1 Mental Health Care in the Prison System](image)
A BRIEF HISTORY OF PRIMARY CARE

Primary care as a function in health care has been elusive in its definition. Barbara Starfield (Donaldson, Yordy, & Vanselow, 1994; Starfield & Shi, 2007), a pioneer and researcher in primary care, and pediatrician from Johns Hopkins defined primary care as “the provision of integrated, accessible healthcare services by clinicians who are accountable for addressing a large majority of personal healthcare needs, developing a sustained partnership with patients, and practicing in the context of family and community” (1994, p. 1). According to Starfield, primary care as a function is meant to be first, foremost, and fundamental to health care, and her definition sets primary care to be the cornerstone of a highly effective and efficient healthcare system (Starfield, Shi, & Mackino, 2005). However, in the United States, there is an imbalance between the generalist function and the specialist function. To this end, Starfield et al. attempted to highlight the key characteristics of primary care in their definition in service to highlighting this critical role (see Table 2.1).

According to Starfield et al. (2005), these characteristics are instrumental in helping enhance and advance primary care. However, for a variety of different reasons, including the hyper-specialization of American medicine, primary care has not held a position of value within the healthcare system. One needs only review the development of the American healthcare system to better grasp the roles hierarchies and power play that position specialized medicine at the top and primary care at the bottom. Paul Starr’s (1982) Pulitzer Prize winning book, The Social Transformation of American Medicine, encapsulates this mentality through his opening sentence: “The dream of reasons did not take power into account” (p. 2).

While it is reasonable to expect that those who do the most in health care would be positioned and rewarded properly, an overtly hierarchical system that rewards specialists more relegates primary care to the lower end of this hierarchy. In terms of dollars, specialists such as orthopedic surgeons ($535,668), cardiologists ($436,849), and hematologists ($376,660) earn more per year on average compared to pediatricians ($206,961), internists ($223,175), and primary care providers (PCPs; $227,541; Hamblin, 2015). In addition to many other factors, imbalanced wages contributes to workforce issues in that there are simply not enough to meet the community demand (Agency for Healthcare Research and Quality, 2011; Bodenheimer & Pham, 2010; Petterson et al., 2012).

While this history of health care has been documented elsewhere, the philosophy around health has indeed shifted significantly. For the purpose of this chapter, four seminal moments that highlight the shift toward primary care as a critical function for addressing all facets of health, including mental

Table 2.1 Starfield’s Key Characteristics of Primary Care

- Primary care is meant to address the provision of comprehensive, coordinated, and continuous services delivered in a seamless fashion combining information about events occurring throughout the patient’s life, over time, and across their life span.
- Primary care is meant to address all health problems that a patient may present with.
- Primary care connects patients to other pieces of the healthcare system.
- Primary care emphasizes continuous care as the heart of the clinician-patient relationship where the patient sees their clinician ongoing throughout their life span.
- Primary care is meant to be available to the patient whenever, however is most appropriate.
- Primary care honors partnerships with patients and families where the clinician works collaboratively to ensure that the patient’s priorities are taken into account.

health, within the larger medical enterprise will be discussed. Of note, as with many movements throughout history, a collective impact began to develop as more evidence and demand emerged. This means that no single event lead to the current integrated care movement; rather, a series of events that built off each other recursively to spawn the culture change that adopts mental health within the purview of primary care.

**Francis Peabody and the Science and Art of Medicine**

In 1927, Francis Peabody insisted on the integration of the “science of medicine” with the “art of medicine” and emphasized the importance of “personal” relationships with patients (Peabody, 1927). In his seminal article, “The Care of the Patient,” Peabody brought forward the idea that at the heart of all healing was the relationship:

Disease in man is never exactly the same as disease in an experimental animal, for in man the disease at once affects and is affected by what we call the emotional life. Thus, the physician who attempts to take care of a patient while he neglects this factor is as unscientific as the investigator who neglects to control all the conditions that may affect his experiment. The good physician knows his patients through and through, and his knowledge is bought dearly. Time, sympathy and understanding must be lavishly dispensed, but the reward is to be found in that personal bond which forms the greatest satisfaction of the practice of medicine. One of the essential qualities of the clinician is interest in humanity, for the secret of the care of the patient is in caring for the patient.

(p. 882)

**George Engel and the Biopsychosocial Model**

The increased attention to the relationship and other social factors really began to emerge in 1977 when George Engel (1977) introduced the biopsychosocial model. Essentially, Engel’s biopsychosocial model was in response to the biomedical model that was so prevalent at the time. According to Engel, the biopsychosocial model took into account missing elements or dimensions not found in medicine. He posited that medical clinicians approached their patients through their conceptual model of the disease and rarely took into account other elements such as a patient’s social status and the unique attributes of the patient as a person. This was a unique view because these factors were historically within the scope of mental health, which, as discussed previously, did not have a prominent role within medical care.

**The Institute of Medicine’s Call for Comprehensive Care**

Nearly 20 years later, a report on primary care in 1996 elaborated the evolving role of medicine, specifically primary care, to the more comprehensive needs of the patient (Institute of Medicine, 1996). When the Institute of Medicine (IOM) released a report on primary care with an appendix on mental health, the connection to primary care was made apparent. Similarly, DeGruy (1996) used decades of evidence to highlight the inseparability of mental health from primary care:

In this paper I will make the case that a major portion of mental health care is rendered in the primary care setting, and always will be, sometimes despite strong disincentives; that a sensible
vision of primary health care must have mental health care woven into its fabric; that the primary care setting is well suited to the provision of most mental health services; that despite suboptimal recognition and management of mental disorders and attention to mental health, the structure and operation of primary care can be modified so as to greatly augment the provision of these services; and that the efforts under way in the United States to reform the health care system offer an opportunity to find the most effective of these modifications and to discover fruitful collaborative structures both within the primary care setting and between primary care clinicians and mental health professionals.

(p. 1)

The Patient Protection and Affordable Care Act

More recently, in the Patient Protection and Affordable Care Act, there were specific provisions meant to enhance primary care. It was thought that as more patients received healthcare coverage, the demand on primary care would increase. While little addressed the role of behavioral health, there was at least a recognition of the important role primary care was to play. The medical home was mentioned 14 times in the law, with primary care being mentioned 113 times.

In fact, the patient-centered medical home (PCMH), a redesign of primary care, incorporated these elements into its seven core principles. These seven joint principles were created by the four major medical societies—the American Academy of Family Physicians, the American Academy of Pediatrics, the American College of Physicians, and the American Osteopathic Association—somewhat in response to the challenge of having a definition for the medical home.

The PCMH is, first and foremost, predicated on the PCP as the personal physician who has an ongoing relationship with patients and leads a physician-directed medical practice consisting of a team of individuals who collectively take responsibility for the ongoing care of patients. This care embodies a whole-person orientation whereby the PCP is responsible for providing for all the patient’s healthcare needs or taking responsibility for appropriately integrating or coordinating care with other qualified professionals. Enhanced care access and high-quality, evidence-based care are hallmarks of the PCMH and are continuously subjected to quality improvement initiatives that are based on patient preferences and national guidelines. Finally, the guidelines stipulate a payment framework that reflects the value of work that falls outside of the fee-for-service reimbursement model, provides opportunities for practices to share in savings from reduced hospitalizations, and permit additional payments for achieving measurable and continuous quality improvements (Ferrante, Balasubramanian, Hudson, & Crabtree, 2010).

Due to the exclusion of behavioral health from these principles, many articles appeared flagging the need for behavioral health. For example, one article by Petterson et al. (2008) used the cost of untreated mental health to make the case that the medical home should be prepared to address mental health issues. Other authors have also flagged the necessity of these principles, including behavioral health (Ader et al., 2015). However, it was a group of family medicine leaders who made this need more explicit. In an article entitled “Joint Principles: Integrating Behavioral Health Care Into the Patient-Centered Medical Home,” behavioral health was added to each principle to provide a clearer sense of what should be expected around behavioral health in the PCMH (Baird et al., 2014). While aspirational, the clarity of the recommendations made an impact and eventually showed up in accreditation standards for the medical home (National Committee for Quality Assurance, 2014). These changes, however aspirational, helped usher in a new way of thinking around the medical home. More specifically, with new standards that hold practices accountable for behavioral health, there is a higher likelihood of adoption.
BUILDING A BRIDGE, EXPANDING A MOVEMENT: INTEGRATION AS THE SOLUTION

A call to action to bring about better integration of mental health began in the 1990s with the IOM report, but has continued as the evidence has mounted. In Blount’s (1998) field-defining book, Integrated Primary Care: The Future of Medical and Mental Health Collaboration, he lays out a vision for having mental health clinicians work side by side with primary care clinicians. The goal is for a more seamless experience for patients around mental health and primary care. However, as with all emerging fields, the evidence around the particular delivery models to employ has been slow to mount. In fact, it took several years for a more complete understanding of what was and what was not integrated mental health.

Blount (2003), in his article on organizing the evidence, provided a typology for integrating care to alleviate confusion around the components of each program, who did what, and which patient population received systematic care. His framing was done in the context of distinguishing approaches to integrate behavioral health into three categories: coordinated, co-located, and integrated. Coordinated services are those that happen in different locations, but are connected by clinicians attempting to communicate about their patient. Co-located may be mental health services provided on-site in primary care, but not as a part of the primary care team. Usually in these scenarios the mental health clinicians see their own patient, and not those of the primary care team. Finally, for integrated approaches, mental health is fully embedded on the primary care team.

To clarify the distinction between these three approaches, Blount (2003) stated:

The categories distinguish between services that are coordinated, but exist in different settings, services that are co-located, both being provided within the same practice location, and services that are integrated. Integrated services have medical and behavioral health (and possibly other) components within one treatment plan for a specific patient or population of patients. Technically, it is possible for services to be co-located but not coordinated or to be integrated but not co-located, so the most precise definition of these descriptions would be that they are dimensions of collaborative care, not mutually exclusive categories.

(p. 26)

Further, Blount (2003) broke down the evidence around integration into categories that helped differentiate programs that were targeted versus nontargeted, specified versus unspecified, and small scale versus implementation. This distinction in level of integration and model composition facilitated the development of methods to assess the myriad examples of integrated care that were emerging throughout the country. Indeed, this framing was one of the first heuristics to organize the various approaches to furnishing whole person care in the same location. With more specificity in describing the scope of integrated care, attention then turned to further defining metrics to evaluate outcomes of integrated care.

Miller, Mendenhall, and Malik (2009) built off Blount’s article by adding specific elements associated with each level of integration to be considered for outcome measurement. The authors categorized models of integration into those that were referral based (co-location and care management) from those that were more consultant based like integrated care. However, it was the conclusion of the article that was the precursor for what was to become:

The field must incorporate a standard language when describing models, components, and metrics of collaborative care. We challenge the field to begin to consistently use a unified vocabulary and break
down the various models of collaborative care to individual metrics. When we combine our care delivery sites to measure these individual metrics, we begin to move beyond a conceptual level and more clearly attempt to empirically discriminate—and ultimately advance—collaborative care.

(p. 29)

Building off this need for more measurement and research, in the fall of 2010, a small conference funded by the Agency for Healthcare Research and Quality (AHRQ) convened in Denver, Colorado, to create a research agenda for advancing integrated behavioral health and primary care. This report highlighted key gaps in the evidence base and also included a lexicon for integrated behavioral health and primary care (Miller, Kessler, Peek, & Kallenberg, 2011). The gathering concluded that it was difficult if not impossible to research integration without having a consistent definition to use to allow for empirical discrimination. Peek (2013) supplied this definition in the AHRQ Lexicon:

The care that results from a practice team of primary care and behavioral health clinicians, working together with patients and families, using a systematic and cost-effective approach to provide patient-centered care for a defined population. This care may address mental health, substance abuse conditions, health behaviors (including their contribution to chronic medical illnesses), life stressors and crises, stress-related physical symptoms, ineffective patterns of health care utilization.

(p. 2)

The Evidence Base for Integration

Evidence for the integration of behavioral and primary care began coalescing in the late 1970s through the early 1990s as it became increasingly clear that a substantial portion of the US population obtained their mental health care in primary care (Katon & Schulberg, 1992; Regier et al., 1978; Schulberg, 1991). These observations triggered a national focus on the importance of screening for depression and providing subsequent quality care (Depression Guideline Panel, 1993). The influence of the biopsychosocial model (Engel, 1977) and the IOM’s (1996) report, coupled with increased recognition to treat co-occurring depression and other chronic health conditions, led to collaborative care models for depression that became very popular.

Large initiatives to develop, evaluate, and refine collaborative care models for depression emerged rapidly (see Table 2.2). Several systematic reviews and meta-analyses ensued and summarized these data that contributed to a strong evidence base for collaborative care for the treatment of depression in primary care. Gilbody and colleagues’ (2006) meta-analysis of collaborative care for depression concluded that compared to usual care, collaborative care for depression results in better outcomes at posttreatment and up to 5-year follow-up. Moderators included patient adherence to treatment and case managers with backgrounds in mental health in combination with regular supervision. Oxman and colleagues’ (2005) systematic review focused more on the practical aspects of implementing these more comprehensive, population-based approaches to depression care. Concerned more with effectiveness outcomes in terms of intervention transportability, implementation, and sustainability, Oxman and colleagues (2005) emphasized the importance of fostering the professional relations between primary care and mental health clinicians in an effort to improve reach.

The AHRQ sponsored a 266-page review of the evidence for integrated behavioral health and primary care and concluded that, overall, this model of care improves outcomes but that integration quality is not related to outcome in a dose-response fashion (Butler et al., 2008). This significant heterogeneity in the effects of integrated care on clinical outcomes is an important aspect of the evidence
Craven and Bland (2006) identified several factors that clinicians and delivery system leadership believed to be important to improved outcomes—practice readiness, co-located behavioral health services with interdisciplinary collaboration, systematic follow-up, psychoeducation, and strategies to improve patient engagement. However, it must be noted that these are qualitative and represent the beliefs of the queried participants. A retrospective meta-regression (Bower et al., 2006) extended these findings by identifying four factors associated with improved depression outcomes as the result of integrated care—non-US setting, recruitment method (explain), case managers with mental health backgrounds, and regular case manager supervision. A more rigorous test of “active ingredients” that drive improved outcomes would involve adequately powered quantitative methods designed to prospectively test for mediation.

Subsequent reviews have attempted to analyze the relationship between integration quality and outcome by focusing on more granular aspects of integration. For example, Foy et al. (2010) noted

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Elements</th>
<th>Main Finding</th>
</tr>
</thead>
</table>
| IMPACT (Unützer et al., 2002) | • Collaborative care team consisting of care manager, PCP, and consulting psychiatrist  
|                         | • Stepped-care delivery model of systematic screening, psychoeducation, brief psychotherapy, and medication management | Increased receipt of treatment and response to treatment                      |
| PROSPECT (Bruce et al., 2004) | • Collaborative care team consisting of care manager (nurse, social or psychologist) and PCP focused on suicide prevention in the elderly  
|                         | • Patient-centered approach offering citalopram or IPT by the care manager            | Increased access to depression care, greater declines in suicidal ideation, earlier treatment response, and higher rates of remission at 4, 8, and 24 months |
| RESPECT-D (Dietrich et al., 2004) | • Collaborative care team consisting of care manager, PCP, and consulting            | Increased rates of remission at 3 and 6 months                                 |
|                         | • Program includes psychoeducation, engagement strategies, assessment of treatment response, and care coordination |                                                                              |
| PRISM-E (Areán et al., 2008) | • Co-located mental health and primary care clinicians                               | Increased access but comparable outcomes to enhanced referral and worse outcomes than enhanced referral for patients with major depression |
| DIAMOND (Solberg et al., 2013) | • Collaborative care team consisting of care manager, PCP, and consulting psychiatrist  
|                         | • Stepped-care delivery model of systematic screening, psychoeducation, brief psychotherapy, and medication management  
|                         | • Global budget payment mechanism in the state of Minnesota                           | Results not published                                                        |
increased effect sizes in the small to medium range when integrated care involves direct, personal contact between PCPs and behavioral health clinicians, an observation confirmed by an independent, national review of Medicare Demonstration Projects (Nelson, 2012).

**Limitations of the Evidence Base**

Kwan and Nease (2013) provide an excellent summary of the evidence on integrated care that synthesizes findings from these systematic reviews and meta-analyses, and importantly, highlights four major gaps in the evidence base. First, in evaluating the importance of structural features of integrated care, information technology, training, practice policies, and physical space factors have received less attention than care team composition, screening strategies, and treatment protocols. Further research is needed to clarify the extent to which these “brick and mortar” factors contribute to clinical and financial outcomes. Second, the evidence is limited to a near exclusive focus on depression at the expense of attention paid to other important clinical problems and to whole-person versus disease-specific care. Although depression is a highly burdensome disease, other diseases and conditions (anxiety disorders, substance use disorders, subthreshold syndromes, etc.) are also prevalent and warrant further attention. Third, the bulk of studies employ randomized controlled trial designs within large academic, healthcare settings that may be less useful and generalizable to “real-world” settings where integrated care is sorely needed. For example, the State Innovation Models initiative (Centers for Medicare and Medicaid Services, 2015) is seeking to improve the delivery and payment for innovative services at the state population level (e.g., integrated behavioral health and primary care). And fourth, outside of grant-funded research and implementation projects, there are few established business models capable of financially sustaining integrated care. Examples of this include the Sustaining Healthcare Across Integrated Primary Care Efforts (SHAPE) to evaluate six integrated primary care practices—three operating under a traditional fee-for-service payment mechanism (control practices) versus three practices that received a global payment for services (intervention practices) in western Colorado.

**BARRIERS**

Despite such a robust literature to support integrated care, substantial barriers remain that inhibit widespread adoption. While immediate indicators point to the separation of mental health from the larger healthcare system, ironically, many barriers to integrating care begin with the lack of investment that has been made in primary care. On average, the US healthcare system invests only about 5% of general healthcare spending on primary care (Phillips & Bazemore, 2010). Compared to other aspects of healthcare delivery (e.g., specialists, emergency department, hospitals), it does not take much to understand the problems with this approach and how it is not in support of primary care. At its core, integrated, whole-person care relies on the capacity of our healthcare linchpin—the primary care system—to function effectively. When one separates out mental health from primary care, one sees a comprehensive and generalist function that is limited in scope and leads to higher cost, less satisfaction, and poorer outcomes (deGruy, 1996). And this separation begins with how care is paid for.

**Payment Models**

Experts are clear that financing is the most significant barrier for advancing integrated behavioral health and primary care (Kathol, Butler, McAlpine, & Kane, 2010). Due to the US healthcare delivery
system traditionally paying for mental health services through financial mechanisms different than those in the medical sphere, integrating care has been stymied by an antiquated payment structure that supports individual clinicians within a fee-for-service model rather than teams of clinicians working together within capitated or global payment models (Hubley & Miller, 2016). This arrangement, in so many ways, has perpetuated the false dichotomy between mind and body that is reflected in separate clinical, financial, and educational systems for mental health and medical clinicians. This duality has created distinct cultures between clinicians who treat the mind and clinicians who treat the body, insufficient training models to prepare clinicians to work together on multidisciplinary teams, and payment models that will support many functions of behavioral health clinicians working in integrated settings.

Financial barriers are often listed as the most significant in achieving integrated care. Multiple authors have described the challenges with paying for behavioral health in primary care. Most recently, Hubley and Miller (2016) outlined the various payment models that are in support and sometimes problematic for integrating care. Specifically, they highlight the pros and cons of four common payment models: fee-for-service, which provides reimbursement for individual services; pay-for-performance, which provides reimbursement for meeting benchmarks of care process or outcomes; bundled payments, which provides one large reimbursement for a group of services for a particular condition or treatment; and global capitation, which provides one large reimbursement for the comprehensive services needed for a defined population.

In another example, the AHRQ’s (2015) Guidebook of Professional Practices for Behavioral Health and Primary Care Integration: Observations From Exemplary Sites highlighted several different strategies for exemplary integrated practices and how they approached financing their behavioral health efforts. For example, negotiating with external financial stakeholders and managing internal finances emerged as the two key, overarching strategies. The report further specified recommendations on how to engage key stakeholders and manage internal finances to assure some level of sustainability.

It is clear that to better support integration, there needs to be payment models that invest in primary care, include behavioral health, and begin to migrate away from volume-based payments to those that are more value-based (Kathol, deGruy, & Rollman, 2014).

**Different Cultures**

Changing the culture of care delivery is often fraught with challenge due to the historical inaccuracies in how one defines health. As discussed, the mere terminology of mental and physical that are attached as adjectives to our health care does not support the idea of whole health care that is comprehensive and integrated. The paradigm of creating integrated care means that assumptions about health and definitions are reexamined and redefined, respectively. It is no surprise that the barriers to integration are profound, as somewhat a disruption to traditional care delivery, and that these barriers require attention to be paid to how we deliver, pay for, and operationalize care (Peek, 2008). Integration requires us to think more specifically about how behavioral health services are delivered outside of specialty settings, and ways that this can be financially supported by policy and paid for through new and novel mechanisms.

**Training**

Blount (1998) has written extensively on the need to prepare our behavioral health workforce for primary care in ways that our current training and education programs do not. He laid out a vision for a workforce that collaborated side by side in primary care. In fact, he went so far as to say we will
have a workforce crisis if we do not better prepare behavioral health for working in primary care settings (Blount & Miller, 2009). While opportunities to receive graduate and postdoctoral training in integrated care settings is slowly increasing, the majority of behavioral health and primary care clinicians receive training within their own specialty and have fewer opportunities for working in multidisciplinary teams, especially those that include strong representation from mental health. For successful integration, there needs to be more connection between our training programs to allow for collaboration and exposure to other disciplines throughout training.

**Summary**

There are substantial barriers that prevent integrating care. At a clinical level, integrating care requires that each clinician, from their respective backgrounds and trainings, learn to work together in a shared environment with a shared purpose for a shared population. In a recent framework released by the Robert Wood Johnson Foundation (Miller, Gilchrist, Ross, Wong, & Green, 2016), six different elements were laid out to support the broader adoption of behavioral health and primary care. All six of these recommendations can be seen in Figure 2.2. These interdependent elements can be used as a roadmap for scaling much of the integration work nationally.

As challenging as it is to navigate the multiple logistics of trying to integrate behavioral health in primary care, overcoming different cultures of care and training backgrounds is not the most formidable barrier. Historically, care delivery is based less on the evidence of what works and more on payment models that incentivize service-driven health care. Payment reform is needed to support the implementation and sustainability of integrated care models that we know work.

The barriers to advancing integrated care as the de facto model can be overcome when they are framed as opportunities to improve the quality and sustainability of integrated care. With all healthcare-related efforts in need of satisfying the Triple Aim, now is the time to correct some of the missteps health care has made on mental health. Now is the time to acknowledge the crucial role that primary care plays in the delivery of mental health in particular, and how a team-based approach may be much more efficient and effective to assist with the Triple Aim goals within the broader healthcare context.

![Figure 2.2 Culture of Whole Health Recommendations](image_url)
CONCLUSIONS

Despite a history of fragmentation, the future of health care is integrated. At no other time in history have there been more attempts to bring together the disparate pieces of health care, mental and physical, as we are witnessing now. Integration is a movement, and like all movements this one requires organization and sustainment. From the beginning of our understanding of what is health, we have thought of the mind and body as separate. Evidence no longer allows this to be true. In fact, we are now attempting to integrate care in spite of our history and culture.

REFERENCES

Kathol, R. G., Butler, M., McAlpine, D. D., & Kane, R. L. (2010). Barriers to physical and mental condition integrated service delivery. Psychosomatic Medicine, 72(6), 511–518. doi:10.1097/PSY.0b013e3181e2c4a0


Primary care medicine has been a continuously evolving and increasingly important aspect of health care. From a societal standpoint, one of the more significant evolutions in health care has been the recent implementation and subsequent revisions of the Patient Protection and Affordable Care Act (ACA; P.L. 111–148). The ACA has begun to address both long-standing perceived weaknesses of past healthcare delivery systems (e.g., imbalances in coverage, isolated care services, overall costs of medical care) and localized patient needs in an ever-changing demographic and resource landscape (e.g., aging baby-boomers; Garfield, Zuvekas, Lave, & Donohue, 2011).

One of the major weaknesses of the healthcare system has been the lack of coverage for the mentally ill. The National Institutes of Mental Health has noted that a large proportion of the US population—1 in 5 adults or 43.8 million individuals—suffers from mental illness in a given year (Abuse, 2014). With regard to insurance, contemporary projections have estimated that as much as 21% of those with severe mental illness do not possess the necessary or needed healthcare coverage. Efforts such as the ACA have begun to fill this deficit by expanding the reach of public coverage to allow for as many as 39 million new patients to receive coverage and utilize needed mental healthcare services over the next several years (Garfield et al., 2011; Rozensky, 2013).

As more individuals receive healthcare coverage, and the overall US population grows, providers within primary care must contend with a larger and ever-growing number of patients as well as noted physician shortages in the face of foreseeable increases in workload. Colwill, Cultice, and Kruse (2008) have predicted that between the years 2005–2025, family and general medicine physicians will be faced with a 29% workload increase but will only see a 2% to 7% increase in the number of physicians in these areas. Reverting or directing the flow of patients to specialist care in order to abate these pressures is not a low-cost alternative nor the most effective, as the United States already boasts one of the highest number of specialists compared to other industrialized countries (Starfield, Shi, & Macinko, 2005). Despite these unceasing structural changes and pressures, the primary care setting has continued to strive for and provide effective patient care, and continues to be an integral part of healthcare delivery in the United States (Ferrer, Hambidge, & Maly, 2005; Starfield et al., 2005).

This chapter will provide an overview and insight into how psychologists and psychological assessment has become a significant part of the primary care mission. In doing so, the following pages will cover important aspects such as the implementation and integration of behavioral health specialists in the primary care setting; the implications, advantages, and use of psychological assessments and screeners in the primary care clinic; and the effects of managed care on selection and choice of psychological assessment measures. Finally, the chapter will conclude with a brief look at assessment procedure data collected from an integrated primary care clinic, and a description of a proposed mini-battery of measures that has been found to be useful in this setting.
THE PATIENT-CENTERED MEDICAL HOME AND INTEGRATED CARE

In the current healthcare environment, attempts to meet the healthcare needs of the public have taken multifaceted forms designed to balance both effective care for patients with complex medical problems as well as system resources. Of note, one of the more recent and better-known ways primary care has attempted to acquire this balance within the passing of the ACA has been to follow care delivery models, such as the Patient-Centered Medical Home (PCMH) (Henderson, Princell, & Martin, 2012; Sklar, Aarons, O'Connell, Davidson, & Groessl, 2015).

The PCMH is designed specifically around primary care medicine, with primary care physicians (e.g., family medicine, pediatrics, internal medicine, OB/GYN) serving as the main source in which patients are able to receive both their general medical care and referrals to specialists when needed. Typically, the PCMH model attempts to offer increased benefits and improvements to patients in various ways, such as timely scheduling of appointments and extended access to their physicians, while in keeping within a cost-effective approach. This also includes major shifts in treatment perspectives, as the PCMH strives to be more proactive in preventative care and early detection of disease. While the early conceptions of the PCMH can be traced to the 1960s, it is more recently that mental health and psychosocial factors have been brought into consideration as a main component of this model. This has become more important as the most recent revision of the Joint Principles of the Patient-Centered Medical Home (JPPCMH) has underscored the importance of integrated behavioral health care (Baird et al. 2014; McDaniel & deGruy, 2014; Sklar et al., 2015). Specific to behavioral health care, the JPPCMH stipulates that in the effort to view the patient as a "whole person" rather than in specialized segments, the PCMH must take into account and acknowledge behavioral, psychological, and psychosocial factors. These recommendations also include collaboration between behavioral health providers and primary care physicians, as well as providing on-site behavioral health professionals in order to meet the biopsychosocial needs of patients (Baird et al., 2014).

In part, organizational and policy changes related to mental health coincide with the continuing trends of high prevalence rates of mental disorders among primary care patients. In the United States, large proportions of patients are increasingly seeking treatment for their mental health needs in the primary care setting (Mechanic & Bilder, 2004; Norquist & Regier, 1996; O'Donohue et al., 2005). Of note, studies of prevalence rates have indicated that the proportion of primary care patients meeting criteria for a mental disorder typically ranges from 42.5% (Ansseau et al., 2004) to upwards of 53.6% (Roca et al., 2009). The most common disorders, subthreshold conditions, and symptoms involve anxiety, depression, alcohol and substance use, and somatoform disorders. In turn, primary care providers now prescribe 70% of psychotropic medications to these patients (Ansseau et al., 2004; Hunter, Goodie, Oordt, & Dobmeyer, 2009; Lute & Manson, 2015; Porcerelli et al., 2013).

Primary care patients with behavioral health complaints or psychiatric disorders are not always one-dimensional, as similar rates of comorbidity have been highlighted in the literature (Didden, Philbrick, & Schorling, 2001). The two most often seen psychiatric disorders in primary care (i.e., generalized anxiety and depression) are also the most common comorbid conditions, accounting for upwards of 40% to 60% of patients with psychiatric disorders. For patients with common chronic medical conditions, such as diabetes mellitus, rates of depression are almost twice as high in comparison with healthy individuals (Piette, Richardson, & Valenstein, 2004). Other studies have also identified chronic diseases such as lupus (Achtman, Kling, Feng, Okawa, & Werth, 2016) and somatic conditions such as cardiovascular disease (Gili et al., 2010) as having similarly high rates of co-occurring psychiatric illness.
Patients with co-occurring physical and behavioral health conditions also have an impact on the doctor-patient relationship. Patients with these complex conditions are often experienced as “difficult” by their primary care physicians (Hahn, Thompson, Wills, Stern, & Budner, 2004). When confronted by such patients, physicians often become more directive and less collaborative (Noyes, Longley, Langbehn, Stuart, & Kukoyi, 2010; Seaburn et al., 2005). These patients also require more visits to healthcare providers and possess a higher overall cost of care (Didden et al., 2001; Lin et al., 2006; Porcerelli, Bornstein, Markova, & Huprich, 2009; Simon et al., 2000). Relatedly, these patients are also at a higher risk for developing more serious medical conditions (Miller, Paschall, & Svendsen, 2006; Thombs et al., 2006).

The primary care setting serves an important role in addressing the needs of patients unable to receive treatment from currently overwhelmed or absent mental health resources (e.g., psychiatry and community mental health clinics). Often, it may serve as an important first line of treatment and/or referral base. Given these concerns, some organizational shifts and restructuring have attempted to adopt “integrated care” or “primary care behavioral health” models in order to meet these needs. Taking a whole-person perspective, integrated care approaches set out to accomplish several goals. First, in order to assist with the high rates of patients with mental illness previously mentioned, behavioral health specialists (e.g., psychologists) are able to extend treatment (e.g., offering brief psychosocial interventions) to patients within the primary care setting (Hunter et al., 2009). By utilizing behavioral health specialists on-site, the cumbersome process of referring patients to separate outpatient mental health services, where high numbers of patients do not follow up, is avoided. In addition, mental health providers often fail to provide feedback to referring physicians regarding the psychosocial care of patients. This is an especially important concern for patients that may lack both the resources and the wherewithal to follow up at outpatient psychiatric service clinics, which are often overwhelmed with high numbers of patients.

The potential level of communication and consistency of patient care is one of the benefits of integrated models (Cohen et al., 2015). Within this approach, clinicians of different specialties (e.g., family medicine physicians and behavioral health providers) are able to share proximity and work collaboratively by providing each other and colleagues with up-to-date information about shared patients. Additionally, this shared space also offers familiarity for patients as they are able to consistently see the same behavioral health and medical specialists and are encouraged to keep their medical appointments. This can be an essential issue for patients that are dealing with chronic or comorbid health concerns and may need frequent monitoring of healthy lifestyles and choices.

Studies that have investigated the implementation of the integrated care approaches have noted several key benefits over standard practice (i.e., primary care physicians referring to mental health specialists outside of the office). In particular, some have noted that overall healthcare costs associated with patients in the integrated care setting tend to be lower than typical primary care practice (Hought, 2002). Benefits of the integrated care approach are also seen at the patient level. For instance, after receiving consultations from psychologists in the primary care clinic, patients exhibited minimized symptoms and reported high rates of satisfaction and strong therapeutic relationships with providers (Angantyr, Rimner, Nordén, & Norlander, 2015; Corso et al., 2012). The benefits of the integrated primary care setting are not only realized by patients, as physicians working within this model generally feel satisfied with the integration of psychologists, have been able to reduce their own levels of stress associated with medical care, and feel that it has improved overall patient care in their clinics (Miller-Matero et al., 2016).

The use of behavioral health services and their integration into standard practice is one of the many ways the primary clinic has adapted to the necessity of conceptualizing and treating the whole
FILLING THE GAP

The functions of psychologists in primary care have included administrative roles, teaching (primarily residency education), direct patient care, and research (Fischetti & McCutchan, 2002). Despite these advantages, many psychologists may be faced with initial difficulty in having what they can offer and their expertise fully realized by medical staff. Some studies have found that physicians in integrated primary care settings may underutilize mental health professionals, in part, because many lack a full awareness of what psychologists are able to offer their patients (i.e., fully appreciating the large scope of symptoms and health-related behaviors that psychologists are able to provide expert assistance and collaboration with) (Beacham, Herbst, Streitwieser, Schu, & Siber, 2012; Hunter et al., 2009; William, Eckstrom, Avery, & Unützer, 2015). An unfortunate result of underutilized behavioral health professionals in clinical settings, or in clinics that do not possess mental health staff, is that often there is difficulty with recognizing psychiatric concerns in their patients.

While these high prevalence rates have been noted in the literature, a significant number of patients leave the primary care office undiagnosed and untreated. In part, the accuracy of physician identification and diagnosis of mental disorders has played an important role in this problem. Studies have found that approximately 50% of patients that qualify for a mental illness diagnosis, such as depression, are identified by their physicians (Carey et al., 2014; Dilts, Mann, & Dilts, 2003). Similarly, methods of evaluating a patient solely by interview have also yielded subpar results when compared to multimethod approaches (e.g., interview and brief screening instrument) when attempting to gain a full understanding and diagnostic picture of a patient (Meyer et al., 2001). Obviously, with the previously mentioned comorbid and complex presentation of many of these patients, these rates of unrecognized mental health problems cause unnecessary patient suffering and drive up healthcare costs. If more primary care clinics are able to adopt integrative approaches, the recognition and needs of these patients will be fully met at a greater rate. By collaboratively treating these patients, physicians are able to spend less time diagnosing and to ruling out cases and more cost-effective referrals can be made to required specialists when needed (Ferrer et al., 2005).

A large number of issues have been raised as to why physicians might have difficulty with routinely identifying and diagnosing mental disorders in their patients. Usually, this has been related to how the physicians are trained (i.e., not enough attention is given to mental illness during training) but also to structural or environmental concerns, such as less time available to see their patients, lower rates of personal satisfaction because of increasing demands, and/or patients themselves not being able to effectively communicate their mental health needs (Bodenheimer, 2006; Hartley, Korsen, Bird, & Agger, 1998). In addition to these factors, the emotional attunement or perceptibility of physicians appears to be a significant mediator in the recognition of mental illness in medical patients, with physicians who are more perceptive to nonverbal expressions and who possess a higher level of interest in the psychosocial concerns of their patients tending to exhibit better performance in accurately diagnosing psychiatric illness (Robbins, Kirmayer, Cathébras, Yaffe, & Dworkind, 1994). It seems that the wide array of medical concerns that primary care physicians must face may also play a role in hampering their abilities in the detection of mental illness. This issue is most prominent in disorders with high prevalence rates, such as depression, where primary care physicians might use different criteria for diagnosing depression (e.g., being perceptive of depression symptoms related to physical health, or viewing psychiatric symptoms as a response to difficult physical complications
or disease) than compared to psychiatrists, who may be more open to the psychological concerns of their patients (Armstrong & Earnshaw, 2004).

Roadblocks and stymied progress in the integration of psychologists working collaboratively in the primary care setting cannot be entirely pinned on the lack of awareness of medical professionals. Rather, psychologists must reevaluate how they not only approach other health professionals, but also how they work with patients in this setting. While practice in the primary care setting shares some similarities with how a psychologist operates in private psychotherapy practice, there exist several pertinent distinctions. Most notably, the intensity in which psychologists might treat patients in the common mental health setting differs markedly in an integrated primary care clinic. For instance, many psychologists in private practice might work with patients under more prolonged circumstances (e.g., amount of time spent with each patient and the number of sessions with each patient), whereas in the primary care clinic many patients are seen for brief consultations (5–10 minutes). If formal psychotherapeutic interventions are offered, they tend to be shorter in length (e.g., 15- to 30-minute sessions). For example, in one integrated, university-based, primary care residency training clinic, the modal number of counseling sessions offered by psychology trainees (doctoral graduate student) within a 36-month period was two (Porcerelli & Jones, 2015). In large part, this difference can be attributed the voluminous and expedited flow of patients in the primary clinic (Nash, McKay, Vogel, & Masters, 2012).

Another difference in the integrated primary care setting is the focus on early detection and preventative assessment (e.g., screening procedures), whereas psychologists working in private practice or outpatient mental health clinics employ assessment techniques and treat existing mental health concerns. The psychologist may now function as the safeguard for patients in between their visits with physicians. This could include not only up-to-date communication about progress to the physician, but also follow-up appointments with patients in order to reinforce healthy living habits and address immediate stresses and psychological concerns (McDaniel & deGruy, 2014; Nash et al., 2012). Therefore, it is imperative that psychologists utilize both their expertise in brief mental health interventions and reliable and valid brief screening/assessment instruments.

A CASE FOR PSYCHOLOGICAL ASSESSMENT

Of the variety of services that psychologists may serve within the integrated primary care, perhaps one of the most important resides in the practice of psychological assessment and screening. In part, this importance comes from the fact that psychological assessment has been one of the unique achievements in the field of psychology, as well as one of psychology’s launch-points into clinical work. Advanced expertise in evaluation and testing is what often sets psychologists apart from other potential behavioral health specialists in the integrated care setting, such as licensed clinical social workers. In recent decades, as testing and psychometric development has advanced, many psychological tests now share similar levels of validity on par with commonly used medical tests (Kubiszyn et al., 2000; Meyer et al., 2001).

How do integrated care models utilize psychological assessment in their repertoire of services? Systematic reviews of service usage have not always found promising results. For example, Martin, White, Hodgson, Lamson, and Irons (2014) noted that a relatively small number (N = 11, 14.4%) of studies (N = 75) investigated the uses of psychological assessment or screening. These low rates raise a few questions pertaining to the importance of psychological assessment within the parameters and goals of primary care. This issue of low use of psychological assessment is intriguing given the fact that psychological assessment is characterized by a number of benefits. From the integrated care
perspective, uses of assessment methods have been linked with early detection of mental illness and lower overall healthcare costs (Derogatis & Lynn, 2000). Therefore, an evaluation of the potential uses of assessment measures in primary care is warranted and differs from what psychologists are normally trained to cover and accomplish (e.g., extensive history, cognitive functioning, achievement abilities).

In particular, the psychologist operating within the integrated primary care model must take into consideration the overall goals of the primary care setting. In this effort, Talen and Valeras (2013) list four important behavioral health dimensions that should be a part of the psychological assessment in this setting, such as (1) promoting wellness and healthy behavior, (2) identifying mental health and substance abuse factors and symptoms, (3) identifying medical conditions that affect behavioral health functioning, and (4) identifying co-occurring biomedical and psychosocial conditions. While each of these dimensions is within the purview of the services offered by the psychologists, it is important to note that medical staff (e.g., physicians, medical assistants, and clerical staff) can be a useful initial gauge of patients that are likely to need assessment services (Talen & Valeras, 2013).

The recommended re-evaluation of assessment in primary care settings necessarily takes on the form of a utility analysis of assessment protocols and procedures. In order for assessment instruments to be used in primary care settings, utility must include ease of use, relatively low cost, and quick to score and interpret. This includes a need for assessment materials to be translatable to other professionals often working alongside the psychologist in the integrated setting. For example, in some situations, a physician must convey assessment findings to patients during follow-up appointments. It is imperative that measures used do not require advanced psychometric knowledge in order to understand and interpret. This point rests in the fact that after proper identification of behavioral health concerns, treatment often involves multiple medical professionals other than the behavioral health provider (e.g., referrals outside of the integrated care clinic) (Dobmeyer & Miller, 2014; Talen & Valeras, 2013). Chapter 5 of this book provides a comprehensive discussion on and recommendations for test selection criteria.

While full or comprehensive assessment is an essential tool of the behavioral health professional, another related and vitally important function involves the use of brief screening measures. By using a wide variety of assessment tools, the behavioral health professional is able to expedite the process of correctly identifying common coexisting psychological disturbances as well as communicating this information in an effective way so that it can be used within the treatment team. Explicitly, screening patients accomplishes two significant goals. First, patients that present to physicians with mental health concerns (e.g., symptomology related to depression, anxiety, substance use) can be quickly screened by either the physicians or the mental health professional. Knowledge of screening is crucial for physicians who may find themselves working within rural settings without the assistance of behavioral health providers.

Similar to utilizing larger assessment measures that are easy to understand, useful screeners should be characterized by ease of administration for both behavioral health professionals and physicians. Screeners should effectively balance streamlined assessment with meaningful findings. Using screening methods in order to assess commonly seen disorders in primary care allows clinicians to expeditiously gain an adequate picture of the mental functioning of a patient within the fast-paced primary care environment. In addition, clinicians can also strategically use screeners. Selecting specific criteria (e.g., high number of clinic visits, body mass index [BMI], presence of vague symptoms) with which to use as a basis for employing screeners helps to hone medical clinicians’ skills in identifying patients with potential mental health concerns and cuts down on needless screening (Talen & Valeras, 2013). Additionally, screening instruments that require a low level of reading ability can easily and reliably be completed by patients without assistance, eliminating time demands on physicians or other medical staff.
There are also several benefits in training medical staff to utilize screeners. Having a physician conduct a brief screener with their patients allows for the continued reinforcement and development of the physician–patient relationship. It may prove beneficial for patients to gain a perception of their physicians as agents that are attempting to help them with all areas of their lives, including concerns with mental functioning. In turn, by helping physicians become more aware of not only the mental health prevalence rates in their clinics but how to properly assess mental disturbance, many of the aforementioned deficits seen in physician detection of mental illness may be lessened.

The second goal of the use of mental health screeners is to allow for quick progress assessment for follow-up appointments. For example, some measures, such as the Beck Depression Inventory (2nd ed.; BDI-II; Beck, Steer, & Brown, 1996) or the Patient Health Questionnaire 9-item Depression scale (Kroenke, Spitzer, Williams, & Löwe, 2009), can be reliably given to patients during follow-up appointments at little or no cost. This periodic monitoring provides the basis for communication between the mental health professional, medical staff, and patients as well as a historical picture of the patient’s functioning over time. A more in-depth discussion of screeners is offered later.

Early Implementations of Psychological Assessment in Medical Care

In many respects, the origins of psychological testing in the United States can be traced to James McKeen Cattell, who was one of the first to create an early physical and mental battery of tests. Additionally, Lightner Witmer’s clinic at the University of Pennsylvania is often cited as one of the first to provide treatment for school-aged children who experienced behavioral and scholastic difficulties (Benjamin, 2005; DuBois, 1970). Historically, however, the pivotal moment in the use of psychological assessment in medical settings can be pinpointed near the end of World War II. It was during this time that medical professionals were being faced with an overwhelming number of soldiers exhibiting a cluster of symptoms known as “shell shock,” a similar problem faced by clinicians returning from the battlefields of World War I (Vander Weg & Suls, 2014). As a result, clinical psychologists were called upon to provide services (e.g., psychotherapy, psychological assessment) for large numbers of returning soldiers, and federal organizations, such as the Veterans Administration, created training programs that bolstered the number of clinical psychologists in medical settings (Benjamin, 2005; Vander Weg & Suls, 2014). Initially, during this time, psychologists mainly served to fulfill the capacities of diagnosis and assessment, with the treatment of patients being left to psychiatrists. The assessment procedures and batteries at this time largely consisted of early measures of intelligence and various aptitude tests, with the eventual integration of personality measures as well.

Stemming from these early insights and promising utility of educational and military psychological assessment, these methods quickly began to expand outward toward other medical settings. In part, the increasing use of psychologists in medical settings could also be due to a shift in medical thinking toward the consideration of the “whole” patient. Engel (1977) postulated that the dominant form of medical practice and care did not readily accept psychosocial concerns and influences on patients’ health, which warranted a shift in perspective that was guided by both psychiatrists and clinical psychologists. Notably, primary care medicine served as one of the places that mental health professionals, especially psychologists, began to be an integral part of patient care (Pace, Chaney, Mullins, & Olson, 1995). Primarily, this integration has been comprised of two main areas. First, many psychologists have been employed by family medicine departments as educators in an effort to train resident physicians in assessing and treating psychosocial issues of their patients and providing
wellness interventions for residents (e.g., leading Balint groups). This role would also include the
task of dissemination of new advancements or research findings to residents and faculty members.
Second, the mental health professional also functioned collaboratively with other medical staff and
began to assist with the often burdensome and at times difficult task of identifying patients dealing
with mental illness (Pace et al., 1995). Given these concerns, efforts have been made to provide phys-
icians and medical staff with diagnostic guidelines or standardized procedures in order to increase
rates of accurate identification of mental disorders.

Effects of Managed Care on Psychological Assessment

Since its implementation, managed care has effectively covered a large proportion of patients in the
United States, with upwards of 160 million patients being enrolled in some form of managed care
(National Conference of State Legislatures, 2011). Overtime, studies have investigated the effects of
managed care on medical practice and have found generally mixed reviews. Often, general practice
physicians have felt that managed care changes to practice have led to pressure to see a higher vol-
ume of patients, increasingly restricted their autonomy to competently care for patient needs, and
slowly lessened their overall positive perception of the future and practice of medicine (Deckard,
1995; Murray et al., 2001; Stoddard, Hargraves, Reed, & Vratil, 2001; Williams & Skinner, 2003).
An important aspect of the restriction of autonomy is related to the overt managed care mission to
control and reduce costs of treatment. This has been a noted concern, as the United States has one
of the most costly healthcare systems, with estimated amounts of approximately $8,000 per person
(Martin, Lassman, Whittle, Catlin, & National Health Expenditure Accounts Team, 2011). Managed
care has attempted to reduce overall healthcare costs by limiting access to and reliance on specialist
referrals.

For example, when treating a primary care patient for chronic major depression, many primary
care physicians may feel pressured to provide an initial psychopharmacologic intervention first before
referring them to psychiatry. In turn, this has encouraged primary care physicians to broaden their
prescribing scope in order to provide treatment for their patients (Schreiter et al., 2013; Schulberg,
also affect their patients. Predictively, studies have found that patients with satisfied and confident
physicians tend to possess higher rates of medical adherence and feel that their physicians are trust-
worthy (Grembowski et al., 2005; Williams & Skinner, 2003).

Similarly, as the directives and changes brought on by managed care have had significant influ-
ence over physicians, it has also reshaped the practice of behavioral health care and psychological
assessment. Within the larger context of mental health care, many practitioners have found managed
care to be both a negative and potentially detrimental aspect to the practice of psychological assess-
ment (see Acklin, 1996; Cerney, 1994). Past examination of managed care and psychological practice
have found that practitioners frequently have difficulty in obtaining payment for assessments (Mur-
phy, DeBernardo, & Shoemaker, 1998), and that decreasing payments by third-party payers has only
served to starve the mental health system (Appelbaum, 2003).

The changes brought on by managed care have had rippling effects on payment for services and
the selection and use of psychological tests. This can be observed longitudinally, as Cashel (2002)
notes that surveys undertaken during the 1960s found that measures such as the Wechsler Adult
Intelligence Scale, Rorschach, Thematic Apperception Test, the Bender-Gestalt Visual Motor Test, the
Minnesota Multiphasic Personality Inventory, and the Draw-A-Person test were widely used in com-
prehensive batteries, with significant proportions of this battery being continuously used into the
latter half of the twentieth century (see Camara, Nathan, & Puente, 2000).
Despite this relative continuity of selected tests, more recent surveys on psychological test use have found significant changes as the result of managed care. Overall use of psychological testing to inform clinical decisions has declined over the years, with psychologists opting for briefer, symptom targeted measures that are less costly (Piotrowski, Belter, & Keller, 1998). Notably, areas of testing, such as projective or free-response (e.g., Rorschach, Thematic Apperception Test) and cognitive functioning (e.g., Wechsler scales), have seen drops in usage because of the time demands placed upon clinicians by these tests as well as their associated costs. In addition to this, it seems that many managed care providers (i.e., those tasked with approving the reimbursement of psychological tests) do not appreciate the utility or the benefits provided by commonly used tests (Ambrose, 1997). Contemporary practice now includes a smaller proportion of tests even though many practitioners find that tests that are no longer viable in the managed care environment (e.g., Rorschach, WAIS) are still useful and important to psychological testing overall (Piotrowski, Belter, & Keller, 1998).

While a significant driving force behind the implementation of managed behavioral health care is to reduce costs, perceptible improvements in the quality of services have also been appreciated. Managed care has had relatively benign effects in settings outside of primary care, with some clinicians reporting little change in their sense of satisfaction, compensation, and autonomy since its implementation (Isett, Ellis, Topping, & Morrissey, 2009). For patients, managed care also did not decrease a patient’s access to mental health care and has not negatively affected health outcomes (Grembowski et al., 2002). In fact, others have also pointed out that under managed care, a patient’s access to treatment has increased overall (Mechanic & Bilder, 2004).

It appears that while many changes in the practice of mental health care have followed the implementation of managed care, it has had the positive benefits of provoking a concrete and concerted look toward assessment practices for the integrated care psychologist. Fortunately, Turchik, Kapenko, Hammers, and McNamara (2007) have outlined essential guidelines for the practice of assessment within low budget and rural settings under managed care, which are helpful for the psychologist in the integrated care setting. Specifically, clinicians need to be wary of giving psychological tests in a routine or unfocused manner, and should opt for brief screeners as a first-line approach before engaging in a more substantial assessment. Additionally, other factors such as using tests that primarily yield clinically relevant information is also vital.

The use of reliable and valid tests in the public domain or at relatively low cost to clinics, as well as asking patients to cover a portion of the costs of assessment, are essential for the integrated care psychologist to consider when attempting to initiate assessment methods under managed care (Turchik et al., 2007). Despite this, new focuses on psychological assessment, especially within the integrated primary care setting, have been brought on by the ACA. In particular, some have noted that new psychologists need to receive enhanced training in assessment methods, which are oriented toward prevention and recovery rather than symptom reduction in order to remain as a potent contributor to patient health care in the era of the ACA (Beidas & Manderscheid, 2014; Chor, Olin, & Hoagwood, 2014). In turn, integrated primary care mental health assessment will continue to emphasize preventive methods and evidence-based outcomes of assessment as increasing numbers of training programs, as well as PCMHs, attempt to fit the ACA model of health care.

**BEHAVIORAL HEALTH AND THE PRIMARY CARE SETTING**

The following sections give a practical description of the role of the behavioral scientist in the integrated primary care setting. This includes an overview of one of the first screening tools available for this setting as well as modifications and extensions to the original instrument.
Behavioral science has had a rich history in primary care medicine, especially with family medicine residency programs. Family medicine is the first specialty in medicine to require its residencies to include a behavioral science faculty member. Behavioral scientists include providers from several different mental health fields, including psychology, family therapy, social work, and psychiatric nursing. The primary role of the behavioral scientist was precepting, which involved reviewing patients within the clinic setting with residents and faculty physicians in real time. Since the early days, the role of behavioral scientists within residency programs has expanded to include teaching (didactics and clinical/ambulatory), precepting, outpatient care, inpatient consultation, scholarship, and administration (Fischetti & McCutchan, 2002). This expansion has prepared psychologists to be collaborators and leaders in the integrated care movement and PCMH (Kaslow, Kapoor, Dunn, & Graves, 2015). From a clinical perspective, psychologists working in integrated healthcare settings are now working shoulder to shoulder with other primary care teams of providers. They serve to reinforce a whole-person view of patients by providing a psychological perspective to wellness, disease prevention, and chronic disease management, as well as patients with mental health symptoms and disorders. Although psychologists use their knowledge of assessment in all their clinical endeavors, traditional psychological assessment instruments have not been well integrated into the primary care setting. The lack of use of traditional assessment instruments is likely due to the fast pace of the primary care environment and the restricted amount of total time spent with each patient. The history and transformation of the Primary Care Evaluation of Mental Disorders (PRIME-MD; Spitzer et al., 1994) into the Patient Health Questionnaire (PHQ-9) serves as a striking example of what has happened to traditional psychological assessment instruments in the primary care setting.

The History and Importance of the PRIME-MD

One of the earliest primary care assessment instruments is the Primary Care Evaluation of Mental Disorders (PRIME-MD; Spitzer et al., 1994). The PRIME-MD was developed as an interview-based measure to assist primary care physicians in assessing the presence of mental disorders. Specifically, the PRIME-MD was designed to measure five different types of mental disorders that were felt to be the most common in this setting (i.e., somatoform, mood, anxiety, alcoholism, and eating). The PRIME-MD was divided into two parts: a brief 26-item yes/no questionnaire of symptoms of these disorders and a semi-structured interview for the physician to administer to patients who reported symptoms on the aforementioned questionnaire. The self-report questionnaire was to be completed and scored prior to the patient’s visit with the physician.

The development and implementation of the PRIME-MD was brought about for several clinically relevant reasons. First, as previously mentioned, given the voluminous rate at which physicians are required to assess patients, less time and energy is available to commit or devote to a thorough psychological assessment. One of the first studies on the utility of the PRIME-MD was able to highlight the problem of nonrecognition among physicians, as 60% of patients that were unrecognized by their physicians were identified by the PRIME-MD (Spitzer et al., 1994). Although the PRIME-MD was a psychometrically acceptable assessment instrument, it was never fully adopted by primary care physicians because of the time that it took to administer to patients. Eventually, the screening questionnaire that was designed to alert physicians to the presence of a psychological disorder was modified and turned into a comprehensive self-report instrument called the Patient Health Questionnaire (PHQ; Spitzer, Kroenke, Williams, & Patient Health Questionnaire Primary Care Study Group. 1999).
The PHQ is a 78-item questionnaire made up of eight scales—somatoform disorder; major depressive syndrome; other depressive syndrome; panic syndrome; other anxiety syndrome; bulimia nervosa; binge eating disorder; alcohol abuse—and additional questions regarding stressors and menstrual/pregnancy/childbirth. Although the PHQ was a comprehensive assessment instrument, it was also deemed “too long” by primary care clinic administrators and clinicians. Therefore, some of the individual scales were extracted and/or modified to create research instruments (e.g., PHQ-15 somatization scale) and clinical scales, such as the PHQ-9 for the assessment of major depressive disorder and the GAD-7 for the assessment of generalized anxiety disorder. It should be noted that the PHQ-9 and GAD-7 have been shortened into the PHQ-2 and GAD-2 for initial screening purposes. The PHQ-9 is addressed in Chapter 9 of this book while the GAD-7 is discussed in Chapter 12.

The history of the PRIME-MD illustrates some of the complexities involved in the use of psychological assessment instruments in primary care settings. Several important lessons have been learned from the PRIME-MD story. Thus, in order for psychological assessment instruments to be adopted by primary care providers and administrators, they need to (1) be brief and focused, (2) be easily scored and interpreted, and (3) assess the most common behavioral health disorders encountered by clinicians in primary care.

**TYPES OF ASSESSMENT IN PRIMARY CARE**

In the following, the types of assessment used in a university-based integrated primary care residency-training clinic will be described. Note that not all types of assessment methods and measures described here are uniformly used in all integrated primary care settings. However, what is provided is an overview of an outline of psychological assessment (e.g., measures used, variants of assessment) employed in this setting.

**Screening**

In a fast-paced primary care setting with access to direct behavioral health providers, screening is essential for identifying, treating, and monitoring patients with psychological symptoms and disorders. Brief measures such as the 9-item Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001) for depression, the 7-item Generalized Anxiety Disorder questionnaire (GAD-7; Spitzer, Kroenke, Williams, & Löwe, 2006) for generalized anxiety, and the 10-item Alcohol Use Disorders Identification Test (AUDIT-10) for alcohol abuse, are commonly used screens in primary care settings. They are easily completed by patients and scored by providers. Clear cutoff scores indicating levels of severity make them especially useful. Some of these instruments have been shortened to provide initial screens. For example, there is a 2-item PHQ (PHQ-2; Kroenke, Spitzer, & Williams, 2003) which includes the mood and anhedonia items from the PHQ-9. Some settings will only give the PHQ-9 if a patient has a positive screen on the PHQ-2. The same is true for the 2-item GAD (GAD-2; Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007) and the 3-item AUDIT (AUDIT-C; Bush, Kivlahan, McDonell, Fihn, & Bradley, 1998). For psychologists interested in practicing in primary care settings, knowledge of screening instruments (i.e., their sensitivity, specificity, and clinical utility) is essential. Many screening instruments are available online and can be downloaded without a charge. However, such measures should be carefully evaluated before using them in a clinical setting to ensure their psychometric soundness.

Primary care physicians are often the first providers to come in contact with individuals with subthreshold (i.e., having only a few symptoms without meeting all of the criteria for a particular
disorder) and full-blown psychological and psychiatric disorders. Thus, primary care psychologists need to become familiar with many different types of mnemonics to aid them with screening interviews as well as screening instruments for a wide range of disorders, including measures that assess the degree of impairment brought about from such disorders. Most commonly, the mnemonic S-I-G-E-C-A-P-S (i.e., Sleep, Interest, Guilt, Energy, Concentration, Appetite, Psychomotor, and Suicide) is used when assessing for depression (Thomas, Michael, Gary, Pharm, & Eugene, 2001).

In addition to screening instruments such as the PHQ-9, GAD-7, and AUDIT-10, reliable and valid screening instruments are available for other problems and disorders commonly seen by primary care providers: ADHD child/adolescent (e.g., Vanderbilt ADHD Teacher Rating Scale; Pliszka, 2007) and adult (e.g., ADHD Rating Scale Attention Deficit/Hyperactivity; DuPaul, Power, Anastopoulos, & Reid, 1998); somatization disorder (e.g., Somatic Symptom Scale-8; Gierk et al., 2014); drug (e.g., Drug Abuse Screening Test; Skinner, 1982) and opiate abuse (Screen and Opiate Assessment for Patients with Pain; Butler, Budman, Fernandez, & Jamison, 2004); posttraumatic stress disorder (e.g., Primary Care-PTSD Screen; Prins et al., 2003) and traumatic events (e.g., Life Events Checklist-5; Weathers et al., 2013); intimate partner violence (e.g., Brief Conflict Tactics Scale; Feldhaus et al., 1997); suicide (e.g., Columbia Suicide Assessment Rating Scale; Posner et al., 2011); insomnia (e.g., Insomnia Severity Index; Morin, Belleville, Belanger, & Ivers, 2011); personality/psychopathology in child/adolescent (e.g., Jellinek et al., 1999) and adult (e.g., Personality Inventory for DSM-5; Krueger et al., 2012); dementia (e.g., Saint Louis Mini-Mental Status Exam; Tariq, Tumosa, Chibnall, Perry, & Merley, 2006); mild cognitive impairment (e.g., Montreal Cognitive Assessment; Nasreddine et al., 2005); pain (e.g., Brief Pain Inventory; Cleeland & Ryan, 1994); and pain disability (e.g., Pain Disability Index; Tait, Chibnall, & Krause, 1990).

Psychologists as well as other primary care providers need to understand that screening is just the beginning of the diagnostic process. Mindless use of screeners to make diagnostic or treatment decisions by providers may lead to inaccuracies that may lead to unnecessary patient suffering. For example, a patient with a PHQ-9 score in the severe range who is not also screened for a history of bipolar illness could become manic if prescribed an antidepressant without a mood stabilizer.

It is also important that psychologists and other behavioral health providers not only assess symptom severity and psychiatric disorders but also assess the patient’s functional status (Robinson & Strosahl, 2009). This is of particular importance because of the brief nature of psychosocial interventions in primary care and the reality that primary care providers have a greater chance of lessening a patient’s functional impairment than curing a (chronic) psychiatric diagnosis.

A description of screening in primary care practices would not be complete without mention of the clinic huddle. The clinic huddle is a brief, 5- to 10-minute meeting prior to the start of the morning involving the entire clinic staff for the purpose of reviewing all scheduled patients with their presenting problems in order to prepare for those who are likely to require behavioral health assessments and/or interventions (Porcerelli et al., 2013). Thus the huddle may help a behavioral health provider to prepare by collecting the types of screening instruments and interview protocols that may be needed to perform a brief yet comprehensive assessment for a particular patient.

**Pull-ins and Warm Handoffs**

The terms “pull-in” and “warm handoff” are commonly used in integrated primary healthcare settings. A pull-in refers to when a physician literally pulls a psychologist or other behavioral health provider into the exam room with a patient to help out with an assessment or intervention (Porcerelli et al., 2013). Common questions by physicians include (from global to specific): “Can you help me figure out what’s going on with this patient?” “Is the patient suffering from a mental illness?” “Is the
patient depressed?” “This patient’s family is wondering if their father is showing signs of dementia.” Pull-ins can occur within a standard 15-minute medical visit or can be done at the end of a physician’s visit with a patient, with the expectation that after the behavioral health provider is introduced to the patient, he/she will work with the patient while the physician proceeds to the next patient on their list. This latter example is referred to as a warm handoff. The warm handoff by a physician who has already developed trust and rapport with the patient transfers that trust to the behavioral health provider. The warm handoff prevents delays in care and reduces the need for return visits (Hunter et al., 2009). It is one of the more common procedures provided by behavioral health providers within a primary care setting and one of the most important activities to lessen patient flow problems.

During pull-ins, behavioral health providers rely on their knowledge of bread-and-butter disorders and problems seen in primary care settings. The ability to quickly and accurately assess behavioral health needs of patients is one of the core competencies for providers practicing in integrated healthcare settings (Bluestein & Cubic, 2009).

**Consultations**

When a behavioral health provider is not available, a primary care physician will request that the patient schedule a separate session with a behavioral health provider to obtain or clarify diagnostic questions. These diagnostic questions typically involve those that can be addressed by a semi-structured interview and one or more screening instruments. Consultations tend occur with patients presenting with more complicated medical (e.g., uncontrolled diabetes mellitus) and behavioral/psychiatric (e.g., major depression) problems. Common referral questions include, “My patient, Mr. X is depressed. Please rule out bipolar illness” or “Robert is a 15-year-old ninth grader with a BMI of 35. His parents are wondering if his overeating is the result of his peer-related anxiety.” Both of these examples would require not only knowledge of the common behavioral health disorders and interventions in primary care but also the biopsychosocial knowledge of the health impact of co-occurring medical and psychological disorders (Blount & Miller, 2009; Bluestein & Cubic, 2009).

**Initial Assessments as Part of Behavioral Health Interventions**

It is common for primary care physicians to refer their patients for specific behavioral health interventions (e.g., cognitive behavioral therapy, or CBT, for anxiety). It is important not to assume that the initial assessment conducted by a primary care physician was comprehensive. Reviewing the initial diagnosis with the patient and performing an assessment may be necessary in order to choose and carry out an appropriate brief intervention.

**Psychological Testing**

Psychological assessment has a long and rich history. Formal psychological assessment came into being over 100 years ago when intellectual measures were designed to identify children who needed special academic attention. In the 1920s formal psychological testing addressing individual differences in psychological adjustment, especially in response to military service, emerged (Weiner & Greene, 2008). With the onset of World War II in 1941 and the subsequent expansion of clinical psychology, psychological assessment took hold in academic, military, forensic, and psychiatric settings. In some psychiatric and medical settings (e.g., rehabilitation units in hospitals), comprehensive psychological and neuropsychological test batteries ranging in length from 3 to 6 or more hours were not uncommon.
Managed care organizations have questioned the necessity for such costly batteries, which has made such practices less commonplace. In primary care settings, formal psychological assessments that include intellectual, achievement, psychopathology, and personality measures are rare. Psychologists who offer psychological assessment services are considered specialists. When assessment psychologists work in academic medical centers, they tend to be affiliated with departments of psychiatry.

For the past three years in an academic family medicine training clinic, the use of a mini-psychological test battery has been tested. In a recent study examining the types of assessment that took place within a primary care clinic between 2012 and 2014 by doctoral trainees, King, Jones, Melnyk, and Porcerelli (2016) reported a total of 691 assessments. A total of 471 (48%) assessments were conducted by interview alone, 151 (22%) were conducted by a combination of interview and one or more screening instruments, and 69 (10%) underwent psychological testing using a mini-battery. The mini-battery consists of a measure of intelligence, the Wechsler Abbreviated Scale of Intelligence II-Second Edition (Wechsler, 2011), a measure of academic achievement, the Wide Range Achievement Test-Fourth Edition (Wilkinson & Robertson, 2006), brief neuropsychological screens (e.g., Montreal Cognitive Assessment), and one or more brief psychopathology measures (e.g., Personality Assessment Screener; Morey, 1997). When a free-response (projective) measure was included in the mini-battery, the brief Early Memory Index (Karliner, Westrich, Shedler, & Mayman, 1996) was administered because of its ease of administration and scoring as well as its performance in urban and suburban primary care settings (Porcerelli et al., 2016). These batteries can be administered in 2 to 3 hours and the instruments can be scored in an hour. For the primary care setting, reports by trainees cannot exceed one page. When restricted to one page, physicians will read the report in its entirety (and not just the diagnostic summary).

What has been learned about the use of a mini-battery in primary care settings is that it provides the most comprehensive information to the referring physician. From an educational standpoint, physicians and healthcare teams learn the most about their patients when IQ, achievement, and psychological aspects are reported on. However, mini-batteries may not be practical if there is only one psychologist covering an entire clinic or where reimbursement for more comprehensive testing is low or nonexistent. In such instances, the psychologist may only be able to use a single brief measure (e.g., an alcohol abuse screen) or refer those patients who need a more comprehensive assessment to an assessment specialist. In settings that include practicum students or interns, comprehensive assessments may be justified for training purposes as opposed to a reimbursable service.

THE FUTURE OF ASSESSMENT IN PRIMARY CARE

Psychological assessment remains perhaps the most important activity of psychologists and other behavioral health providers in primary care settings. Despite the fact that the amount of time behavioral health providers have to spend with each patient, as compared to traditional mental health settings, an accurate assessment must be made in order to provide effective evidence-based brief interventions and/or proper referral to a community agency or specialists such as a psychiatrist.

Psychologists working in primary care need to be knowledgeable and have the interviewing skills to screen and assess a wide range of behavioral and psychiatric symptoms and disorders. Knowledge and access to screening instruments is also required within a fast-paced setting. There is an ongoing need to develop and validate new measures specifically for primary care settings (Porcerelli, Kurtz, Cogan, Markova, & Mickens, 2012). The issue of clinical utility is thus paramount. One promising area for gathering assessment data is the use of patient portals. Patient portals involve having patients complete health- and assessment-related information remotely (i.e., from home or clinic waiting room).
This allows for automated scoring and reports of assessment finding for providers prior to patient visits. A discussion of computer-based, in-office and distance administration, scoring, and reporting of psychological tests is presented in Chapter 5 of this book.

More systematic research is needed to develop best practices for monitoring clinical improvement and outcomes over time within the integrated PCMH (Croghan & Brown, 2010; Kearney, Wray, Dollar, & King, 2015). There are measures such as the PHQ-9 and GAD-7 that are brief, reliable, valid, and easy to score that can alert clinicians about improvements in or worsening of symptom severity.

Psychologists can contribute to the development and ongoing validation of new measures that assess constructs that are important to primary care providers. For example, the issue of health literacy has become an important component of the PCMH, as large numbers of patients, approximately 30% (Kutner, Greenburg, Jin, & Paulsen, 2006), struggle with low levels of health literacy. Patients’ ability to understand and make use of important health information and recommendations made by their healthcare providers is dependent upon their level of health literacy (Institute of Medicine, 2004).

New measures, such as the 5-item Brief Health Literacy Screen have been developed as initial validation studies, have found promising results (Sand-Jecklin & Coyle, 2014). Several other constructs have proven to be important components of patient and provider wellness and satisfaction, such as quality of life (of both patient and providers; see Chapter 20 of this book) and the patient–provider relationship (as rated by both patient and providers). With training in the science and practice of assessment, psychologists are well positioned to make substantial contributions to the practice of assessment in integrative primary care settings.

REFERENCES


Piette, J.D., Richardson, C., & Valenstein, M. (2004). Addressing the needs of patients with multiple chronic illnesses: The case of diabetes and depression. American Journal of Managed Care, 10(2; PART 2), 152–162.


Psychologists have an ethical responsibility to strive to benefit those with whom they work and take care to do no harm. To meet this responsibility, psychologists must be aware of their boundaries of competence, especially in emerging areas of practice where recognized standards for preparatory training do not exist. Primary care psychology is an example of an emerging area of practice. Although these data have not been collected, it is expected that the vast majority of licensed psychologists in the United States do not have the competencies to work effectively in primary care. Although state licensure may allow psychologists to provide care within a primary care context, this licensure does not mean they are competent to do so. Psychologists can become competent through a number of avenues, but for most, it takes training beyond licensure.

Competence can be defined in a number of ways and may vary depending on the level of competence needed to practice at a given level. Epstein and Hundert (2002) developed a dynamic definition of competence for medical schools that is commonly cited and applies equally well to competencies for psychological practice; thus “Professional competence is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individual and community being served” (p. 226). Although mastery of a professional knowledge base is a necessary component of competence, it cannot serve as a proxy for competency in practice (Nelson, 2007).

The importance of competency-based training is not new to psychology. Nelson (2007) details a 30-year history of professional psychology moving from a knowledge-based focus on training to one that focuses on applying knowledge to produce observable clinical behavior and performance. More recently, competency-based education and training in psychology has come to a sharp focus as a result of the following: (1) a need for accountability to ensure psychologists to benefit the public and not do harm; (2) the importance of outcome-based education and learner-based outcomes; and (3) newly trained psychologists being able to fully enter the workforce upon the attainment of their doctorate (Fouad & Grus, 2014). Toward these efforts the American Psychological Association (APA) sponsored the Assessment of Competency Benchmarks Work Group that produced a set of benchmarks for understanding and measuring competence in professional psychology (Fouad et al., 2009). Those competencies were updated in 2011 into six overarching clusters containing 16 core competencies (Hatcher et al., 2013). Even with general consensus on core competencies for psychological practice, areas of specialty (e.g., clinical health psychology, geropsychology, clinical neuropsychology, rehabilitation psychology) recognize that additional competencies are needed in specialization and have created specialty core competencies to guide training and practice (Fouad & Grus, 2014).

Despite the efforts to develop general competencies for professional psychology practice as well as detailed competencies for specialty areas, psychologists sometimes believe that because they are competent for general practice, or perhaps a specialty like clinical health psychology, that they are
compotent to practice in primary care. This notion could not be further from the truth. For over 15 years we (JLG and CLH) have seen psychologists without the appropriate primary care competencies fail spectacularly when attempting to work in primary care. A recent study (Hall et al., 2015) confirmed this experience with the examination of 19 US practices that varied in size, ownership, location, and experience delivering integrated care. Results showed organizations had difficulty finding clinicians who had the skills and experience necessary for working in an integrated practice. Newer practices underestimated the time and resources needed to train new clinicians; through trial and error, practices learned that clinicians needed relevant training to work effectively as an integrated team. They concluded that “Clinicians lack appropriate training and experience needed for effective behavioral health and primary care integration” (Hall et al., 2015, p. S50).

The need for a set of agreed upon competencies for psychology practice in primary care has not gone unnoticed. In 2012, as APA president, Dr. Susan Bennett Johnson formed an interorganizational workgroup to delineate competencies for primary care psychology. The workgroup comprised representatives from nine national organizations with a central focus on education or practice in primary care psychology and delineated six broad areas of competence with detailed subareas. Each competency includes the essential knowledge, skills, and attitudes as well as behavioral anchor examples (McDaniel et al., 2014). These competencies were approved by the APA council in 2015, with the competencies serving as aspirational goals for psychologists that would guide training programs’ curriculum development and psychologists’ self-monitoring (APA, 2015a). These competencies are discussed in more detail later.

To ethically work in primary care as a psychologist and provide efficient and effective team-based care, additional training needs to take place around the competencies delineated in the APA competencies for primary care psychology. Doing so improves the chances that primary care psychology practice will be successful and produce the desired healthcare impact in the patients being served. In this chapter we describe the observations of students as they move from learning about traditional mental health practices to learning and experiencing practice in primary care. Also, we describe the foundational guidelines that shape practice in primary care, discuss the necessary competencies for psychologists working in the primary care setting, and explain how to measure whether those competencies are met and maintained.

STUDENTS’ PERSPECTIVES: EXPECTATIONS VERSUS REALITY

Perhaps the most important perspectives on practicing in primary care come from the current and future students who will be practicing in these environments. Graduate students are able to identify unique and surprising aspects of primary care work, which is sometimes forgotten or devalued by those with years of experience in these settings. These perspectives help to identify and guide necessary trainings for primary care psychologists.

Learning the skills for primary care psychology presents both challenges and surprises for clinical psychology students. The first author (JLG) taught a doctoral level course in integrated behavioral health in primary care, along with a family physician who had several years of clinical collaboration with the first author. The third author (CLH) lectured in the class, and the second author (CMW) was a third-year doctoral student in the course. The course covered a range of topics including history; core concepts and models of integrated primary care; understanding the context of primary care; policies and implementation; ethics, diversity, training and research considerations; assessment and documentation; and core intervention strategies for conditions commonly seen in primary care. The required coursework, beyond assigned readings and class discussions, included memorizing a
verbal introduction and informed consent, writing a paper describing how to approach a condition commonly seen in primary care and comparing/contrasting its treatment in specialty care with appropriate treatment in primary care, and conducting a 30-minute role play demonstrating an initial assessment using the primary care behavioral health model. The structure of the course and an example of the weekly readings are provided in Appendix A.

Prior to enrolling in the course students had exposure to a broad range of clinical and health psychology-related courses, but almost no practical experience with integrated primary care. Students took the course in their third or fourth year of clinical training after completing at least one year of clinical practicum training in a specialty mental health clinic. They enrolled in the course with varying degrees of understanding regarding the models of integrated primary care and what to expect; however, they all experienced differences between the expectations about integrated primary care and the reality presented during the course. Student reflections on the experience are described next in four general categories: the integrated care model, clinical skills required, adjustments to practice management, and ethical differences. Although these ideas are from a very small sample of students, their observations reflect common anecdotal observations.

**Integrated Care Model**

Students were surprised to learn that “integrated primary care” was more than simply sharing office space with primary care providers and running separate but co-located practices. Instead, it could mean merging psychological treatment into the medical clinic and structure. Further, they were surprised to learn how much effort has been allocated to developing models, describing/practicing techniques, and creating policies and training to expand the knowledge and competence in this area. Students were also surprised to realize that psychology practice in integrated primary care is not “abbreviated psychology” or “triage psychology,” but instead is a type of clinical practice requiring the ability to quickly conceptualize and provide interventions in the first appointment. Thus, a high level of clinical skill is needed to provide assessment and treatment for a wide range of conditions on a moment’s notice. However, along with the challenges of obtaining sufficient generalist skills to practice in primary care, students also observed the tremendous benefits integrated primary care has at the population health level.

**Clinical Skills**

In addition to recognizing the high level of clinical skill needed for assessing and treating a wide range of conditions on short notice, students observed the challenges of evaluating the assessment and treatment literature and summarizing it into the most critical and usable components. Further, students commented on the need to “unlearn” or utilize skills in a completely new way when conducting their clinical skills building training through role play. This experience is consistent with what Robinson and Reiter (2016) describe as an “irresistible urge” (p. 159) to use open-ended questions to gather detailed psychosocial data rather than focus on a functional analysis of the problem and gather a snapshot of a patient’s life in 5–10 minutes. However, students also commented on their perceptions of transferring existing therapy skills to the integrated primary care setting. For example, one student pointed out that it seemed as if it would be a relatively easy transition for a health psychologist or a psychologist trained from a behavioral theoretical perspective to adapt their skill set to this setting.
Practice Management

In addition to differing expectations versus reality in the context of the integrated primary care model and clinical skills required, students also wrestled with some aspects of practice management. For example, students acknowledged the importance of charting during appointments, but struggled with how that might negatively affect the patient–provider rapport and ability to convey empathy and understanding through nonverbal communication. According to Robinson and Reiter (2016), this real-time charting is important for practice management considerations, and they recommend that behavioral health consultants complete at least 75% of the chart note by the end of the visit to preserve the specificity, accuracy, and brevity of the communication. Thus, despite their concerns about charting in this manner, students recognized this is a necessary practice management skill that needs to be developed and practiced.

Ethics

Students also reflected on how different professions view certain common ethical difficulties, and how multiple professions view certain situations from different perspectives. For example, the topic of dual relationships and how they are viewed from different professions sparked interest. Family physicians practice under the American Medical Association’s (AMA) Code of Ethics just as psychologists practice under the APA Code of Ethics (AMA, 2012; APA, 2010). However, each profession approaches some situations differently. For example, physicians have a more accepting approach to dual relationships and treating peers compared to psychologists. Their different approach is not surprising, as the AMA code approves and even encourages treating peers. Primary care psychologists will likely encounter physician peers expecting to enter into a dual relationship, something that psychologists are cautioned against in the APA code. Ethical dilemmas such as this will be discussed in greater detail later in this chapter. However, it is helpful to note that these different perspectives were important for students to observe and discuss during the course.

In addition to these qualitative observations at the doctoral level, Possis et al. (2016) report on 24 individuals, including seven psychology practicum students, interns, and postdoctoral residents who participated in a Veterans Affairs immersion training program in Primary Care Mental Health Integration. These trainees made similar observations about working in the primary care as the students described earlier. Possis et al. (2016) report that trainees commented on the need to develop specific skills to work in primary care, the unpredictable nature and workflow, and the challenges of pacing and documentation.

As new primary care behavioral health training programs are being developed it is critical to consider the existing guidelines and competencies to steer the core educational and experiential content of those programs. Equally important is the ongoing evaluation of trainee experience; what they find important in their training can continue to be emphasized and honed to build the best training possible.

FOUNDATIONS THAT SHAPE CURRENT GUIDELINES AND PRACTICE

The field of primary care psychology is considered an emerging practice area. As such, only recently have specific professional guidelines for practice been published (McDaniels et al., 2014). However, there are overarching guidelines that inform broader psychological practice relevant in the primary care context that are important to understand and that inform the most recently published
professional guidelines (McDaniels et al., 2014) and will serve as a foundation for future guideline publications. The first of two documents providing these guidelines is the APA Guidelines for Psychological Practice in Health Care Delivery Systems (APA, 2013a). The document was prepared with the goal of clarifying the roles and responsibilities that psychologists should consider taking on in the increasingly diverse contexts that psychologists practice. The second document is the APA Ethical Principles of Psychologists and Code of Conduct (APA, 2010). Psychologists in primary care should also be aware that other professions’ ethics codes, such as the AMA Code of Ethics, will also be relevant in guiding clinical practice and resolving some ethical dilemmas. In this section, we will briefly summarize the APA Guidelines for Psychological Practice in Health Care Delivery Systems and highlight relevant sections of the APA Ethics Code. Further, adapting to the environment of primary care and some particularly common ethical dilemmas are discussed.

**APA Guidelines for Psychological Practice in Health Care Delivery Systems**

The APA Guidelines for Psychological Practice in Health Care Delivery Systems refers to aspirational, umbrella guidelines for psychologists operating in a variety of contexts (APA, 2013a). The document offers 10 guidelines spread throughout the following four categories: (1) distinct professional identity within the healthcare delivery system, (2) privileges, (3) integrative and collaborative care, and (4) competency. Although these guidelines are written to apply to many contexts, they also serve as a good prelude to specific primary care psychology competencies discussed later in this chapter.

**Professional Identity**

The first guideline on professional identity reminds psychologists to remain cognizant of ethical and legal obligations as members of a distinct and autonomous profession, to practice within the scope of competence as a psychologist (as required in the APA Ethics Code), and to avoid compromising judgment in response to pressure by other professionals or systemic factors. Second, psychologists should understand the internally and externally imposed expectations and requirements of the systems within which they practice. For example, they should become familiar with the environment, culture, and context of care delivery in order to understand when systemic problems should be addressed. Finally, psychologists should clarify their distinct roles and services and how these relate to those of other healthcare professionals. For example, psychologists in an integrated environment such as primary care must be able to explain how they can contribute in contexts such as prevention, diagnosis, consultation, treatment, rehabilitation, and end-of-life care.

**Privileges**

In the second category, privileges, guidelines include seeking appropriate staff appointments and clinical privileges within healthcare delivery systems including the highest level of staff membership, and developing an accurate understanding of medical staff categories. Further, the guidelines encourage psychologists to be involved in the development of institutional policies on the professional scope of practice and participation in service delivery. This may include an active participation in the administration of healthcare systems, such as department leadership and helping to set criteria for psychologist clinical privileges (APA, 2013a).

**Integrative and Collaborative Care**

The third category, integrative and collaborative care, is particularly relevant to primary care psychology. Within this category, three guidelines are discussed. The first encourages psychologists to function in multidisciplinary positions with diverse roles and responsibilities such as health promotion, disease
prevention, primary care, and acute and chronic care for medical conditions. Psychologists are also encouraged to take on a broad scope of practice including consultation with multidisciplinary teams, realizing that psychology training lends itself to helping in the following areas: enhancing communication with patients, observing behavioral change related to disease progression and symptom management, facilitating decision-making and problem-solving, adjusting practices for developmental, behavioral, or psychiatric conditions as well as cultural considerations, and involving family or other support systems. Overall, psychologists should seek ways to integrate their expertise in diverse aspects of patient care (APA, 2013a).

The second guideline is to promote optimal delivery of psychological services through effective and timely communication with other healthcare professionals. This is accomplished by learning and speaking the language of the systems in which the healthcare professional works, communicating psychological concepts within the language of the medical system and culture, and participating in team meetings, rounds, and case conceptualizations.

The third guideline is to strive to provide collaborative services in the broadest range of healthcare settings, and to apply expertise to the full spectrum of health issues. This requires being on the lookout for a wide range of settings within which patient care can be improved with psychology expertise (APA, 2013a).

**Competency**

In the final category, competency, the first guideline includes gaining and maintaining appropriately specialized competence. As mentioned previously, psychologists are increasingly called to take on a wide range of goals within today’s healthcare delivery systems. Therefore, they must keep abreast of scientific knowledge, skills, and literature relevant to their particular job roles and duties. Second, they are encouraged to offer their expertise in the administration and management of psychological and other professional practice within healthcare delivery systems (APA, 2013a).

**APA Ethical Principles and Code of Conduct**

In addition to reviewing the APA Guidelines for Psychological Practice in Health Care Delivery Systems, it is important to briefly review ethical principles for psychologists in all contexts. The APA Ethical Principles and Code of Conduct (Ethics Code) calls psychologists to conduct their professional work under the aspiration of five guiding principles: beneficence and nonmaleficence, fidelity and responsibility, integrity, justice, and respect for people’s rights and dignity (APA, 2010). Further, the code outlines 10 sections of Ethical Standards in the following areas: resolving ethical issues, competence, human relations, privacy and confidentiality, advertising and other public statements, record-keeping and fees, education and training, research and publication, assessment, and therapy (APA, 2010).

Although all 10 sections are important for psychologists to understand and abide by, ethical issues in primary care are particularly likely to fall into the categories of privacy and confidentiality, therapy (for informed consent), and human relations (for multiple relationships). However, navigating the ethical foundations of these rather basic ethics code tenets is not always simple in primary care. For example, members of multiple professions work together in the same setting, some following ethics codes with different protocols than psychologists for informed consent and confidentiality as well as other ethical issues (Hudgins, Rose, Fifield, & Arnault, 2013). At times, psychologists may find themselves unprepared for the nuances of primary care ethical dilemmas, in part because the APA Ethics Code has assumed specialty mental health care as the baseline and gaps remain between what the code can provide and what primary care psychologists need (Kanzler et al., 2013). The following three sections will briefly address these common ethical concerns as well as introduce a decision-making model that can assist psychologists in critically evaluating a variety of actions that could be pursued.
Informed Consent
The need for informed consent is described by the APA Ethics Code in Section 3 (Human Relations) and Section 10 (Therapy). In Section 3, the requirement for informed consent is established, and in Section 10 further information regarding informed consent for therapy is outlined (APA, 2010). In specialty mental health care, informed consent is usually accomplished verbally and in writing at the first appointment, making sure that patients understand and agree to the proposed treatment as well as the limits of confidentiality (APA, 2010).

However, informed consent within primary care may differ based on the role of the psychologist and level of team integration. Some primary care psychology standard operating procedures such as those detailed in the US Air Force’s integrated behavioral health manual (Air Force Medical Operations Agency, 2014) are setting the precedent that a separate consent process similar to what occurs in specialty mental health settings is not necessary when the psychologist is acting as a consultant to the primary care provider and is considered a primary care treatment team member. However, informed consent should be an ongoing discussion between the patient and the psychologist in primary care settings to make sure that they understand the psychologist’s role in the primary care behavioral health service delivery (Hodgson, Mendenhall, & Lamson, 2013).

Confidentiality
In addition to informed consent, confidentiality is an ethical concern that often presents difficulties for primary care psychologists. The nature of primary care clinics, which include a whole-family treatment focus, a team-based approach to care, and a longitudinal care delivery model make it particularly prone to presenting ethical challenges in confidentiality (Reiter & Runyan, 2013). More challenges in primary care also exist because of multiple providers treating the same patient in a shared medical setting (Hodgson et al., 2013). Further, in co-located integrated primary care settings, rules of confidentiality are generally guided by the AMA Code of Medical Ethics, which most primary care providers abide by (Hodgson et al., 2013). Thus, primary care psychologists may have difficulty navigating between multiple governing codes when it comes to confidentiality issues in collaboration and record-keeping (Hodgson et al., 2013). Further, they will need to delicately balance patient privacy protection with sharing enough information for decision-making among care team members (DiTomasso et al., 2010).

Human Relations—Multiple Relationships
Primary care psychologists are likely to experience an interprofessional environment where providers adhere to different ethical standards. These differing standards are likely to cause ethical dilemmas related to multiple relationships. Kanzler et al. (2013) provide an example that highlights this challenge by describing a situation in which a peer (primary care physician) approaches a primary care psychologist seeking mental health care. The APA Ethics Code recommends caution against entering into a multiple relationship by treating a peer (particularly if objectivity could be impaired), while the language in the AMA Code of Medical Ethics seems to encourage physicians to do so (AMA, 2012). Because it is common for physicians to treat their peers, it is likely that primary care psychologists will experience physician peers approaching them for treatment more often than would happen in specialty mental health care (Kanzler et al., 2013).

Ethical Decision-Making in Primary Care
Working on an interprofessional team with different professional codes will challenge psychologists to balance the APA Ethics Code, other professional codes (e.g., AMA), as well as the clinic culture when deciding how to act. It will not be uncommon for psychologists in primary care to face gaps in the APA Ethics Code that demand careful thought on how to remain true to the APA code while
adapting appropriately to the environment (Kanzler et al., 2013). The authors provide a series of recommended steps for resolving such ethical dilemmas in primary care:

1. Consider your context;
2. Consider ethics code;
3. Determine risk/benefits;
4. Critically interpret/implement your code;
5. Consider different perspectives;
6. Identify other’s expectations;
7. Clarify your role;
8. Discuss your concerns.

Using these recommendations to guide decision-making in situations that may be ethically complicated may help psychologists consider unique aspects of working within the primary care context and engage in ethically defendable clinical decisions that benefit the patient.

Together, the APA Guidelines for Psychological Practice in Health Care Delivery Systems and APA Ethical Principles and Code of Conduct provide a foundation for guiding psychology practice in primary care. However, the specific core competency skills required for effective practice within primary care are guided by current and future professionally established competencies.

GENERAL COMPETENCIES FOR WORKING IN PRIMARY CARE SETTINGS

As mentioned earlier, psychology has increasingly moved toward an emphasis on competency-based education (Hatcher, Fouad, Campbell, McCutcheon, Grus, & Leahy, 2013). APA has established core competencies for health service psychology (i.e., clinical, counseling, and school psychology) organized under six domains: science, professionalism, relational, applications, education, and systems (Health Service Psychology Education Collaborative [HSPEC], 2013). These competencies have shaped the standards of accreditation and required competencies (i.e., research; ethical and legal standards; individual and cultural diversity; professional values, attitudes, and behaviors; communication and interpersonal skills; assessment; intervention; supervision; consultation and interprofessional skills) for doctoral, internship, and postdoctoral programs (APA, 2015b).

Some guidelines exist for evidence-based care across mental health, substance use, and clinical health psychology; however, psychologists working in primary care need an additional set of competencies to guide clinical practice, education, research, and training. Recently these foundational competencies for primary care psychology practice have been established.

Competencies for Psychologists in Patient-Centered Medical Homes

Nash, Khatri, Cubic, and Baird (2013) discuss essential competencies for psychologists within a patient-centered medical home (PCHM) environment. Nash and colleagues comment on the importance of establishing competencies for primary care, asserting their importance in assisting psychologists in taking advantage of growing opportunities for work in primary care and to be prepared for a practice environment very different than specialty mental health settings. Further, Nash et al. describe a primary care psychology ethic that promotes the values of appreciation, openness, and willingness to
engage. Fitting within this primary care ethic are two overall categories of competency: *foundational* and *functional*. These two categories, according to Nash et al., are necessary to inform psychology training curriculum development, guide research, serve as a foundation for psychologists seeking to establish credibility within PCMHs, and inform postdoctoral and board certification processes.

*Foundational competencies* are relevant to multiple psychology roles within primary care, and prepare psychologists for seamless adaptation to the primary care context. Further, these foundational competencies help primary care psychologists effectively perform their roles, which are greatly expanded from specialty care. The foundational competency category includes the following: (1) an approach to care that is population-based, patient-centered, and evidence-based; (2) an attitude of team-based care that acknowledges the importance of systems, teamwork, and communication; (3) professional behavior and identity as a psychologist; and (4) reflective practice, diversity, and ethical, legal, and professional practice standards (Nash et al., 2013).

The *functional competencies* described by Nash et al. (2013) articulate the skills psychologists working in primary care need in order to meet the foundational competencies. The functional competency category includes the following: (1) assessment, intervention, and consultation; (2) supervision and training; (3) management and administration in patient care and practice operations; and (4) research methods for quality improvement and outcomes initiatives. An example of the assessment, intervention, and consultation competency includes brief and evidence-based assessment and intervention methods along with good consultation skills for providing timely expertise to PCMH clinicians. APA has further defined these competencies (i.e., McDaniel et al., 2014).

**APA Task Force and Interorganizational Work Group**

A task force established by the APA’s Board of Educational Affairs in 2011 determined that there was no broadly accepted list of competencies to guide the education and training required for psychologists working in primary care. In response to the lack of formal articulation of such competencies, an Interorganizational Work Group (IWG) was formed and produced a document (built initially off work completed by the Substance Abuse and Mental Health Services Administration in 1998) regarding the establishment of competencies for psychology practice in primary care (McDaniel et al., 2014). This document, titled *Competencies for Psychology Practice in Primary Care*, reflects significantly updated and expanded content on necessary competencies in primary care (McDaniel et al., 2014).

The IWG authored a comprehensive report on the outcome of their work, established as an initiative of then APA President Suzanne Bennett Johnson. These competencies are established for broad use in the following areas: informing graduate psychology education and training programs, assisting students and practitioners in evaluating available primary care training and certificate programs, and informing policy makers and other health professionals about primary care psychologist competencies (APA, 2013b; McDaniel et al., 2014).

The six clusters of competencies described in McDaniel et al. (2014) include: science, systems, professionalism, relationships, application, and education. For each cluster, McDaniel et al. offer specific competencies, essential components of the competency, and sample behavioral anchors to demonstrate examples of competence. It is important to note that the IWG does not expect all primary care psychologists to be experts in all of the competencies. Rather, primary care psychologists should at least be familiar with them. The following section will briefly review each cluster.

**Science**

The science competency cluster is founded on the principle that the sustained integration of science and practice is important in the identity of psychology as a profession. Within primary care, two
general areas make up the science cluster: the importance of a scientific foundation related to the biopsychosocial approach, and the importance of research and evaluation. The scientific foundation of primary care psychology encompasses components such as knowledge of the biological, cognitive, and affective components of health and illness. Included in this competency is a greater knowledge of topics such as physiology, basic pharmacology and psychopharmacology, epidemiology, and public health policy. The second component, research and evaluation, refers to competencies such as the ability to conduct research as part of an interprofessional team and to select and monitor outcome measures for program evaluation. It is important to note that these scientific knowledge-related competencies are greater than what is normally expected of psychologists in general mental health practice (McDaniel et al., 2014).

Systems
The system cluster acknowledges that psychologists in primary care work in a constantly changing and evolving system at local, regional, and national levels. Three components comprise the system-related competencies: understanding interprofessional systems of care, leadership and administration, and advocacy. Skills within this competency include the ability to promote effective communication among staff operating in an interprofessional environment at multiple levels (staff, clinical, and organizational); an appreciation for the complex environment that surrounds patient care; and an understanding of how healthcare policy effects clinical, operational, and financial aspects of health care (McDaniel et al., 2014).

Professionalism
The professionalism cluster refers to a wide range of professional attributes, including professional values and attitudes; individual, cultural, and disciplinary diversity; ethics; and reflective self-practice, self-assessment, and self-care. Part of professional comportment in primary care includes practice management flexibility—that is, managing time to facilitate interruptions and match a fast-paced primary care environment with limited space resources. Further, it entails navigating the unique ethical concerns related to informed consent and confidentiality (McDaniel et al., 2014).

Relationships
The relationship competency cluster contains items related to interprofessional collaboration and contributions to effective team functioning. According to McDaniel et al. (2014), as well as authors of other competency documents (Air Force Medical Operations Agency, 2014; Nash et al., 2013), relationship building with other primary care providers across professions is essential to fostering an environment of collaboration.

Application
The application cluster encompasses four discrete competency areas: efficient practice management, assessment, intervention, and clinical consultation. This is another area of significant overlap with other competency documents. Efficient practice management refers to adaptation to a fast-paced primary care environment, and delivering care in a way that fits the individuals served within the financial, time, and service delivery constraints of the clinic. Primary care health teams often follow patients over a long period of time, thus the primary care psychologist may provide a variety of different services over time as patient and family needs change. Assessment should be continuous, flexible, and dynamic. When assessment measures are used, the psychologist has a responsibility to select measures that are appropriate for answering the diagnostic or functional impairment questions, to consider the reliability and validity of measures, and to consider the strengths and limitations of measures, particularly within the context of a primary care environment. Interventions are not just focused on mental health conditions but also on health behaviors, adherence to chronic disease
prevention and treatment, and conditions that contribute to high healthcare utilization. Primary care psychologists are also in a good position to facilitate bidirectional communication—helping the provider understand the patient’s life context and helping the patient understand biopsychosocial factors relevant to his or her health. Finally, clinical consultation refers to the ability to respond to clinical questions in a collaborative manner, offering recommendations that are tailored to the primary care pace and environment. Further, recommendations during consultation are specific, evidence-based, and can be implemented in a brief amount of time (McDaniel et al., 2014).

Education

Finally, commitment to education and training is important for psychologists in primary care. This competency is expressed through a commitment to developing learning opportunities for psychology students in primary care as well as through educating other healthcare providers about the role of psychology in primary care (McDaniel et al., 2014).

Overall, the Competencies for Psychology Practice in Primary Care serves as a foundational guide. Whether training is occurring at the graduate or the professional level, it details what competencies individuals should demonstrate when working in primary care settings.

Competencies in Special Populations

In addition to the two broadly applicable documents mentioned earlier, competency guidelines applicable to more specialty populations such as the Department of Defense and substance abuse populations have also emerged.

Department of Defense

The Department of Defense (DoD) serves over 3 million individuals (active duty members and their families and retired active duty members and their families) enrolled by name to primary care providers in primary care clinics on military installations. In August 2013, the DoD established guidance for the integration of behavioral health personnel in primary care through DoD Instruction (DoDI) 6490.15, Integration of Behavioral Health Personnel (BHP) Services Into Patient-Centered Medical Home (PCMH) Primary Care and Other Primary Care Service Settings. This instruction outlines the minimum requirements for internal behavioral health consultants (IBHCs) within military medical treatment facilities. Further, DoDI 6490.15 lists IBHC core competencies in the context of a training approach for IBHCs, and covers the following six categories: (1) clinical practice knowledge and skills, (2) practice management skills, (3) consultation skills, (4) documentation skills, (5) administrative knowledge and skills, and (6) team performance and skills.

Substance Abuse and Mental Health Services Administration Competencies

The Substance Abuse and Mental Health Services Administration and Health Resources and Services Administration (SAMHSA-HRSA) Center for Integrated Health Solutions published a competency document to help prepare the behavioral health workforce for delivering integrated care (Hoge, Morris, Laraia, Pomerantz, & Farley, 2014). The SAMHSA-HRSA document outlines nine core competencies to be used for shaping training, informing job descriptions, recruiting and orienting employees, and assessing performance (Hoge et al., 2014). Several of the same broad categories of competency are covered, including strong communication skills with patients and interprofessional team members, teamwork, clinical skill in assessment and intervention, and systems-oriented practice. However, the SAMHSA-HRSA competencies, particularly in their detailed description, are tailored toward comorbid substance use and mental health conditions. For example, the competencies include care planning
and coordination skills with a social work focus, and intervention skills specific to treatments of comorbid substance use and mental health conditions (Hoge et al., 2014).

Summary

Each of the major competency documents described earlier lists a number of broad categories and there is significant overlap among the documents. A shared belief is evident that competency for primary care work requires not only clinical skills of assessment and intervention in a fast-paced primary care environment, but also broader skills to function effectively in a population-based framework of care. These broader skills include functioning as an interprofessional team member and effectively and efficiently providing expert consultation to other primary care providers. Further, there are increasing opportunities for psychologists to be involved in the management and administration of primary care clinics; therefore, it is important that primary care psychologists strive to obtain competency in practice management roles and in developing policies and procedures for integrating behavioral health care in primary care clinics.

Examples of Training to Competencies

In addition to providing an overview of the competency documents, it is useful to briefly mention the experiences of three programs that trained students to these competencies. Larkin, Bridges, Fields, and Vogel (2015) acknowledge the importance of training to the shifting environment of healthcare delivery systems, and describe three programs using the six competency clusters established by APA's IWG (McDaniel et al., 2014). These programs are the Doctoral Program in Clinical Psychology at the University of Arkansas, the Doctoral Psychology Internship Program at the West Virginia University School of Medicine, and the Postdoctoral Fellowship Program at the Consortium for Advanced Psychology Training affiliated with Michigan State University.

Larkin et al.’s (2015) article provides examples of how these programs trained students in both skill-based and knowledge-based competencies for integrated behavioral health care as well as their plans for future development. Examples of training steps taken toward knowledge-based competencies include seminars and didactics on integrated care topics, grand rounds attendance with the family medicine department, lectures covering general medical practice issues as well as behavioral health topics, a highly structured 12-month didactic clinical health psychology core curriculum, and a 12-month series of lectures on key principles of pharmacology relevant to clinical health and primary care settings. These programs took care to add didactic content as necessary to fulfill training across the six competency clusters established by APA's IWG. Examples of training to obtain skills-based competencies included practicum placements in integrated primary care clinics, supervising and providing feedback on medical student interactions at a medical simulation center, and hands-on training in a Primary Care Behavioral Health model including introductions via “warm handoff.”

SPECIFIC CLINICAL COMPETENCIES FOR WORKING IN INTEGRATED PRIMARY CARE

It is beyond the scope of this chapter to summarize the competencies required for targeting all possible problem presentations in an integrated primary care setting. Yet, it is important for psychologists to consider their competence and to determine whether they appropriately and ethically can target behavioral health concerns within the primary care context. Multiple books, including Hunter, Goodie, Oordt, and Dobmeyer (2017), Robinson and Reiter (2016), and Talen and Valeras (2013),
describe assessment and intervention strategies for primary care settings in far more detail than is provided below. Often there are no established competencies or even agreed upon evidence-based interventions for behavioral health (e.g., depression and anxiety), clinical health psychology (e.g., obesity and chronic pain), and substance misuse. However, there are common conditions and skills that those who choose to work in primary care settings must know to be successful.

**Behavioral Health**

The cornerstone of training in psychology is learning methods for the assessment and treatment of a range of behavioral health problems, particularly depressive and anxiety disorders. It remains unclear exactly what assessments and interventions work best for many behavioral health concerns in the primary care environment. Most interventions have relied on modified techniques from acceptance and commitment therapy, behavioral, cognitive, solution-focused, and problem-solving therapies (Hunter et al., 2017; Robinson & Reiter, 2016; Strosahl, Robinson, & Gustavsson, 2012); therefore, it is helpful for individuals to have a background in these treatment modalities before working in primary care settings. Additionally, having functional analysis skills are critical for guiding assessment and treatment development. Although training in these modalities is necessary, they are not sufficient for knowing how to deliver adapted primary care appropriate evidence-based interventions in primary care. Following are examples of the types of skills that can be helpful for targeting common behavioral health presentations.

**Assertive Communication**

Poor communication often interferes with functioning across a variety of domains. Helping patients to distinguish between passive, assertive, and aggressive communication and how to engage in assertive communication skills (e.g., nonverbal behaviors, word choices, “I” statements), can help them to manage a variety of situations.

**Behavioral Activation**

Behavioral activation offers a straightforward and effective strategy for beginning to target depressive symptoms (Cuijpers, van Straten, & Warmerdam, 2007). Developing a plan with patients to engage in valued activities can lead to significant improvements in functioning.

**Managing Cognitions**

Whether psychologists are going to use techniques derived from cognitive theory (e.g., cognitive disputation) or methods derived from Acceptance and Commitment Therapy (e.g., cognitive defusion), psychologists should have training in methods for targeting unhelpful patterns of thinking. Depending on the target of the treatment, using strategies such as worry logs and exposure hierarchies may also be valuable for targeting problems (e.g., generalized anxiety, panic).

**Clinical Health Psychology**

Clinical health psychology is a recognized specialty by the Council of Specialties in Professional Psychology that

> applies scientific knowledge of the interrelationships among behavioral, emotional, cognitive, social, and biological components in health and disease to the promotion and maintenance of health; the prevention, treatment and rehabilitation of illness and disability; and the improvement of the health care system.

(APA, 2016)
The skills typically associated with clinical health psychology, such as health behavior change (e.g., tobacco cessation, weight loss, increasing physical activity) and helping to manage health conditions (e.g., cardiovascular disease, chronic pain, diabetes, sleep disorders), are extremely useful in primary care settings. However, working in primary care settings does not, by itself, define a practitioner as a clinical health psychologist.

The Council of Clinical Health Psychology Training Programs has developed a Clinical Health Psychology Competencies Rating Form and a specialty taxonomy that defines levels of opportunity (i.e., major area of study, emphasis, experience, and exposure) for training (Council of Clinical Health Psychology Training Programs, 2016). These levels of opportunity are similar to the levels that are discussed for integrated primary care training later and are summarized in Table 4.1 for doctoral level training. Ideally those who are working in primary care settings will have had at least some “exposure” or “experience” level training in clinical health psychology. Those working in primary care should be familiar with the common health conditions (e.g., chronic pain, diabetes, hypertension, insomnia) that present in primary care and respond well to behavioral intervention. It is valuable to have a working understanding of the physiological mechanisms associated with these conditions as well as the common medical treatments (e.g., medications, procedures, surgical interventions). There are several skills, described later, associated with a broad range of health conditions that are useful for those working in primary care.

### Table 4.1 Definitions of Level of Primary Care Psychology Experience Used in the American Psychological Association’s Directory of Training Programs With Training Opportunities in Primary Care Psychology

<table>
<thead>
<tr>
<th>Level</th>
<th>Doctoral</th>
<th>Internship &amp; Postdoctoral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major</td>
<td>Highest level of education and training opportunities to develop the knowledge, skills, and attitudes at a level of intensity and involvement that allows the development of those competencies (e.g., 2–3 years of didactics, supervised experience, and dissertation or research project in primary care)</td>
<td>Two or more major rotations (16–40 hours per week) in primary care</td>
</tr>
<tr>
<td>Emphasis</td>
<td>Structured, in-depth opportunities for knowledge acquisition and practical experience in a specialty area (e.g., 4 courses and 2 practica)</td>
<td>One major rotation (16–40 hours per week) in primary care</td>
</tr>
<tr>
<td>Experience</td>
<td>Type and intensity of training opportunities goes beyond acquainting a student with a specialty (e.g., 1–2 courses and a practicum)</td>
<td>One or more minor rotations (1 day a week) in primary care</td>
</tr>
<tr>
<td>Exposure</td>
<td>Education and training opportunities are limited in type and intensity, but provide an opportunity to acquaint students with specialty area (e.g., 1–2 courses)</td>
<td>Limited introduction (1–4 hours per week) to primary care</td>
</tr>
</tbody>
</table>

*Note. Adapted from the American Psychological Association’s directories of doctoral, internship, and postdoctoral training programs with training opportunities in primary care psychology, available from [http://www.apa.org/ed/graduate/primary-care-psychology.aspx](http://www.apa.org/ed/graduate/primary-care-psychology.aspx)*
Promoting Physical Activity

It is recommended that adults engage in 150 minutes a week in moderate intensity or 75 minutes a week of vigorous intensity aerobic activity a week (US Department of Health and Human Services, 2008). Additionally, adults should engage in muscle strengthening exercises at moderate or vigorous intensity two or more days a week. Together, these physical activity guidelines promote both physical and behavioral health. Supporting these recommendations is critical when working in primary care settings.

Sleep Behaviors

In individuals who demonstrate problems with insomnia (i.e., difficulty falling asleep or staying asleep, resulting in impaired functioning) there are well-established interventions that have been shown to be effective. Most importantly, stimulus control (e.g., creating a bedroom environment conducive to sleep, using the bed and bedroom for sleep and sex only; avoiding lying awake in bed), sleep restriction (e.g., setting a consistent wake-up time and limiting time in bed to the amount of time spent sleeping), and sleep hygiene (e.g., avoiding behaviors, such as caffeine or alcohol use, that would interfere with sleep) are effective strategies for improving sleep in primary care (e.g., Falloon, Elley, Fernado, Lee, & Arroll, 2015; Goodie, Isler, Hunter, & Peterson, 2009).

Relaxation Skills

Physiological arousal and the stress response are related to a broad range of health conditions. Being able to teach simple relaxation strategies within the context of brief appointments is useful for targeting a variety of presenting problems (e.g., chronic pain, hypertension). Simple deep breathing techniques, abbreviated strategies for progressive muscle relaxation, or guided imagery are useful tools that patients can use in a variety of contexts.

Goal Setting and Pacing

Helping patients establish goals that are Specific, Measurable, Attainable, Relevant, and Time-based (i.e., “SMART” goal setting—an acronym for the component parts of effective goal setting) can help patients achieve a variety of behavior changes (e.g., increasing physical activity). Helping patients to set these goals and then having the primary care team monitor progress toward these goals encourages accountability and sustained progress. It is also critical for some patients (e.g., chronic pain) to learn how to pace activities using goal setting. Patients may need to learn to engage in activities for specified periods of time despite discomfort or pain. Learning to pace activities (e.g., refraining from doing too much when feeling little or no pain, continuing to engage in activities when in pain) can be critical for promoting functional improvement.

These are some of the fundamental skills that are helpful for targeting a range of health conditions in primary care settings. Trainings focusing on these skills along with education regarding the breadth of common medical conditions seen in primary care greatly enhance an integrated behavioral health providers’ value within this setting.

Substance Misuse

Alcohol and prescription medication misuse are common problems in primary care (SAMHSA, 2014). Approaches to these problems may be taught outside of the traditional curriculum for other behavioral and clinical health psychology problems. For alcohol misuse, it is important for psychologists to have the skills to use appropriate screening and assessment measures for primary care like the Alcohol Use Disorders Identification Test Consumption (AUDIT-C). Psychologists should know the diagnostic
criteria for alcohol use disorders (AUD) and be able to ask specific questions to help confirm whether a patient is positive on any of the AUD criteria. The ability to provide primary care services to those with mild AUD severity is a must. Psychologists should be able to advise patients and the rest of the team on a range of services, both in and out of primary care and the potential pros and cons of those options. Evidence-based primary care intervention skills should include ability to provide education on alcohol, to include discussion of standard drink size, population norms on alcohol consumption, and risky drinking levels. The psychologist should be able to help develop interventions consistent with paced drinking, harm reduction, stimulus control, and anxiety and stress management skills. See Hunter et al. (2017) for more detailed information.

Similarly, there are a number of skills a psychologist should have to assist the patient and team with managing prescription medication misuse. The psychologist should be able to effectively screen for patients that are at increased risk as well as current prescription medication misuse—for instance, using the evidence-based questions like:

1. Do you drink alcohol? If the answer is “Yes,” do further screening for alcohol misuse.
2. In the past year, have you used an illegal drug or used a prescription medication for nonmedical reasons?

A response of at least one time to this single-question screening test for drug use in primary care has been shown to have good sensitivity (100%) and specificity (73.5%) for detecting those that might have a drug use disorder (Smith, Schmidt, Allensworth-Davies, & Saitz, 2010). Knowing the potential signs of misuse (e.g., losing medication, running out of medication earlier than expected based on number prescribed and dose per day) is an important skill, as well as is assisting the PCP and the team in recognizing individuals who may be misusing medication. Other helpful skills might include reviewing the pros and cons of using potentially addictive medications, developing an agreement or plan to take medication on a schedule, or following up with the PCP monthly to refill the prescription to minimize abuse or dependence risk. Being able to assist with other self-regulatory strategies for stress reduction, as well as for managing and coping with chronic pain and discomfort, should also be in the primary care psychologist’s skill set.

**Primary Care-Specific Skills**

Primary care is a vastly different culture in comparison to the average psychologist’s operating environment. This new culture includes different language, goals, pace, population focus, and team delivered/managed services. In order to be successful, psychologists must adapt to this new culture in order to be accepted as part of the team. Hunter, Goodie, Oordt, and Dobmeyer (2009) give a detailed description of primary care specific skills for successful integration that fall under the areas of (1) getting your foot in the door, (2) being a team member, (3) building key relationships, (4) being accessible/available, and (5) learning the primary care culture. The following is a summary of skills in these areas.

Getting your foot in the door refers to being able to effectively market your services. Psychologists need to have skills that involve selling what they have to offer in a way that fits with what is valuable to the team. It is important to identify the priorities of the clinic and establish services that immediately impact challenges PCPs have. Being able to discuss/demonstrate how primary care psychology services will improve the overall impact on health outcomes, improve efficiency and effectiveness of clinic service delivery, and potentially manage cost more effectively can be important (see Corso, Hunter, Dahl, Kallenber, & Manson, 2016 for additional detail).
Being a team member can be a challenge for many psychologists who are used to working alone, one-on-one with patients, where only the patient and the psychologist are privy to the ongoing clinical problems, interventions, and rate of improvement. Collaborating with primary care team members in meeting the needs of the patient is one of the keys to successful integration. It is important that psychologists realize they have two customers; although the patient is obviously an essential customer, the primary customer is the PCP and support staff. Psychologists are there to help the team manage its patients. It is important for psychologists to be open to accept any referral and find the best way to assist the team and that individual. There is likely going to be a need for most psychologists to expand their clinical skill set to include areas related to health behavior change for problems like tobacco use, poor physical activity, unhelpful eating behaviors, sleep difficulties, chronic pain, and medical adherence challenges. Being skilled in providing useful feedback to the referring team member typically by the end of the day, via notes and verbal feedback, is a must. Verbal feedback generally takes less than 60 seconds and includes significant findings and recommendations regarding patient interventions that the PCP and team members can support and evaluate in follow-up appointments.

Building key relationships is important in gaining immediate acceptance and having an impact in the clinic as well as sustaining a level of close, integrated teamwork. Psychologists should be able to build key relationships with administrative staff, technicians, nurses, and PCPs. This involves interacting with them during lunch, parties, between patients as well as direct patient care. Psychologists should get to know the staff, their roles, history, and challenges they face. They should shadow a PCP, nurse, medical technician, and administrative staff for half of a day. Such actions show interest to the rest of the team, gets face time that cannot be obtained by just doing clinical work, and allows psychologists to gain a better understanding of team members’ day-to-day activities. Those without these skills often make the mistake of isolating themselves in their office when not seeing patients, keeping their door closed, and not participating in team activities. The idea is to be perceived as a team member—a friendly, accessible, go-to helper that the team can access whenever needed. Gaining buy-in and support from informal leaders (i.e., those individuals who may not have formal leadership positions, but who have significant influence on the behavior of others), by assisting them in performing their jobs more effectively, can also do a lot. Working with these informal leaders is often the best way to promote primary care psychology services—as they see your helpfulness, they spread the word and help integrate you into the team.

Psychologists need to be accessible to anyone in the clinic at any time, potentially carrying a cell phone to be easily accessible by any clinic staff. Psychologists should expect and encourage staff to interrupt them if needed when they are with a patient. They should be available to step out of lunch, a break, or after a standard end of the day clinic to see emergent or acute needs. Psychologists should also have the skills to see unscheduled patients during the day, similar to the expectations of PCPs.

Psychologists must be able to adapt to primary care and not expect primary care to adapt to them. This involves skills like using primary care terms and language. For example primary care has patients and appointments, not clients and sessions. Using a “foreign” language keeps the psychologist as a cultural outsider. Psychologists should avoid using psychological jargon and keep up with medical journals and texts to learn the medical terms and latest science in which their team members are immersed. Patient appointment times should mirror PCP appointment times (e.g., 15 and 30 minutes) to ensure prompt access to services. Psychologists need to have the mindset and skills to provide brief, focused interventions to large numbers of patients in the clinic, instead of providing comprehensive care to a small number of patients as they are accustomed to in specialty mental health.
MODELS FOR TRAINING

Training for psychologists working in primary care has been infused throughout the career cycle; there are opportunities for training at the doctoral, internship, postdoctoral, and professional levels. The SAMHSA-HRSA Center for Integrated Health Solutions (http://www.integration.samhsa.gov/) maintains a variety of resources for behavioral health providers working in integrated primary care settings, including information and links to relevant education and training opportunities. The site represents a good starting place for those who are looking for up-to-date information regarding training opportunities.

The APA maintains lists of doctoral, internship, and postdoctoral programs offering training in primary care psychology; these lists are available at http://www.apa.org/ed/graduate/primary-care-psychology.aspx. These lists not only include contact information and accreditation status for the sites, but also classify the level of primary care psychology experience and list the models of primary care psychology offered. Table 4.1 summarizes the descriptions used to define whether the level of experience is classified as major, emphasis, experience, or exposure. Data for doctoral programs are based on results of a survey emailed to program directors, while internship and postdoctoral data are based on information available on the Association of Psychology Postdoctoral and Internship Centers’ (APPIC) website (http://www.appic.org). These lists provide a good starting place for identifying programs offering some training in primary care. However, the type of primary care psychology and the model of integration are inconsistent and poorly defined. For example, primary care psychology models such as “multidisciplinary, team-based,” “biopsychosocial,” or “collaboration” do not provide a clear understanding of the types of training or the competencies that will be learned in these programs.

Graduate Programs

Integrated primary care training in graduate programs has increased over the past 20 years. In APA’s directory available in March 2016, 44 doctoral programs reported having some training in primary care. Two programs reported a major level of experience, 10 an emphasis level, 26 an experience level, and six an exposure level. This is likely an underestimate of the number of psychology programs currently offering training in primary care. Increasingly programs are recognizing the value of training in integrated primary care and offering courses and practicums related to the primary care setting. In graduate programs the goal is often to expose students to the concepts (e.g., population health, stepped care) that influence the delivery of care in the primary care environment and to help students begin to develop skills necessary for providing care within this context.

The American Psychological Association’s Society for Health Psychology (i.e., Division 38) has prepared an entire curriculum for teaching integrated primary care psychology to graduate students. Modules, complete with lesson plans, instructor’s guides, and presentation materials are available, for free, from the society’s webpage: http://www.health-psych.org/. The modules, which cover foundational (e.g., the context of primary care, core knowledge, and skills) and specific topic (e.g., chronic pain, depression, substance misuse) areas, are designed to be used by all faculty, regardless of their previous experience in primary care settings. Faculty can choose to create an entire course, or pick specific modules to present during an existing course. If implemented, these extraordinary resources will significantly expand the opportunities for graduate students to learn fundamental aspects of practicing in primary care settings.
An interesting development is the offering of a Doctor of Behavioral Health (DBH) at Arizona State University that is earned through online courses (https://chs.asu.edu/programs/schools/doctor-behavioral-health). The DBH program offers degree concentrations in clinical or management. The concentrations are designed to promote improved delivery and management of integrated behavioral health care. The clinical concentration is designed for individuals who are already licensed clinicians (e.g., with a master’s degree in counseling, social work, psychology), and the management concentration is designed for executives and managers involved in the implementation of integrated behavioral health programs. The program offers a method to gain training related to enhancing behavioral health in primary care.

**Internship/Residency**

In the APA directory that was updated on February 2016, 140 internship sites were identified as having some training in primary care. Thirty-four programs reported a major level of experience, 48 an emphasis level, 27 an experience level, and five an exposure level. Twenty-six programs did not indicate type of training. The number of internships offering primary care training, and the high number that are described as having a “major” or “emphasis” level of training, highlights the importance of internships in training psychologists for work in primary care settings. To truly appreciate and develop the skills and requirements for working in primary care settings, trainees must spend sufficient time in the primary care setting. Readings, books, and webinars are insufficient substitutes for training that involves providing care in these settings. To successfully create a competent workforce for integrated primary care settings, it is critical that these training sites adopt the primary care competencies and to ensure trainees are meeting them.

**Postdoctoral Training**

Postdoctoral training, which embeds the learner into the primary care setting, likely provides the best opportunity to master the competencies necessary for working in primary care. Seventy-two postdoctoral sites, according to the APA June 2015 directory, reported offering training in primary care psychology. Thirteen sites were described as offering a major level of training, four reported an emphasis level, three an experience level, and one an exposure level. The training levels for 51 sites were not identified. A postdoctoral experience that allows a trainee to spend the majority of time in integrated primary care settings while supervised is not realistic for most psychologists. However, those trained in this manner are most likely to master the competencies for primary care practice.

**Professional Trainings**

Integrated primary care workshops and trainings for the full range of professionals are offered often at professional conferences. Professional organizations including the APA (http://www.apa.org), Association of Behavioral and Cognitive Therapies (http://www.abct.org), Collaborative Family Healthcare Association (http://www.cfha.org), and the Society of Behavioral Medicine (http://www.sbm.org) offer institutes and workshops that teach the competencies for working in integrated primary care. Although these events may not be sufficient by themselves for ethical practice in integrated primary care settings, they often provide an understanding and basic skill set for initial work in these environments.
Certificate Programs

Institutions have created certificate programs designed to train professionals to provide assessment and interventions within the primary care setting. One example is the University of Massachusetts Medical School Center for Integrated Primary Care (http://www.umassmed.edu/cipc), which offers a certificate program in primary care behavioral health for behavioral health providers. The program, which historically was provided through live webinars, is now moving toward using prerecorded content, offers six days of training covering a range of topics, including how to create and manage an integrated care practice; how to target a range of health behaviors; chronic pain, serious mental illness, and substance abuse; how to work with underserved and diverse populations; and how to work with children and families.

ASSESSING FIDELITY TO SERVICE DELIVERY

Assessing psychologist fidelity to service delivery standards involves assessing a range of clinical, practice management, consultation, documentation, and administrative behaviors at the initiation of primary care psychology services as well as ongoing assessment to ensure that fidelity to important service deliverables does not fade over time. Optimally, assessing fidelity would involve (1) assessing documentation in the medical record to ensure that tasks are being completed according to standards/benchmarks, and (2) direct observation of the psychologist with patients, documentation of patient encounters in the medical record, and interaction with team members. Many psychologists do not like to be observed during their clinical activities, which is understandable. However, their discomfort is not an excuse to forgo observation. A red flag should go up when a psychologist resists or avoids being observed. They should be able to demonstrate adequate skills and be open to ongoing improvement based on expert observation and feedback. Assessing documentation in the medical record is something that should happen in any clinic; it will give the evaluator a quick snapshot on easily accessible medical record data. For instance, if the standard is to administer a screening or assessment instrument to every patient on every appointment, randomly pulling a certain percentage of psychologist appointments will allow an evaluator to easily tally the percentage of patients with documentation of those screenings in the record. It is a process metric—it allows one to determine if tasks are being done. However, it does not necessarily provide information on that task being done well or the information being documented accurately.

Appendix B contains a list of competency areas that are considered by the authors to be important for psychologists to have. Most of these competencies will require direct observation to confirm they are being done with fidelity—for instance, “Summary and Formulation” under Clinical Practice Skills. One could see in a clinical record that the psychologist documented “summarized understanding of patient concerns and patient agreed,” but one could only determine through observation if that was done in 1–3 minutes including key symptoms and functional impairment.

It has been the authors’ experience (JLG and CLH) that even when psychologists can meet necessary competencies demonstrated through role-play training, those competencies do not completely transfer to real-time primary care psychology work. When psychologists are able to demonstrate real-time primary care psychology competencies at the start, sometimes those competencies fade over time. As such, having an expert trainer observe and train competencies in real-time clinical work, as well as ongoing fidelity checks, ensures the primary care psychology services are being delivered with evidence-based fidelity. It also bolsters the chances that the results wanted from these services will occur.
One measure, the Primary Care Behavioral Health Provider Questionnaire (PPAQ), was developed as a self-report measure of behavioral health provider fidelity to integrated primary care practice (Beehler, Funderburk, Possemato, & Dollar, 2013). Although not a substitute for direct observation, the PPAQ can help to assess whether a behavioral health provider is engaging in behaviors that are consistent or inconsistent with integrated primary care practice.

SUMMARY

Practicing in primary care settings requires a breadth of necessary training and skill development to be effective. There is no perfect training system or pathway that can be used to train enough psychologists to fulfill the need to enhance behavioral health care in the environment where most people go for such care (i.e., primary care). The APA competencies provide the best guidance for ensuring that primary care psychologists are trained to a minimal level and have appropriate expectations for working in primary care. However, a larger question to be addressed in future discussions is how are those skills are assessed and maintained over time to ensure that psychologists meet their ethical responsibilities in this emerging area of practice.

AUTHORS’ NOTE

The opinions and statements in this chapter are the responsibility of the authors, and such opinions and statements do not necessarily represent the policies of the Department of Defense, the US Department of Health and Human Services, or their agencies.

The authors would like to thank Omni Cassidy, Amanda Gehrke, and Sarah McCreight for sharing their expectations and reactions to learning about working in primary care settings.

NOTES

1 Population-based care is defined as an approach that allows one to assess the health status and health needs of a target population, implement and evaluate interventions that are designed to improve the health of that population, and efficiently and effectively provide care for members of that population in a way that is consistent with the community’s cultural, policy, and health resource values.

2 A needs analysis is the process of identifying and evaluating needs in a community or other defined population of people. The identification of needs is a process of describing “problems” of a target population and possible solutions to these problems.

3 See Hunter, Goodie, Oordt, & Dobmeyer (2017) for examples of adapted evidence-based primary care appropriate interventions.

REFERENCES


APPENDIX A

SAMPLE SYLLABUS

TEXTS AND READING
Course readings will include select chapters from the books listed below, augmented by a selection of articles and reports assigned by each instructor or lecturer.

Required Texts

Selected Readings
The class readings will be made available in PDF format, at least two weeks prior to each lecture.

Recommended Texts

WEEKLY TOPICS AND ASSIGNMENTS

Week 1  Topic(s): Introduction and History and Foundational Concepts

Topics: Course introduction; history of integrated primary care; Core Concepts: biopsychosocial model, chronic care model, stepped care, PCMH

Readings: DiTomasso (Chap. 1)

History & Scope

Biopsychosocial Model

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.

**Chronic Care Model**


**Triple Aim**


**Stepped Care**


**Other Preclass Assignments:**

Familiarize self with the AHRQ Lexicon (i.e., skim).


**Week 2  Topic(s): Context of Primary Care**

**Readings:** DiTomasso (Chaps. 3, 6)


**Week 3  Topic(s): Models of Integrated Care and Barriers**

**Readings:** DiTomasso (Chap. 4)

**Models of Integrated Care**


**Barriers**


**Preclass assignment:**


**Week 4  Topic(s): Policies and Implementation of Integrated Primary Care in the DoD; Ethical, Diversity, and Military Considerations**

**Readings:** DiTomasso (Chaps. 2, 7, 8); Hunter (Chap. 4)

**Policies and Implementation in the DoD**

Department of Defense (2013, August). Integration of behavioral health personnel (BHP) services into patient-centered medical home (PCMH) primary care and other primary care service settings (DOD Instruction 6490.1S). Washington, DC: Author.


**Ethics, Diversity, and Military Considerations**


© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
Week 5  Topic(s): Training and Research

Readings:

**Training**


**Research**


**Preclass assignment:**

1. Review Competencies checklist
2. What are the pros and cons of integrated primary care compared to specialty mental health care? (5 pages)
3. Submit paper/presentation topic

Week 6  Topic(s): Integrated Care in the VA and the First Appointment: Assessment & Documentation

Readings: Hunter (Chaps. 2, 3)

**Integrated Primary Care in the VA**


**The First Appointment: Assessment & Documentation**


© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.


Other Preclass Assignments: Memorize Introductory Script, Review interview format

**Week 7 Topic(s): Core Intervention strategies**

**Readings:** DiTomasso (Chaps. 9–12); Hunter (Chap. 3)


**Other Preclass Assignments:**

Review (skim) the Air Force BHOP Manual.

**Week 8 Topic(s): Depression, Anxiety, & Sleep**

**Readings:** DiTomasso (Chaps. 16–18, 31); Hunter (Chap. 5)


**Week 9 Topic(s): Health Behaviors (eating, activity, tobacco)**

**Readings:** DiTomasso (Chap. 15); Hunter (Chap. 6)


**Week 10 Topic(s): Chronic Disease & Chronic Pain**

**Readings:** DiTomasso (Chaps. 24–26, 28); Hunter (Chaps. 9–11)


© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
Other Preclass Assignments: *All papers due before the start of class.*

**Week 11  Student Presentations & Role Plays**

**Readings** TBD

**Week 12  Student Presentations, Role Plays, & Future Directions**

**Readings:** TBD
APPENDIX B

SAMPLE TEMPLATE FOR COMPETENCIES

BEHAVIORAL HEALTH CONSULTANT (BHC) CORE COMPETENCY TOOL

**BHC: Date: Trainer: Training Phase: I or II or Other**

Certified BHC Trainer rates the BHC trainee skill level based on their observations of each dimension. Trainees must demonstrate ability to perform the *minimal* benchmark behaviors and/or other clinical knowledge/skill that is consistent with minimal benchmark. A **“Pass” rating on every element is required to satisfactorily complete training**.

For Phase I Training, only unshaded items will be rated. A trainer must observe a minimum of four new patient appointments (role play or in vivo).

For Phase II Training, all items (shaded and unshaded) will be rated. A minimum of eight patient appointments (in vivo, including five new patient appointments) and a minimum of five PCP consultative feedback interactions must be observed. All BHC chart notes for patients seen during Phase II training must be reviewed, along with an additional five chart notes from prior appointments.

Note. BHCs are expected to have foundational skills in evidence-based assessment, treatment, consultation, and documentation through their graduate education and postlicensure work experiences. These prerequisite skills are NOT comprehensively taught during the Phase I BHC training course. A strong foundation in these areas is typically necessary to successfully acquire and demonstrate the competencies required to function as a BHC as assessed during the Phase I and II evaluations.

---

**Dimension: I. Clinical Practice Knowledge and Skills**

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
</table>
| **1. Role definition:** Says introductory script smoothly; conveys the BHC role to all new patients. | 1a. Accurately describes per standardized script:  
   i) Who they are and their role in the clinic,  
   ii) How long the appointment will be,  
   iii) What will happen during the appointment,  
   iv) What types of follow up might occur,  
   v) That the appointment note will go in medical record,  
   vi) That the PCP will get feedback,  
   vii) Any reporting obligations. | | Fail | Pass |

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
### Dimension: I. Clinical Practice Knowledge and Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1b.</td>
<td>Delivers the script in 2 minutes or less.</td>
<td></td>
<td></td>
<td>Fail</td>
</tr>
<tr>
<td>1c.</td>
<td>If interrupted by the patient during the introductory script, the BHC answers questions and appropriately redirects to complete the introductory script.</td>
<td></td>
<td>Pass</td>
<td></td>
</tr>
</tbody>
</table>

**2. Rapid problem identification:**
- **Rapidly reaches agreement with the patient on identifying the primary problem.**

#### Sample Benchmark Behaviors

- **2.** Confirms and clarifies consultation issue and obtains initial patient engagement in addressing consultation issue within 60 seconds after completing the introductory script.

#### Sample Behavioral Anchors

- “It looks like Dr. Hunter would like me to assist the two of you in better managing or targeting your depressed mood. Is that what you see as the main problem or is it something different?”
- “I’ve read Dr. Hunter’s note but want to make sure we are all on the same page. Before we get started on our appointment, what is the problem that you’re here to address?”

---

### Dimension: I. Clinical Practice Knowledge and Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a.</td>
<td>Assessment of functioning:</td>
<td></td>
<td></td>
<td>Fail</td>
</tr>
<tr>
<td></td>
<td>Asks questions appropriate to the presenting problem about how the following are influencing the identified problem and patient’s functioning in</td>
<td></td>
<td>Pass</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Dimension: I. Clinical Practice Knowledge and Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>the home, social, work, recreational, and spiritual areas of life:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>i) physical condition/physical response,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ii) behaviors/habits,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) thoughts,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>iv) emotions,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>v) environment/social interactions.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b.</td>
<td>Assessment of symptoms:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uses assessment measures appropriate to primary care (BHM-20 for every patient every appointment). If additional measures (e.g., PHQ-9, GAD-7) are administered, they are appropriate for primary care, scored, interpreted, and documented correctly.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3c.</td>
<td>Asks about duration, frequency and intensity of physical sensations, behaviors/habits, thoughts, and emotions, as appropriate to presenting problem. Asks about biopsychosocial factors that coincided with onset of or change in symptoms.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• “How long have you been feeling depressed?”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• “How many days per week are you feeling down?”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• “When during the week have you noticed that this is NOT a problem or is NOT occurring?”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• “On a scale of 0–10 with 0= not feeling depressed at all and 10= the most depressed you’ve ever felt, over the last two weeks what would be the average number for how you’ve felt? . . . what is the highest it has been in the last two weeks? . . . what is the lowest it has been?”</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Dimension: I. Clinical Practice Knowledge and Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>3d.</td>
<td>Appropriately assesses and manages risk of harm to self/others:</td>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td></td>
<td>i) Assesses SI and HI verbally at every initial visit,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ii) At follow-up appointments, assesses SI and HI in accordance with Service-specific standards,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) Initiates appropriate disposition for patients with elevated risk, following Service-specific guidance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Does not routinely attempt to maintain higher risk patients in primary care consultation and makes efforts to elevate high-risk patients to specialty care in a timely manner.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. <strong>Problem focus</strong></td>
<td>4a. Limits problem definition/assessment (focuses on presenting problem). Does not assess other areas (except suicidal and homicidal ideation) until assessment of initial referral problem is complete and as time allows. Exceptions may be made when a patient clearly indicates a preference for talking about a different problem area and an unwillingness to discuss referral problem.</td>
<td>After data for presenting problem is collected, BHC might opt to assess other areas identified or suspected as problematic: For example, for a patient presenting with anxiety problems: “You mentioned that gaining weight has also been a problem for you. How much weight do you think you’ve gained in the last year?”</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td></td>
<td>4b. Completes focused assessment in 15 minutes or less for 75% of new appointments.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. <strong>Summary and formulation</strong></td>
<td>5. Provides patient with succinct summary of assessment information and biopsychosocial impressions/formulation of problem. Appropriately integrates key biological, psychological, social, or environmental factors in the formulation. Summary and formulation takes 1–3 minutes and occurs between the assessment and intervention phases of appointment.</td>
<td>Summary includes key symptoms and functional impairment. Formulation includes discussion of biological, behavioral, cognitive, emotional, and/or environmental factors that led to (or maintain) the problem. Example: discusses potential links between medical problems or medications and mood; environmental factors and sleep; cognitions and emotions.</td>
<td>Fail</td>
<td>Pass</td>
</tr>
</tbody>
</table>

| 6. **Population-based care** | 6a. Able to verbally describe principles of population-based care, such as timely access to care, viewing PC panel as his/her own panel, value of prevention and early intervention, and stepped care approach. | 6b. Performs needs analysis (working collaboratively with PCMH staff members or other resources as needed) to identify defined subpopulations of enrolled beneficiaries who might benefit from an BHC appointment. Reviews data from needs analysis with PC leadership and proposes an implementation plan incorporating a clinical pathway for one or more of the conditions identified as a priority by the needs analysis. | Obtains assistance from practice manager, clinic leadership, or population health personnel, to generate list (e.g., through EMR or registry) of top 5 medical conditions seen in clinic (e.g., chronic pain, uncontrolled diabetes, obesity). Prioritizes conditions with behavioral health components. During discussion of needs analysis with leadership, BHC reviews clinical pathways as an option for improving population health. | Fail | Pass |

(Continued)
<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Biopsychosocial approach: Understand relationship of medical and psychological aspects of health (e.g., biopsychosocial model of physiological disorders)</td>
<td>7. Describes to the patient the relevant factors (physical, behaviors, thoughts, environment, interactions with others) impacting symptoms and functional impairments.</td>
<td>Articulates that thoughts can lead to changes in emotions, physiology, behaviors. Articulates that engagement in behaviors can impact thinking, emotions, physiology. Articulates that changes in physiology can lead to changes in behaviors, emotions and thinking. • “When people have thoughts like yours, that can lead to direct physical changes like increased heart rate, blood pressure and muscle tension, which can lead to prolonged feelings of stress/anxiety, decreased social interactions (just don’t feel like it), and work performance (can’t concentrate). This can also interfere with your ability to get restful sleep, which then may further impact mood, energy and concentration.”</td>
<td>Fail Pass</td>
</tr>
</tbody>
</table>

---

Dimension: I. Clinical Practice Knowledge and Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating Comments</th>
</tr>
</thead>
</table>
| 8. Uses evidence-based recommendations and interventions suitable for primary care for patients and PCPs. | 8a. For Phase I training, BHC can demonstrate at least 2 of the following interventions in the categories below: Phase II training must demonstrate at least 3 of the following in the categories below: i) Adapted cognitive interventions, • Adapted cognitive interventions: Questioning unhelpful thinking, developing new ways to think that are consistent with values and goals. • Adapted behavioral interventions: Behavioral activation, behavior change to improve management of sleep, effectively target | Fail Pass | (Continued)
### Dimension: I. Clinical Practice Knowledge and Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ii) Adapted behavioral interventions,</td>
<td>chronic pain, diabetes,</td>
<td>Fail</td>
<td></td>
</tr>
<tr>
<td></td>
<td>iii) Adapted physiological management</td>
<td>obesity, tobacco use, EtOH</td>
<td>Pass</td>
<td></td>
</tr>
<tr>
<td></td>
<td>interventions (e.g., relaxation training),</td>
<td>use, communication skills training.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>iv) Adapted Motivational Interviewing (MI)</td>
<td>• Adapted physiological management interventions:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>interventions,</td>
<td>Relaxed breathing, cue</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>v) Acceptance and Commitment Therapy (ACT) interventions.</td>
<td>controlled relaxation,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>imagery, progressive muscle relaxation, distraction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8b.</td>
<td>Interventions are specifically (operationally) defined and supportable by primary care team members.</td>
<td>• Adapted MI interventions: Using decisional balance, emphasizing personal choice, eliciting change talk, using readiness ruler, developing a change plan.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increase fun activities (dinner with friend 1x/wk).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Increase exercise (Mon–Fri from 1700–1730, 30-minutes on stair-stepper).</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Use relaxation skills (deep diaphragmatic breathing, starting at bedtime).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8c.</td>
<td>Interventions are collaboratively developed with the patient.</td>
<td>Intervention is targeted to improve an objective functional outcome or symptom. Changes are measured through self-report of frequency or duration, report of intensity/quality using 0–10 rating scale, standardized assessment measure, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9a.</td>
<td>Designs interventions to improve functional outcomes and /or reduce symptoms measurably.</td>
<td>• Decrease work absenteeism.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9b.</td>
<td></td>
<td>• Finish specific work tasks by the end of work day.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9c.</td>
<td></td>
<td>• Improve performance on specific responsibilities at home.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
### Dimension: I. Clinical Practice Knowledge and Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
</table>
| **9. Intervention design (continued)** | • Increase frequency of social interactions.  
 • Increase exercise, enjoyable or spiritual activities.  
 • Improve sleep duration, efficiency or decrease time to sleep onset.  
 • Decrease pain exacerbation, improve mood. | | | |

### Dimension: I. Clinical Practice Knowledge and Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>9b.</strong></td>
<td>9b. Uses self-management, home-based practice as the prime method for intervention. The majority of intervention is done by patient outside of the consultation appointment.</td>
<td>Examples: Deep breathing, cue controlled relaxation, cognitive disputation, sleep hygiene, stimulus control, eating behavior changes, increased physical activity, problem-solving, and assertive communication.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| **10. Psychoeducation Class Skills** | 10. Provide classes and/or group medical appointments with a format and content appropriate for primary care. Minimum of 1 class must be scheduled within the first twelve weeks of being in clinic full time. | • Sleep enhancement class  
 • Relaxation class  
 • Drop-in stress management class  
 • Group medical visit for chronic condition  
 • Classes consistent with any of the BHC clinical pathways | | |
11. **Pharmacotherapy**

11a. Can identify common psychotropic medications, the indications for the medication, and common side effects. Can address common myths about psychotropic medication.

11b. Knows indications for medications frequently prescribed for health conditions commonly treated in PC.

11c. Seeks additional information on psychotropic medications from appropriate sources when needed (e.g., PCP, reference material).

11d. Stays within professional license and scope of practice for nonprescribers. BHCs who hold prescription privileges do not prescribe.

---

**Dimension: II. Practice Management Skills**

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Visit efficiency</td>
<td>1. Uses 30-minute appointment efficiently.</td>
<td>Identifies problem, conducts assessment, summarizes understanding of problem at approximately the 15 minute point, and uses the next 10 minutes to develop a behavioral change plan.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Time management</td>
<td>2a. Appointments are kept to 30 minutes or less.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Phase I training standard:</strong> At least 50% of appointments are completed within 30 minutes on the final training day.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Phase II training standard:</strong> At least 85% of appointments are completed within 30 minutes.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2b. Keeps on schedule with consecutive appointments.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If one appointment runs 5 minutes long, the BHC compensates by shortening future appointments (as appropriate) in order to maintain schedule.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Follow-up planning</td>
<td>3a. Appointments are spaced in a manner consistent with a population-health model</td>
<td>• Plans follow-up for two weeks or one month, instead of every week.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Dimension: II. Practice Management Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>as well as individual patient needs. Schedules weekly follow-up intervals only when clinically indicated.</td>
<td>• Schedules less than 2 week follow-up as indicated; e.g., a depressed patient with mild risk of self-harm; patient with daily panic attacks; patient needing bridging care during interval to specialty care appointment, etc.</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3b.</td>
<td>Spaces out BHC appointments by alternating follow-ups with other members of PC team (e.g., PCPs, dietitian).</td>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>4. Intervention efficiency</td>
<td>4a. Sees most patients for 4 or fewer appointments per episode of care. 4b. Structures behavioral change plans consistent with time-limited treatment (i.e., selects interventions that reasonably can be implemented in 4 or fewer appointments).</td>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
</tbody>
</table>

### Dimension: II. Practice Management Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. Triage</td>
<td>6a. Demonstrates knowledge and clinical judgment in determining which patients need referral to specialty mental health. Uses a stepped-care approach.</td>
<td>• Provides BHC intervention over several appointments, assesses progress, recommends referral to specialty mental health if additional care needed. • Encourages referrals for patients with a high severity level of symptoms and low functioning, higher risk category for risk to self or others, alcohol dependence, manic or psychotic symptoms.</td>
<td>Fail</td>
<td>Pass</td>
</tr>
</tbody>
</table>

(Continued)
### Dimension: II. Practice Management Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>6b.</td>
<td>Provides intervention for most common mental health and medical conditions in primary care, recommending specialty mental health care only when indicated.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. PCBH care coordination</td>
<td>7a. BHC and Behavioral Health Care Facilitator (BHCF) staff shared patients (and potentially shared patients) at least weekly (N/A if no BHCF).</td>
<td>Patients with depression or anxiety who are not yet receiving BHCF services.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7b.</td>
<td>Refers patients with depression and/or anxiety to BHCF for care facilitation services, when indicated.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Community resource referrals</td>
<td>8. Has information on military installation resources and community-based resources, and refers patients when indicated.</td>
<td>• Family support services, chaplaincy services, childcare options, legal assistance, etc. • Local support groups, civilian credit or finance counseling, community centers, senior centers, health promotion organizations, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Continuity consultation</td>
<td>9a. Provides continuity consultation to patients with chronic conditions or high medical utilization who might benefit from ongoing, periodic consultation with the BHC.</td>
<td>Patients with obesity, diabetes, chronic pain, etc., seen monthly or quarterly for monitoring progress and reinforcing long-term behavior change plans.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Dimension: II. Practice Management Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>9b. Continuity consultation visits focus primarily on monitoring progress, reinforcing behavior change plans, maintaining gains, and preventing relapse.</td>
<td>Continuity consultation is not provided for the purposes of providing long-term psychotherapy.</td>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>9c. When continuity consultation occurs, the BHC documents the rationale in medical record.</td>
<td>“Patient seen for continuity consultation to promote longer term adherence to recommended lifestyle changes for diabetes management.”</td>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
</tbody>
</table>

---

### Dimension: III. Consultation Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Referral clarity</td>
<td>1. When referral question is unclear, seeks clarification from referring medical note, PCP, and/or patient.</td>
<td></td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>2. Curbside consultations</td>
<td>2a. Available for “on demand” consultation requests from PCPs, BHCFs, and other PCMH staff.</td>
<td>• When BHC is with a patient and PCP knocks on door to speak about another patient, BHC briefly steps out of office to meet with PCP for 1 to 2 minutes to address PCP’s concerns. • BHC does not use a “do not disturb” sign on door.</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td></td>
<td>2b. Clearly and concisely provides assessment results and recommendations in response to PCP “curbside consultation” request. Typically this is no more</td>
<td>Example: PCP approaches about a patient who has been asking for a third renewal of zolpidem. Although you’re running a few minutes behind, you step into the PCP’s office to provide brief consultation based on your recent appointment: “Mrs. S’s sleep problems haven’t improved. This seems related to her difficulty</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### Dimension: III. Consultation Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>than 3 minutes. Often includes specific recommendations for PCP’s course of action with a patient.</td>
<td>giving up her daytime naps. Although she wants more zolpidem, I’d recommend you encourage her to follow through on her plan to eliminate naps and get up at the same time each morning. You could also remind her to stop at the desk to schedule her follow-up with me.”</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>3. Verbal consultative feedback</td>
<td>3a. Provides same-day verbal feedback to PCPs for every appointment, unless PCP requests alternate method and/or frequency. If PCP is not available for same-day verbal feedback, BHC uses alternate means to provide feedback.</td>
<td>Alternate means for feedback could include secure email, secure messaging, copy of EHR note.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3b. PCP feedback is concise (less than 3 minutes), avoids jargon, and includes the BHC’s problem assessment, the behavior change plan, and recommendations for PCP’s care of the patient.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Assertive follow-up with PCPs</td>
<td>4a. Interrupts PCP/team member when needed to address urgent patient needs. Interruptions are kept as brief as possible to minimize impact on PCP workflow. Examples of urgent patient needs: Concerning medication side effects, urgent patient questions regarding medications or medical recommendations, safety concerns.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4b. Avoids interrupting PCPs for nonurgent needs. Uses other PC team members when appropriate for nonurgent issues. Other PC team members include nurses, med techs, admin assistants, etc. Examples of nonurgent needs: Questions regarding appointment scheduling, routine medication refill requests, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Dimension: III. Consultation Skills (continued)

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. <strong>PCP education:</strong> Educates primary care staff on how to optimally use BHC services.</td>
<td>5a. Clearly articulates the following via formal and informal avenues (curbside consultations; i.e., informal or unofficial consultations): i) BHC visit expectations (25 minute consultations, brief episodes of care), ii) referrals (emphasizing what conditions and when to refer), iii) how to best use BHC (alternating appointments to reduce high utilization, teaming up forcontinuity consultation for chronic medical conditions).</td>
<td>Examples of formal avenues: PCMH staff meetings/briefings, EMR notes. Example of informal avenue: Curbside consultations (see Element III, 2b—Consultation Skills)</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>5b. Provides training to PCMH team members on role of BHCs in addressing behavioral health concerns.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. <strong>Value-added recommendations:</strong> Recommendations are tailored to the pace of primary care.</td>
<td>6. Provides PCPs with patient recommendations that the PCP could implement with the patient in 3 minutes or less and are intended to reduce PCP visits and workload. Recommendations are: i) achievable for the patient, ii) evidence-based, iii) brief (explained in less than 3 minutes), iv) concrete, and v) intended to reduce PCP visits and workload.</td>
<td>Example of reducing PCP visits and workload: Follow up with BHC instead of PCP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
## Dimension: IV. Documentation Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Concise, clear charting using appropriate format</strong></td>
<td>1a. BHC-specific EMR template is used.</td>
<td>Example of appropriate EMR template.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1b. Documentation is initiated in the medical record during the clinical encounter.</td>
<td>See examples in service clinic practice manuals.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1c. Clinical notes are written specifically for PCP and team members in a succinct and jargon-free manner.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1d. Documentation includes pertinent history of presenting problem, focused review of systems and MSE, information on functional impairment, and clear clinical impression with specific evidence-based recommendations and follow-up plan for patient and PCMH team.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1e. Ensures AHLTA notes are accessible to the PCMH team and maintained as part of the patient’s medical record.</td>
<td>Example: Avoiding restrictive demarcations, such as checking the box to make the note “sensitive,” which is, in most cases, not appropriate for primary care.</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>2. Written recommendations to PCP:</strong></td>
<td>2a. Completes at least 80% of clinical notes by the end of the same day the patient is seen. Completes 100% of notes within 24 hours of patient’s appointment.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Written feedback provided to PCP and team members.</td>
<td>2b. Content of clinical note is succinct and tailored to the PCP, and avoids psychological jargon.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2c. Written recommendations are actionable by PCP while not adding significantly to PCP’s existing workload.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
### Dimension: V. Administrative Knowledge and Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
</table>
| 1. **BHC policies and procedures:**  
*Uses appropriate scheduling, templates, and coding for PCMH work.* | 1a. Ensures their template is structured to allow for flexibility of scheduling:  
   i) includes time for same-day/walk-in appointments in daily schedule/template,  
   ii) adjusts template based on clinic needs and patient flow,  
   iii) continues to meet the requirement of service-specific available appointments (e.g., minimum of 10 slots per day). | Example of appropriate codes: Health and Behavior procedure (CPT) codes, E/M codes. Examples of inappropriate codes: Any CPT code indicating out-of-scope practices (e.g., formal assessment, review of records, etc.). | Fail | Pass |
| 1b. Uses appropriate MEPRS code, reflecting primary care work. | | | | |
| 1c. Uses correct CPT and/or E/M codes in accordance with Service clinical practice manual. | | | | |
| 1d. Ensures templates reflect appropriate timeframes for appointment type:  
   i) individual and couple appointment slots are no more than 30 minutes,  
   ii) class length appropriately reflects content. | | | | |
| 1e. DoD BHC peer review items are used in regular peer review process. | | | | |

(Continued)
### Dimension: V. Administrative Knowledge and Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Risk management protocols</td>
<td>2a. Accurately describes informed consent procedures in primary care behavioral health consultation to patients.</td>
<td>Examples of specific SOPs addressing risk pathways: • Service practice manuals, relevant service regulations, instructions, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2b. Appropriately uses risk pathways in primary care and any involvement with specialty mental health clinics and/or emergency departments as applicable, as outlined by service-specific SOPs.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2c. Works with PC leadership to formally outline risk assessment, management, and treatment appropriate to PCBH, including writing specific guidelines with detailed pathways/resources.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Dimension: VI. Team Performance Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Fit with primary care culture</td>
<td>1a. Knows the roles of the various primary care team members and articulates their roles/duties in the clinic.</td>
<td>• Group practice manager • PCPs (i.e., physicians, PAs, NPs) • Nurses • Administrative staff • BHCFs</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Dimension: VI. Team Performance Skills

<table>
<thead>
<tr>
<th>Element</th>
<th>Minimal Demonstrated Benchmark Behaviors</th>
<th>Sample Behavioral Anchors</th>
<th>Skill Rating</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1b.</td>
<td>Regularly (90% of clinic meetings)</td>
<td>• Avoids specialty MH jargon (e.g., “session,” “therapy,” “processing,” or “group”). Instead, uses terms such as “appointment,” “visit,” or “classes” to be consistent with PC consultation. • Uses PC space for consultation (e.g., exam room). • Avoids creating unique space akin to specialty MH clinics that are out of the norms of a typical PC clinic, thus promoting a message of “different/ specialty service” (e.g., dim lighting, specialty furniture not found in other PC exam rooms/offices, incense/candles, white noise machine, “do not disturb” sign).</td>
<td>Fail</td>
<td>Pass</td>
</tr>
<tr>
<td>1c.</td>
<td>Uses language and practice habits</td>
<td>• When not seeing patients, actively engages with PCMH team members to encourage referrals. • Door is literally open when not seeing patients. • Does not use a “do not disturb” sign.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 2.     | Maintains flexible attitude and openness | • Avoids specialty MH jargon (e.g., “session,” “therapy,” “processing,” or “group”). Instead, uses terms such as “appointment,” “visit,” or “classes” to be consistent with PC consultation. • Uses PC space for consultation (e.g., exam room). • Avoids creating unique space akin to specialty MH clinics that are out of the norms of a typical PC clinic, thus promoting a message of “different/ specialty service” (e.g., dim lighting, specialty furniture not found in other PC exam rooms/offices, incense/candles, white noise machine, “do not disturb” sign). | Fail | Pass |
| Responsiveness and availability to PC team | in providing consultation: i) Readily provides unscheduled services when needed; ii) Has an “open door” policy encouraging PCMH staff interruptions to promote same-day visits and curbside consultations. | | | |
Phase I Training

Successful completion requires *all unshaded items* rated as “Pass”

**CIRCLE: PASS/FAIL**—(Provide rationale in comments.) Date of completion: ________________

Trainer’s signature: __________________________  BHC’s signature: __________________________

Date: __________________________

Phase II Training

Successful completion requires *all items (shaded and unshaded)* rated as “Pass”

**CIRCLE: PASS/FAIL**—(Provide rationale in comments.) Date of completion: ________________

Trainer’s signature: __________________________  BHC’s signature: __________________________

Date: __________________________

Comments

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
Test publishers and other developers of psychological measures (e.g., researchers, universities, government agencies) regularly release new psychological test instrumentation that can facilitate and evaluate behavioral healthcare treatment in primary care settings. Thus, availability of instrumentation for these purposes is not an issue. However, selection of the appropriate instrument(s) for screening, treatment monitoring, or outcomes assessment is a matter requiring careful consideration. Inattention to an instrument’s intended use, its demonstrated psychometric characteristics, its limitations, and other aspects related to its practical application can result in a misguided diagnosis and/or treatment and potentially harmful consequences for a patient.

Similarly, various modes and methods of automated administration, scoring, and reporting/interpretation of the results of psychological measures that capitalize on the capabilities of today’s technology also are available. These methods require the same level of consideration and scrutiny as the measures they deliver in order to determine which is the most suitable for the settings in which they are implemented.

The purpose of this chapter is to identify and discuss important factors that should be carefully considered when one sets out to determine which among the available psychological test options is optimal for one or more clinical purposes within a primary care setting. The chapter is also intended to similarly identify and discuss factors that should be considered in the selection of psychological testing technologies. Considerations for the selection of psychological measures will be addressed first, followed by a discussion of some of the more common technologies that are available for the administration, scoring, and/or reporting of those measures.

**SELECTION OF PSYCHOLOGICAL TEST INSTRUMENTS**

The selection of a particular psychological test instrument for use in any setting—including primary care settings—is a process that should involve careful thought and deliberation on the part of the individual(s) responsible for making this choice. Perhaps the best place to begin the process is asking whether use of a psychological test for the intended purpose (e.g., screening for depression, monitoring patient recovery over time) is the best solution for the task. While recognizing that psychological tests’ efficiency and objectivity are significant advantages over other methods of clinical information gathering, Urbina recommended doing a cost-benefit analysis similar to one suggested by Goldman (cited in Urbina, 2004), asking questions pertaining to (1) what kind of information will be gained, (2) how that information will be used, (3) how much of that information is available from other
sources, (4) in what other ways that information can be gained, (5) what are the advantages of using psychological tests in addition to or instead of other information sources, and (6) what are the disadvantages of using tests (e.g., cost, time involved) in addition to or instead of other information sources.

Assuming that the benefits and usefulness of using a psychological test outweigh the use of other information gathering methods, it is important to remember that no instrument is best for all settings (Bufka, Crawford, & Levitt, 2002; Ogles, Lambert, & Fields, 2002), regardless of its intended purpose. Consequently, there are many factors to consider as part of the process for evaluating the suitability of one or more tests for a particular purpose. Some of these factors are general in terms of their applicability to various types of tests for various purposes; others are specific to tests that are used for specific purposes and/or with specific patient populations. Both types of factors are discussed in this section.

**General Considerations for Instrument Selection**

Regardless of the type of instrument one might consider using in a primary care setting, psychologists frequently must choose between many product offerings. But what are the general considerations for the selection of any instrument for psychological testing? What should guide the clinician’s selection of an instrument for a specific purpose? As part of their training, psychologists and other appropriately trained professionals from related specialties have been educated about the psychometric properties that are important to consider when determining the appropriateness of an instrument for its intended use. However, psychometric integrity is just one of several considerations that should be taken into account in evaluating a specific instrument for a specific clinical application.

Following are considerations that are recommended for the selection of tests and other assessment instruments for use within primary care settings, regardless of their intended purpose(s). Some of these may seem obvious, but one would be surprised how easily some of these considerations can be overlooked. Following the general considerations are additional considerations for instruments that will be used for specific purposes or under specific circumstances.

Finally, one other point is worth noting. In discussing selection criterion of structured or semi-structured interviews, Summerfeldt and Antony (2002) noted that “no one instrument best fits the requirements of all clinicians and researchers. When selecting an interview, health care workers must consider their specific needs, priorities, and resources” (p. 4). This same caveat holds true not only with regard to interviews but also with regard to the other psychological instruments.

**Relevancy to the Intended Purpose of Assessment**

Evaluating a given psychological test instrument should begin with identifying what is to be measured (Fischer & Corcoran, 1994) and then determining the relevancy of an instrument to the reason for testing the patient. Groth-Marnat (2009) indicated that when selecting a test, the most important consideration is how useful the test is in answering the referral question. This recommendation can be extended to cover other reasons for testing, such as determining personality characteristics important to facilitating a course of psychotherapy or for routine treatment monitoring and outcomes assessment. Administering a PHQ-9 to screen for levels of general distress is not the most appropriate way to accomplish the task—unless, of course, this screener has been empirically shown to be effective when used this way. Regardless of its psychometric integrity, a measure lacking empirical support for the intended use should not warrant any further consideration.

**Source of Information**

With the majority of patients, the clinician will probably want the patient to serve as the source of information (i.e., the respondent to the instrument). There are times, however, when a family
member or other adult must serve as the information source. This is particularly the case with children, but there also are adult cases in which the clinician will have to look to another adult to obtain the desired information (e.g., severely impaired adults, reluctant participants in the assessment process). In those cases, only *instruments specifically developed to obtain information about a patient from another person* such as a family member (e.g., parent, spouse) or other knowledgeable party (e.g., guardian, child’s schoolteacher) should be considered. In these cases, the instrument should be evaluated in the same manner as a patient self-report instrument using the considerations noted in this section. In no instance should the use of a patient self-report instrument be evaluated for use with a collateral as a respondent based on data (e.g., validity, reliability, norms) obtained directly from actual patient samples.

**Instrument Content**

The content of the test—what the test is asking the patient to report on—is also important in achieving acceptability to the patient (e.g., see Bufka et al., 2002). Asking questions that are not considered relevant by the patient (i.e., face invalidity) or that are informative but potentially offensive should be avoided. Patients’ perceptions of what is appropriate or useful to inquire about can have a direct bearing on the probability that they will give a valid response to any or all of the questions being asked.

A related content issue is whether the measure is theory based or is atheoretical. In some cases, the user may want to use a measure that is consistent with a particular theory he or she espouses, or one that is not meant to reflect any particular theory.

**Psychometric Integrity**

Good psychometric properties are a crucial factor for any assessment instrument (Groth-Marnat, 2009). Thus, in selecting a test, one must ask, “Is the instrument a valid and reliable measure of what it purports to measure?” Also, “Should specific aspects of validity and reliability be attended to more than others when evaluating psychological measures?” There is no one clear answer to this latter question. All relevant aspects of validity and reliability are important to consider in selecting a test, but depending on the intended use, particular types of validity and/or reliability might be given more weight in the decision. For example, if the intended use of an instrument is treatment monitoring, one would like to see empirical data demonstrating good test-retest reliability. When used for screening, one should closely consider the instrument’s criterion validity and accompanying diagnostic efficiency statistics (discussed later in this section).

Regardless of the intended use of the measure, there are several aspects of an instrument’s reliability and validity that should be evaluated. The instrument’s internal consistency and test-retest reliabilities are important considerations to ensure consistency of item content and stability of test results over time. Important validity considerations include the instrument’s construct validity, which can be evaluated through means such as examination of the instrument’s content, factor analytic studies, correlations with other measures of the same and dissimilar constructs, ability to discriminate groups of individuals with and without a particular trait or condition, and ability to predict a behavior or outcome in the future.

A detailed discussion of desired psychometric properties of psychological measures is beyond the scope of this chapter. The reader is referred to the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014) for further guidance on these matters. Other resources such as Smith and Archer (2014), Urbina (2004), Cicchetti (1994), and Fischer and Corcoran (1994) provide further discussion and guidance related to the evaluation of the validity and reliability of psychological measures being considered for clinical use.
Availability of Relevant Normative Data
Tied to the relevancy and psychometric soundness of the instrument is the issue of whether normative data are available for the particular population(s) with whom the measure will be used. Using the Geriatric Depression Scale to screen geriatric patients for depression is quite appropriate. However, use of this same instrument with patient populations consisting primarily of young adults and adolescents should be avoided unless its use with other populations has been empirically demonstrated to be valid and relevant norms are available. As Fischer and Corcoran (1994) note, the patient being administered a test should be considered a member of the population on which the test’s norms are based. Moreover, they recommend determining whether the norms are based on a sample that is current or up to date and of a sufficiently large size.

Comprehensibility of Results
Another often overlooked factor in instrument selection is how easily the test results can be understood by the primary care team or other relevant individuals, such as the patient and their family. Aside from the fact that authorization for appropriate treatment may hinge upon the primary care team fully comprehending the implications of the test findings, the psychologist should be able to easily explain the results to the patient and/or the family in order that they more clearly understand the patient’s problems and the implications they have for his or her day-to-day functioning. In some cases, the results may reveal areas of strength, which, if understood by the patient, may facilitate efforts toward positive change. Comprehensibility of results is particularly important when the psychologist utilizes Finn’s (2007) therapeutic assessment technique during the course of treatment.

Actionable Information
Actionable information generally refers to information that one can do something with, that is, information that gives the care provider direction about how to improve the quality of services offered to patients. During the ongoing monitoring of an individual’s progress in treatment, information that is actionable can be used to indicate when therapy isn’t working as well as to suggest the changes that need to be made. However, the patients who provide posttreatment outcomes data that yield actionable information frequently do not benefit from that information because they are no longer in treatment. Instead, the information that they provide can lead to changes that future patients with similar problems will benefit from.

An excellent example of actionable information is the information that is commonly obtained from patient satisfaction surveys. These surveys are almost always administered to patients at the end of an episode of care. Their assessment of specific aspects of the quality and helpfulness of the provider’s services, access to and availability of those services, professionalism of the clinical and support staff, and other aspects pertaining to the quality of the services they received will help the provider and the insurer see where improvement can be made that will result in a more beneficial experience for those patients seen at a later point in time. Similarly, aggregated outcomes data from a multidimensional symptom inventory may show that a clinician does quite well in treating depression but their efforts yield minimal results with anxiety disorders. As a result, the clinician may decide to obtain more supervised training in the treatment of anxiety disorders or, alternatively, limit his practice to patients with depressive disorders.

Usefulness Across the Continuum of Care
The ability to use an instrument across different levels of care (LOCs) may be an important consideration in some cases, depending on the range of services the clinician offers and types of service settings in which the patient is seen (e.g., inpatient, partial hospitalization, intensive outpatient,
outpatient). In particular, outside of the primary care setting, chronically mentally ill patients may cycle through any number of mental health settings due to the waxing and waning of their illness over time. Clinicians wishing to monitor these types of patients will want the instrument they select for this task to be capable of assessing them on the variables of interest (e.g., symptomatology, social functioning), regardless of their fluctuating mental and behavioral status or the setting in which they are being treated. For example, if the psychologist is interested in monitoring symptomatology in bipolar or schizophrenic patients, the instrument chosen to do so should have enough floor and ceiling to account for the variability in symptom severity and frequency (including at the extremes) that is likely to occur over time.

**Ease of Use**

One selection consideration that many professionals may overlook, or at least minimize in their importance, is the ease at which the test is administered and scored by themselves or (in cases where the psychologist is not going to be engaged in these activities) other staff members, such as nurses or administrative support staff. Considerations here can include the psychologist’s training in the use of the instrument (Groth-Marnat, 2009), or how easily the psychologist or other person(s) who will be administering, scoring, and/or interpreting the instrument can become skilled in its use (see Bufka et al., 2002; Fischer & Corcoran, 1994; Newman, Rugh, & Ciarlo, 2004). It also extends to how acceptable it is to the patient’s insurer or other payer, that is, the degree to which they see the value in and are willing to reimburse the clinician or practice for its use.

**Clinical Utility**

There are several ways in which one can define and assess the utility of a psychological measure in clinical settings. For instance, Fischer and Corcoran (1994) indicate that determining the utility of an instrument requires that one consider both the practical advantage of the information the instrument yields as well as whether that same information could be obtained in another way. Smith and Archer (2014) define incremental validity as the primary form of clinical utility, which they define as “the ability of a measure to add a new form of information or improve classification accuracy over and above another established measure of the same construct” (p. 24). In addition to its ease of use and cost-effectiveness (i.e., the information the measure yields compared to other measures), Smith and Archer identify the measure’s diagnostic efficiency as means of evaluating its clinical utility. Diagnostic efficiency has to do with how well an instrument can correctly classify those with the disorder of interest from those without the disorder, typically by comparing the patient’s obtained test score to an established classification cutoff score.

Commonly calculated diagnostic efficiency statistics include sensitivity, specificity, positive predictive power, negative predictive power, overall predictive power, and kappa. Note that the base rate of a particular disorder in a given setting will affect the instrument’s predictive power for that setting. And as Urbina (2004) points out, in situations in which the base rate for a specific disorder is extremely low, the cost of using the test for identifying those with that disorder might not be justified by the relative small gains that can be achieved by its use. Moreover, Urbina cautions that one must recognize that in actual clinical settings, classifications and other clinical decisions are not made solely on the basis of test results. The knowledge and judgment of the clinician also come into play.

**Brevity**

One of the most important characteristics of any instrument being considered for use in primary care settings is that it is brief. Lengthy instruments do not fit well in a primary healthcare delivery system where time-limited, problem-oriented intervention is the primary approach to patient treatment. In
addition, to maximize the patient’s cooperation with the testing process, the length of the instrument should be acceptable to most patients. Keep in mind that what psychologists consider “short” for a test may seem unreasonably long to the patient. The use of lengthy, expensive instruments or batteries of instruments that either provide little useful information or represent overkill with a patient being seen for a limited number of sessions is of little value in primary care settings. An exception may be found in outcomes assessment systems. Generally, however, the shorter the instrument, the better.

In addition to the length of the measure or the time for the patient to complete it, there are two other aspects of brevity: the time to score it and the time to review the information and interpret the results (Bufka et al., 2002). These may become important matters to consider, depending on the test and whether automated (i.e., computer-based) scoring and/or interpretive services are available to the clinician. Ideally, one would want to administer, score, and interpret the results of the test before the patient leaves the office.

It is important to note that brevity is meaningless unless the instrument is valid and reliable. In fact, the briefer the instrument, the more one should be concerned with its psychometric properties given that brief instruments tend to be less reliable and valid than longer instruments. Thus, although the psychometric characteristics of all psychological instruments used in primary care and other settings should be carefully evaluated, particular scrutiny is called for with brief measures of any construct.

**Reading Level**

For instruments that are completed with paper and pencil, online, or any other electronic format in which items are administered via visual presentation, reading level is a major consideration. In general, developers of tests and other assessment instruments have become more sensitive to the issue of reading level vis-à-vis the intended patient population’s expected reading level. Part of this may be due to the fact that software is available to easily determine the reading level of material, thus providing a means of identifying problematic text.

Despite these efforts, one is still amazed at the reading level of some instrumentation that is intended for a general patient audience. Specifically, the reader is referred to McHugh and Behar (2009), who determined the reading grade level for 14 self-report depression measures and 91 self-report anxiety measures using three readability formulas: the Flesch Reading Ease formula (Flesch, 1948, 1949), the SMOG Readability formula (McLaughlin, 1969), and the FORCAST formula (Kern, Sticht, Welty, & Hauke, 1976). For each of the 105 measures, McHugh and Behar present the three estimated reading grade levels for its instructions and the three estimated reading grade levels for its items. The average of each measure’s three reading grade level estimates for instructions and the average of each measure’s three reading grade levels estimates for items for some of the measures that are commonly used in primary care settings include the following: 8.9 and 7.6, respectively, for the Beck Depression Inventory-II (BDI-II); 7.1 and 6.5, respectively, for the Center for Epidemiological Studies Depression Scale (CES-D); 8.6 and 9.1, respectively, for the Patient Health Questionnaire-9 (PHQ-9); 9.6 and 6.2, respectively, for the Zung Self-Rating Depression Scale (SDS); 9.4 and 7.1, respectively, for the Beck Anxiety Inventory (BAI); and 8.5 and 8.7, respectively, for the Generalized Anxiety Disorder-7 (GAD-7). Note that for the 14 depression measures that were investigated, the average reading grade levels for instructions and items were 8.4 and 7.6, respectively; for the 91 anxiety measures investigated, the average reading grade levels for instructions and items were 10.1 and 8.7, respectively.

Primary care settings provide care to patients with various levels of reading ability, so the recommendation is to try to select an instrument with a reading level no higher than eighth grade. Tests
requiring no more than a sixth grade reading level are preferable, but these may be difficult to find. Citing Manly, Jacobs, Touradji, Small and Stern, McHugh and Behar (2009) note that level of educational attainment (i.e., grades completed) cannot serve as a valid estimate of reading grade level. They recommend that an instrument such as the Rapid Estimate of Adult Literacy in Medicine (Davis et al., 1993) be used to make such a determination. For patients reading below the recommended test’s minimal reading level, alternative solutions may include instruments that have available an oral administration mode, such as by audiotape or an interactive voice response (IVR) system (discussed later).

Cost

Cost is always a factor. When funds for psychological testing are limited, one does not want to spend a great deal of that money on expensive testing forms or scoring services. Fortunately, valid, reliable, and useful instruments are becoming more available for little or no cost. Some of these are instruments that have been in the public domain for a number of years; other cost-free instruments have become available only during the past few years. Others may be provided by their developers for a nominal licensing fee. Note that Groth-Marnat (2009) suggests that use of computer-administered measures may help to lower the cost of some instruments.

Overall Feasibility and Practicality

Slade, Thornicroft, and Glover (1999) identified feasibility as a consideration in test selection. In this discussion’s context, they defined feasibility as “the extent to which [the instrument] is suitable for use on a routine, sustainable and meaningful basis in typical clinical settings, when used in a specified manner and for a specified purpose” (p. 245). Feasibility, as conceptualized here, actually is comprised of six characteristics, five of which—brevity, availability, relevance, simplicity, and acceptability—are generally comparable to criteria identified earlier by others. What is unique in this conceptualization is the sixth criterion: value. Value is said to occur when the benefits of using the instrument exceed the costs associated with learning how to use the instrument, implementing its use, and analyzing, presenting, and interpreting the resulting data. Although Slade et al. discussed feasibility as it relates to being a desirable criterion for outcomes instrumentation, the characteristics they identified as being part of this construct are in fact desirable characteristics for psychological assessment instruments regardless of how they are used (e.g., screening, treatment monitoring) in a clinical setting.

Related to an instrument’s feasibility is its practicality. Fischer and Corcoran (1994) identified several aspects to consider when judging an instrument’s practicality. In addition to some of the already discussed considerations (i.e., brevity, content that is acceptable and understandable to the patient, utility, ease of scoring), they point out the importance of the instrument as being sensitive to change, but nonreactive. Reactivity here implies that the act of measurement itself causes changes in the patient. In addition, Fischer and Corcoran recommend using instruments that are relatively direct, that is, those whose scores are signs of problems, not symbols of problems. For example, instruments that ask about the experience, intensity, or frequency of a problem are considered to be direct.

Translated and Adapted Versions of Instruments

Some patients seen in primary care and other healthcare settings present challenges in terms of selecting an appropriate measure for administration. One such group is the US non-English-speaking population and those with English as their second language. Healthcare settings have already begun to establish means of overcoming the language barrier that can prevent the provision of quality medical care. Bilingual staff members, translators at the ready, and materials printed in commonly encountered languages (e.g., Spanish, Chinese, Arabic) provide relatively simple and effective solutions to the problem. However, one area in which finding an easy solution can be problematic is that involving
the use of psychological tests. Simply translating tests that are commonly used in the psychologist’s practice setting, or finding an existing translated version of the test (e.g., by word-of-mouth or on the Internet) without a clear understanding of its psychometric properties or how it was developed is not an acceptable option.

Procedures for and issues related to the development of translated or otherwise adapted versions of existing psychological tests have been presented and discussed by several assessment experts and organizations. In particular, the reader is directed to the Standards for Educational and Psychological Testing (AERA, APA, & NCME, 2014) as well as to the guidelines of the International Test Commission (2001, 2005; also see Acevedo-Polakovich et al., 2007) and the work of Geisinger (1994; Geisinger & McCormick, 2013). Another source of information about translation procedures is Bullinger et al. (1998) who describe the multistage methodology used by the International Quality of Life Assessment (IQOLA) Project. The IQOLA Project sought to develop 14 translations or adaptations of the SF-36 Health Survey that were conceptually equivalent to the US source instrument as well as culturally and linguistically relevant (see Maruish, 2011). Thus, evaluation of any translated or adapted version of a psychological instrument for clinical use should at least include an assessment of how the development of the translation or adaptation compares to the procedures espoused or utilized by one or more of the aforementioned cited sources.

For those requiring help in determining the appropriateness of a particular test translation or adaptation, Fernandez, Boccaccini, and Noland (2007) offer a four-step approach to identifying and selecting translated tests. Briefly, Step 1 is identifying translated tests offered by test publishers. Step 2 involves identifying research that involves the use of translated tests, and then confirming that the research applies to the patient (Step 3). In Step 4, one must determine the level at which the research supports the use of the test with patient(s) in question through available equivalency (construct, metric, and functional) research. Acknowledging the lack of empirical support for many translated tests at the time, Fernandez et al. also provide recommendations for how to deal with such limitations in realistic testing situations. Moreover, the authors also identify cultural issues beyond language that should be addressed during the process of test selection.

General considerations for selecting psychological assessment instruments are summarized in Table 5.1.

**Table 5.1 General Considerations for Selecting Assessment Instruments**

<table>
<thead>
<tr>
<th>Instrument Feature</th>
<th>Important Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevancy to the intended purpose of the assessment</td>
<td>• Is appropriate for measuring the targeted domain(s) in the targeted population(s)</td>
</tr>
<tr>
<td>Source of information</td>
<td>• Was developed to obtain information from the desired source of information (e.g., obtaining patient information from a parent using an instrument developed for administration to parents)</td>
</tr>
<tr>
<td>Instrument content</td>
<td>• Has no irrelevant, inappropriate, or offensive item content</td>
</tr>
<tr>
<td>Psychometric integrity</td>
<td>• Meets generally accepted standards for validity and reliability</td>
</tr>
<tr>
<td></td>
<td>• Has demonstrated responsiveness (for individual data) and/or sensitivity (for group data) to changes in patient status</td>
</tr>
<tr>
<td>Availability of relevant normative data</td>
<td>• Has norms that are appropriate for the targeted population</td>
</tr>
<tr>
<td>Comprehensibility of results</td>
<td>• Results can be easily understood by the provider, patient, family members, and other relevant stakeholders</td>
</tr>
<tr>
<td>Actionable information</td>
<td>• Provides the clinician with information that can be used to improve services to the patient</td>
</tr>
<tr>
<td>Assessment across the continuum of care</td>
<td>• Is appropriate for use with patients receiving care at any level of service (e.g., inpatient, outpatient, partial hospitalization)</td>
</tr>
<tr>
<td>Instrument Feature</td>
<td>Important Considerations</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Ease of use</td>
<td>• Is easy to administer, score, interpret, and provide feedback</td>
</tr>
<tr>
<td>Clinical utility</td>
<td>• Provides information that cannot be obtained any other way</td>
</tr>
<tr>
<td></td>
<td>• Yields classification accuracy that is better than other measures of the same construct</td>
</tr>
<tr>
<td></td>
<td>• Is cost-effective and easy to use</td>
</tr>
<tr>
<td>Brevity</td>
<td>• Is considered short from the patient’s perspective</td>
</tr>
<tr>
<td>Reading level</td>
<td>• Requires no higher than an 8th grade reading level, with 6th grade or lower preferable</td>
</tr>
<tr>
<td></td>
<td>• Or, can be administered via another mode that does not require reading (e.g., audio tape, IVR) and yields comparable results</td>
</tr>
<tr>
<td>Cost</td>
<td>• Inexpensive to use for multiple administrations to a single patient</td>
</tr>
<tr>
<td>Overall practicality and feasibility</td>
<td>• Given all considerations, is practical for use in the intended setting, with the intended population, for the intended purpose(s)</td>
</tr>
<tr>
<td>Translated and adapted versions</td>
<td>• Was developed according to commonly accepted procedures and guidelines</td>
</tr>
<tr>
<td></td>
<td>• Has met psychometric and other standards for equivalency with the original version of the instrument</td>
</tr>
</tbody>
</table>

Adapted from Maruish (2013) with permission.

### Application-Specific Considerations for Instrument Selection

In addition to the general considerations, the selection of test instruments for specific purposes will sometimes require other considerations.

#### Instrumentation for Screening and Diagnosis

Screening for the likelihood of the presence of disorders or (more generally) for the need for additional assessment requires considerations that do not necessarily apply to instruments when they are used for the other assessment purposes addressed in this book. When selecting tests for use in screening for the presence of general psychopathology or specific disorders (as is commonly the case in primary care settings), there are a number of considerations related to sensitivity and specificity that must play into the decision. In general, Ficken’s (1995) requirements of instruments used for screening were noted to include (1) high levels of sensitivity and specificity to DSM or ICD diagnostic criteria; (2) an administration time of no more than 10 minutes; and (3) an administration protocol that easily integrates into the organization’s workflow. Much like Derogatis and Culpepper’s (2004) sequential screening process, cases testing “positive” on the screener then would be administered one or more “second tier” instrument(s) to establish severity and a specific diagnosis.

Narayana and Wong (2015) also discussed the importance of sensitivity and specificity. However, these are only two components of a much more encompassing consideration, that is, criterion validity. Although broadly included in the construct of “validity,” it bears particular attention when evaluating instruments for screening purposes. More specifically, what is being referred to here is the previously mentioned diagnostic or classification accuracy or efficiency, which also includes an instrument’s positive, negative, and overall predictive powers. An instrument’s positive predictive power (PPP) indicates the percentage of patients classified as having the disorder or characteristic assessed by the instrument, who actually do have that disorder or characteristic. An instrument’s negative predictive power (NPP) indicates the percentage of patients classified as not having the disorder or...
characteristic assessed by the instrument, who actually do not have that disorder or characteristic. Overall predictive power represents the total percentage of those administered the instrument who were accurately classified as having or not having the disorder or characteristic. Diagnostic efficiency statistics are discussed in more detail in Chapter 6.

Instrumentation for Treatment Monitoring
Tests that are used for ongoing monitoring of treatment progress over time require considerations that are related to the fact that the instrument will be completed two or more times during the course of an episode of care. Thus, cost becomes a particularly important factor in the selection of a measure to be used for this purpose.

Next, given that the purpose of repeated testing is to detect change in a patient’s status in one or more areas of functioning, it is important that the selected instrument is sensitive to change that has occurred over time. Here, good test-retest reliability is a key indicator of this sensitivity. Lambert and his colleagues (Burlingame, Lambert, Reisinger, Neff, & Mosier, 1995) indicate that the minimum acceptable reliability should be about .70 while Smith and Archer (2014) recommend that the test-retest value should be at least .80. Ideally, one would want to see reported reliability coefficients that are based on a demographically relevant “normal” or community sample. It is difficult to tell what a stated reliability coefficient based on patient sample really means. Is a low to moderate reliability on a given measure (based on a patient sample) due to true change in the patient as a result of intervention, or is it a reflection of error variance that is built into the instrument? And what does a high, patient-based reliability mean? Does it indicate that the instrument is insensitive to change, or was it based on a sample of patients who truly did not change as a result of intervention? From this author’s perspective, one should feel confident with instruments with good community sample-based reliabilities.

In some cases, using a test with alternate forms might be advisable. For instance, if there is a likelihood that the results of subsequent administrations of a given test may reflect a memory or learning effect, it would be helpful to have a different set of questions that have been empirically demonstrated to yield comparable results to those of the original set. Unfortunately, unlike achievement and ability tests, it would be rare to find frequently used treatment monitoring instruments for which there is a validated alternate form.

Outcomes Instruments
Generally, outcomes measurement might be considered an extension or variant of treatment monitoring. It is, in effect, the monitoring of the results of treatment at or beyond the end of treatment. For this reason, those additional test selection considerations for treatment monitoring also apply when tests are being considered for use in outcomes measurement.

Lambert and his colleagues (Lambert & Hawkins, 2004a, 2004b; Ogles et al., 2002) offered a model that conceptualized outcomes measures according to five characteristics: content, social level, source, technology (method), and time orientation. Each of these characteristics should be considered when selecting outcomes measures for use in primary care and other clinical settings. According to Lambert et al., the content of the measure is likely to represent cognition, affect, and behavior related to the measure’s domain(s) of interest (e.g., depression, anxiety). The social level of the measure has to do with whether the measure focuses on intrapersonal, interpersonal, or social role functioning. Source refers to the source of the information that is collected by the measure. This could be the patient, the therapist, a family member, or another. The technology of the change measure relates to the methodology by which change is measured. It involves issues such as measuring change on a global level versus change on a specific symptom or behavior. It may also involve observational information (e.g.,
behavior counts) or status information (e.g., hospital discharge, recidivism, physiological measures). Finally, time orientation has to do with whether the instrument measures the more stable traits or states that are more subject to change. In the first instance (traits), the instrument would be used infrequently but used more frequently in the latter instance (states).

Probably the most thorough and clinically relevant guidelines for the selection of instruments for outcomes measurement purposes comes from the National Institute of Mental Health (NIMH)-supported work of Ciarlo, Brown, Edwards, Kiresuk, and Newman (1986). A synopsis of Newman et al.’s (2004) updated summary of this NIMH-supported work is presented here.

Newman et al. (2004) describe 11 criteria for the selection of outcomes assessment instruments, each of which can be grouped into one of five types of considerations. The first consideration is that of relevance to the target group. The instrument should measure those problems, symptoms, characteristics, and so forth that are common to the population to whom it will be administered. The more heterogeneous the population, the more chance that modifications will be required and that these will alter the standardization and psychometric integrity of the instrument. Related to relevance is the degree to which the results obtained from the instrument’s administration are independent of the type of treatment that is offered to the population being tested.

The second set of general considerations is that of methods and procedures (Newman et al., 2004). Several selection criteria are related to this group. The first is that administration of the instrument is simple and easily taught. Generally, this is more of an issue with clinician rating scales than patient self-report measures. In the case of rating scales, concrete examples or objective referents at each rating level should be provided to the user. This is Newman et al.’s second criterion in this group. The third criterion is that the instrument should allow input not from just the patient but also from other sources (e.g., the clinician, collaterals). The benefits of this include the opportunities to obtain a picture of the patient from many perspectives, to validate reported findings and observations, and to promote honesty in responding from all sources (given that all parties will know that others will be providing input also). The fourth methods and procedures criterion is that the instrument provide information relevant to understanding how the treatment may have affected change in the individual.

The last two methods and procedures criteria, if adhered to, would significantly limit the number of instruments that could be used for outcomes assessment. Neither of the two appears to be critical for use in primary care settings at this point in time. Consequently, one may not wish to place too much weight on their consideration unless there are specific reasons for doing so.

Newman et al.’s (2004) third set of considerations has to do with the psychometric strengths of the instrument. Consistent with what was stated earlier, the NIMH panel of experts indicated that outcomes measures should (1) meet the minimum psychometric standards for reliability (including internal consistency, test-retest reliability, and as appropriate, interrater reliability) and validity (content, construct, and concurrent validity); (2) be difficult to “fake bad” or “fake good;” and (3) be free from response bias and not be reactive or sensitive to factors unrelated to the constructs that are being measured (e.g., physical settings, behavior of the treatment staff). These criteria obviously also apply to other psychological instruments used for purposes other than outcomes assessment.

The fourth group of considerations concerns the cost of the instruments. Newman et al. (2004) point out that how much one should spend on assessment instrumentation and associated costs (e.g., staff time for administering, scoring, processing, and analyzing the data) will depend on how important the data gathered are to assuring a positive return on the functions they support. In the context of the NIMH undertaking, Newman et al. felt that the data obtained through treatment outcomes assessment would support screening/treatment planning, efforts in quality assurance and program
evaluation, cost containment/utilization review activities, and revenue generation efforts. However, this should probably be considered the ideal. At this point, the number and nature of the purposes that would be supported by the obtained data will depend on the individual primary care setting. The more purposes the data can serve, the less costly the instrumentation is likely to be, at least from a value standpoint. In terms of actual costs, Ciarlo et al. (1986) originally estimated that 0.5% of an organization’s total budget would be an affordable amount for materials, staff training, data collection, and processing costs related to outcomes assessment. At the same time, one must be mindful that this recommendation was made over 30 years ago and may not reflect changes in policies, regulatory and accreditation requirements, rate of inflation, and attitudes related to the use of psychological test instruments since that time.

The final set of considerations in instrument selection has to do with the utility of the instrument. Newman et al. (2004) posit four criteria related to utility. First, the scoring procedures and the manner in which the results are presented should be comprehensible to all with a stake in the treatment of the organization’s patients. This would include not only the patient, his or her family, the organization’s administrative staff and other treatment staff, but also third-party payers and legislative and administrative policy makers. Related to this is the criterion that the results of the instrument are easily interpreted by those with a stake in them. Another utility-related criterion is that the instrument be compatible with a number of clinical practices and theories that are employed in the behavioral healthcare arena. This should allow for a greater range of problem applicability and greater acceptance by the various stakeholders in the patient’s treatment.

Also important to consider with regard to utility is that “the assessment instruments should support the clinical processes of a service with minimum interference” (Newman et al., 2004, p. 209). There are two issues here. The first has to do with whether the instrument can support the screening, planning, and/or monitoring activities in addition to the outcomes assessment activities. In other words, are multiple purposes served by the instrument’s results? The second issue has to do with the extent to which the organization’s staff is burdened with the collection and processing of assessment data. How much will the assessment process interfere with the daily work flow of the organization’s staff? Equally important is whether the benefits that accrue from the use of the instrument justify the cost of implementing an outcomes assessment program for whatever purpose.

Although the work of Newman and his colleagues provides more extensive outcomes instrumentation selection guidelines than most, others who have addressed the issue have arrived at recommendations that serve to reinforce and/or complement those found in the NIMH document. For example, Gavin Andrews’ work in Australia led to significant contributions to the body of outcomes assessment knowledge. As part of this, Andrews, Peters, and Teesen (1994) identified six general “qualities of consumer outcome measures” that are generally concordant with those from the NIMH study. First, the measure should meet the criterion of applicability. In other words,

> it should address dimensions which are important to the [patient] (symptoms, disability, and [patient] satisfaction) and useful for the clinician in formulating and conducting treatment, yet the measure should be one which can have its data aggregated in a meaningful way so that the requirements of management can be addressed.

(p. 30)

Multidimensional instruments yielding profiles of scores on all dimensions of interest are viewed as a means of best serving the interests of all concerned.

Acceptability, that is, being both brief and user-friendly, is another desirable quality identified by Andrews et al. (1994). Closely associated with this is the criterion of practicality. It might be viewed as
a composite of those NIMH criteria related to matters of cost, ease of scoring and interpretation, and training required for the use and interpretation of the measure. Again in agreement with the NIMH work, the final three criteria identified by Andrews et al. relate to reliability, validity, and sensitivity to change. With regard to reliability, Andrews et al. specify what they consider to be the minimum levels of acceptable internal consistency reliability (.90 for long tests), interrater reliability (.40), and construct and criterion validity (.50). They also stress the importance of an instrument’s face validity in helping to ensure cooperation from the patient, as well as the importance of self-report instruments having multiple-response options rather than just “yes/no” options. This, of course, would increase the instrument’s sensitivity to small but relevant changes in the patient’s status over time.

Based on the work of Vermillion and Pfeiffer (1993), Burlingame et al. (1995) recommended four criteria for the selection of outcomes measures that also reveal consistency in the thinking of other experts in the field. The first is acceptable “technical features,” that is, validity and reliability. Specifically, these authors recommended that instruments have an internal consistency of at least .80, test-retest reliability of at least .70, and concurrent validity of at least .50. The second criterion is “practicality features.” These include brevity, ease of administration and scoring, and simplicity of the reporting of results. Third, the instrumentation should be “suitable” for the patients that are seen within the setting. Thus, because of the nature of most behavioral health-related problems in primary care settings, it should assess symptomatology and psychosocial functioning. The fourth criterion is sensitivity to “meaningful” change over time, allowing for a differentiation of symptomatic change from interpersonal/social role functional change.

Overall, one would be wise to heed Ogles et al.’s (2002) three overarching recommendations when evaluating psychological measures for outcomes assessment purposes: know the trade-offs involved in the selection of one measure over another, know the audience, and recognize resource limitations.

SELECTION OF TECHNOLOGY FOR PSYCHOLOGICAL ASSESSMENT

Maheu, Pulier, McMenamin, and Posen (2012) noted, “Two potent forces—the technological expansion of health care delivery and heightened consumer expectations—present psychologists with the imperative to efficiently and ethically leverage new tools and systems to enhance practice, research, education, training, and policy” (p. 613). Moreover, being able to conduct psychological assessments from outside of the practice setting via telehealth technologies enables the provision of otherwise unavailable services to certain patients and results in convenience and cost and time savings for some patients (Luxton, Pruitt, & Osenbach, 2014).

Looking back to the mid-1980s and early 1990s, the cutting-edge technology for psychological testing included desktop personal computers linked to optical mark reader (OMR) scanning technologies. There also were those “little black boxes” that facilitated the per-use sale and security of test administration, scoring, and interpretive reports for test publishers, while making in-office, computer-based testing available and convenient to practitioners. As has always been the case, someone has had the foresight to develop applications of current technological advances that we use every day, to the practice of psychological testing. Just as at one time the personal computer held the power of facilitating the assessment process, the Internet, fax, interactive voice response (IVR), and smartphone technologies were later developed to make the assessment process easier, quicker, and more cost-effective. These technologies support the increasing use and popularity of telehealth, “the use of electronic and communications technology to accomplish health care over distance” (Jerome et al., 2000, p. 407). In addition to assessment, telehealth can enable or support
the provision of other clinical services, including crisis response, triage, treatment planning, care management, and even psychotherapy.

Perhaps more than ever, the selection of the type of technology by which psychological tests will be administered, scored, and/or interpreted in a primary care setting requires careful consideration—at least as much as when selecting psychological test measures. Following is a brief overview of technologies that are available to facilitate psychological assessment in primary care and other healthcare settings. This is followed by a discussion of factors that one should consider in the selection of assessment technology for their practice setting.

Common Modalities for Psychological Testing

Following are some of the most current modalities for administering, scoring, and/or interpreting psychological tests.

Paper-and-Pencil Administration Forms

Paper-and-pencil test forms represent the simplest of administration modalities. They have the advantage of not being dependent on the patient having access to a particular device, allowing for the administration of a test just about anywhere and at any time. Commonly, the indicated responses on the answer sheet can be either hand scored, or the patient’s responses can be entered into an automated scoring program via key entry or, using specially developed answer forms, via scanning or faxback technology (see later). Moreover, the paper-and-pencil answer form can be used to record answers to test items presented via a validated audiotape version of the test.

Mail-Out/Mail-Back Administration

Another way of utilizing paper-and-pencil administration is through a mail-out/mail-back process. This administration modality can provide a means of monitoring or assessing the outcomes of treatment for those without access to other available electronic technologies and without needing an in-office visit. It can also serve as a means of screening for a specific psychopathology (e.g., depression) or general psychological distress, or as a means of establishing a baseline prior to an initial office visit.

Internet Technology

An Internet-based assessment process is straightforward. The clinician accesses the website on which the desired instrumentation resides. The desired test is selected for administration, and then the patient completes the test online. There may also be an option of having the patient complete a paper-and-pencil version of the instrument and then have administrative staff key-enter the responses into the program. Alternately, the patient may be provided a link to the test administration website, allowing him or her to complete the test outside of the primary care office. The data is scored and entered into the website’s database, and a report is generated and transmitted back to the clinician through the web. Turnaround time for receiving the report can be only a matter of minutes. The test data can later be archived and subsequently be used for any of a number of purposes, such as regularly scheduled reporting of aggregated data to the primary care practice or the psychologist. Data from repeated testing can be used for treatment monitoring and report card generation, as well as for psychometric test development or other statistical purposes.

The advantages of an Internet-based assessment system are clear-cut. It allows for online administration of tests, which can include branching or other logic for item selection. Instruments available through a website can be easily updated and made immediately available to users. This is in contrast
with disk-distributed software, where updates and fixes are sometimes long in coming. The results of an Internet-based test administration can be made available almost immediately. In addition, data from multiple sites can easily be aggregated and used for the purpose of normative comparisons, test validation, risk adjustment, benchmarking, generation of recovery curves, and any number of other statistically based activities that require large data sets. Moreover, the Internet allows one to access psychological measures that utilize item response theory (IRT) and computerized adaptive testing capabilities (Narayana & Wong, 2015).

There are only a few disadvantages to an Internet-based system. One disadvantage has to do with the general Internet data security issue. With time, this will likely become less of a concern as advances in Internet security software and procedures continue to take place. Moreover, as Tsacoumis (2000) has noted, Internet-based testing can be costly, especially when taking into account fees for Internet access, test administration scoring and reporting services from a reputable publisher or vendor, and equipment maintenance and upgrading. How costly this may become will vary from one provider or organization to another, depending on the particular services and vendors that are used.

Smartphone Technology
The commonly found smartphone provides another means of online administration and response entry. After that, entered data are transmitted to the test vendor for scoring, after which the results can be returned to the smartphone for review by the clinician. The real advantage of the smartphone lies in the ability to conduct an automated administration of a test just about anywhere. Also, it is a device that the majority of the US general population is familiar with and is generally well accepted (Clough & Casey, 2015).

Faxback Technology
The development of facsimile and faxback technology that has taken place during the past few decades has made available an important application for psychological testing. The process for implementing faxback technology is fairly simple. A paper-and-pencil answer sheet developed for a test available through the faxback system is completed by the patient. In those systems in which several tests are available, the answer sheet for a given test contains numbers or other types of code that tell the scoring and reporting software which test is being submitted. When the answer sheet is completed, it is faxed in—usually through a toll-free number that the scoring service has provided—to the central scoring facility where the data are both scored and entered into a data base. The system may include an editing capability that will allow for the correction of problematic respondent entry errors (e.g., stray marks, multiple-responses to single-response items) before final processing of the faxed-in data. A report is generated and faxed or emailed back to the clinician, usually within a few minutes. At the scoring end of the process, the whole system remains paperless. Later, the stored data can be used in the same ways as that gathered by an Internet-based system.

Like Internet-based systems, faxback systems allow for immediate access to software updates and fixes. They also can incorporate the same statistical and benchmarking capabilities. Like the PC-based testing products that are offered through most test publishers, their paper-and-pencil administration format allows for more flexibility as to where and when a patient can be tested. In addition to the types of security issues that come with Internet-based testing, the biggest disadvantage of or problem with faxback testing centers around test identification and linkage of data obtained from an individual patient. Separate answer sheets are required for each instrument that can be scored through the faxback system. The system must also be able to link data from multiple tests or multiple administrations of the same test to a single patient.
An example of how faxback technology can be implemented for screening purposes in primary care settings can be found in a project described by Goldstein, Bershadsky, and Maruish (2000) in the first edition of this book.

**IVR Technology**

Another application of technology to the administration, scoring, and reporting of results of psychological tests can be found in the use of interactive voice response, or IVR, systems. Everyone is familiar with IVR technology. When one calls to order products, address billing problems, find out the balance in his checking account or conduct other phone-enabled activities, one is often asked to provide information to an automated system, either orally or by key entry using the phone’s key pad, in order to facilitate the meeting of the caller’s requests. This is IVR, and its applicability to test administration, data processing, and data storage, is simple.

In essence, IVR test administration is similar to the automated online, desktop administration of psychological measures. However, as Kobak, Greist, Jefferson and Katzelnick (1996) pointed out, the advantages of IVR over PC-based testing include the patients’ comfort with the technology (i.e., touch-tone telephone), easy availability and access to that technology, and the ability of the patient to complete the administration outside of the practitioner’s office at any time of the day or week.

IVR technology’s utility as a test delivery system has been reported in a number of published studies. Kobak and his colleagues conducted several studies that employed the use of IVR technology for administration of various types of psychological instruments. Kobak et al. (1997a, 1997b) found support for the overall validity and use of both IVR and desktop-computer-administered versions of the PRIME-MD as instruments for gathering information from behavioral health patients. Here, the results from the clinician-administered SCID-IV were used as the diagnostic gold standard criteria.

In their review of 10 studies that included the administration of both clinician- and IVR-administered versions of the Hamilton Depression Rating Scale (HDRS; Kobak, Reynolds, Rosenfeld, & Greist, 1990), Kobak, Mundt, Greist, Katzelnick, and Jefferson (2000) found the IVR system to generally yield comparable or better psychometric properties than the clinician version. The combined data across all 10 studies yielded an overall correlation of .81 ($p < .001$) between the scores resulting from the administration of the HRSD in both formats.

Moreover, in their HRDS and other studies, Kobak, Mundt, et al. (2000) reported that up to 90% of patients being assessed by IVR systems reported moderate to very strong acceptance as it related to clarity and ease of use. These findings, in turn, are consistent with Kobak et al.’s (1997) findings in the PRIME-MD study. Findings from these and other studies led Kobak, Mundt et al. (2000) to observe that “several decades of research have shown that people often report more problems of greater severity to computer applications than to clinicians, especially when such problems concern personally sensitive topics” (p. 152). Millard and Carver (1999) arrived at similar findings in their comparison of results obtained from the administration of the SF-12 Health Survey via the IVR and live telephone interview formats.

IVR technology is attractive from many standpoints. It requires no extra equipment beyond a touch-tone telephone or cell phone for administration. It is available for use 24 hours a day, 7 days a week. One does not have to be concerned about the patient’s reading ability, although oral comprehension levels need to be taken into account when determining which instruments are appropriate for administration via the IVR (or other audio) administration format. As with faxback- and Internet-based assessment, the system is such that branching logic or a more sophisticated item selection system (e.g., IRT) can be used in the administration of the instrument. Updates and fixes are easily implemented system-wide. Also, the ability to store data allows for comparison of results from previous testings, aggregation of data for statistical analyses, and all the other data analytic capabilities available through faxback and Internet-based assessment.
In addition, Kobak et al. (2000) identified several methodological benefits of IVR. Among them are the provision of a completely standardized, structured assessment; thorough and complete error checking of data at the time of collection and database entry; immediate real-time access to analyzable data; patient-determined pace of assessment; and increased honesty on sensitive questions.

As for the down side of IVR assessment, probably the biggest issue is that in many instances the patient must be the one to initiate the testing. Thus, control of the testing is turned over to a party that may or may not be interested in or otherwise amenable to psychological testing. With less cooperative patients, this may require costly follow-up efforts to encourage full participation in the process.

Advantages and disadvantages of various assessment technologies and modalities are summarized in Table 5.2.

**Table 5.2 Advantages and Disadvantages of Testing Modalities**

<table>
<thead>
<tr>
<th>Assessment Modalities</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mail-out/ Mail-back</td>
<td>• Does not require any special equipment or software</td>
<td>• Cannot be used with patients with limited or no reading ability</td>
</tr>
<tr>
<td></td>
<td>• Good for research involving large groups, over a large geographic area, and/or repeated administration over time</td>
<td>• Lack of control of testing environment</td>
</tr>
<tr>
<td></td>
<td>• Enables assessment of enduring effects of treatment long after treatment termination</td>
<td>• Costs for postage and follow-up</td>
</tr>
<tr>
<td>Internet</td>
<td>• Immediate access to updated or enhanced versions of software</td>
<td>• May require costly follow-up to obtain data</td>
</tr>
<tr>
<td></td>
<td>• Results immediately available for clinical decision-making</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Enables computer-adaptive test (CAT) administration of measures based on item response theory (IRT)</td>
<td></td>
</tr>
<tr>
<td>Smartphone</td>
<td>• Same as Internet</td>
<td>• Cannot be used with patients with limited or no reading ability</td>
</tr>
<tr>
<td></td>
<td>• Can be used anywhere at any time</td>
<td>• Possible confidentiality or security issues</td>
</tr>
<tr>
<td>Faxback</td>
<td>• Assessment is completed in paper-and-pencil format</td>
<td>• Requires access to the Internet</td>
</tr>
<tr>
<td></td>
<td>• Facilitates data entry for scoring and reporting</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Facilitates database entry for aggregation and analysis of sample or population data</td>
<td></td>
</tr>
<tr>
<td>IVR</td>
<td>• No additional equipment required for patient administration</td>
<td>• Possible security issues</td>
</tr>
<tr>
<td></td>
<td>• Available for patient use 24 hours a day, 7 days a week</td>
<td>• Administration must be initiated by the patient</td>
</tr>
<tr>
<td></td>
<td>• Provides a test administration solution for patients with limited or no reading ability</td>
<td>• May require costly follow-up to obtain data</td>
</tr>
</tbody>
</table>

Adapted from Maruish (2013) with permission.

In addition, Kobak et al. (2000) identified several methodological benefits of IVR. Among them are the provision of a completely standardized, structured assessment; thorough and complete error checking of data at the time of collection and database entry; immediate real-time access to analyzable data; patient-determined pace of assessment; and increased honesty on sensitive questions.

As for the down side of IVR assessment, probably the biggest issue is that in many instances the patient must be the one to initiate the testing. Thus, control of the testing is turned over to a party that may or may not be interested in or otherwise amenable to psychological testing. With less cooperative patients, this may require costly follow-up efforts to encourage full participation in the process.

Advantages and disadvantages of various assessment technologies and modalities are summarized in Table 5.2.

**General Considerations for Technology Selection**

As noted, each of the previously described assessment technologies has its own set of advantages and disadvantages. How does a primary care setting determine which technology is best for its particular needs? For the individual primary care psychologist, there may not be a choice. The primary care setting in which the practitioner provides services may have an assessment system in place—instrumentation,
delivery system, and associated reports—and the practitioner may be required to use it. Alternately, the practitioner may be free to employ any test delivery system that best meets his needs. Regardless of whether the decision-maker is the psychologist or someone else, there are several points that should be considered when selecting from among the available technologies.

Psychometric Integrity of Automated Test Administration
Luxton et al. (2014) warn that those conducting psychological assessments via telehealth technologies must be aware of the psychometric properties of tests that are administered using those technologies, particularly when test administration takes place outside of the clinician’s office. For example, lack of an in-person presence may have an effect on the patient’s clinical presentation and honesty in responding. The technology being employed may be inadequate for the test administration task. Poor network connections, bandwidth limitations, equipment-related distractions, and patient fatigue or discomfort (e.g., eye strain from viewing the monitor) are important considerations when evaluating the test’s psychometric properties. Moreover, patient acceptance of the employed technology as well as cultural and demographic factors must be taken into account when evaluating technologies for psychological test administration.

Noting that the literature reports evidence of both equivalence and differences between paper-and-pencil and Internet administered standardized measures, Luxton et al. (2014) concluded,

there are gaps in the literature that practitioners should consider when selecting particular assessment instruments and [technology] mediums. In particular, the vast majority of available measures and assessment tools are based on norms that were established by employing traditional in-person procedures. The reevaluation of these tools with diverse populations, clinical presentations, and telehealth mediums is necessary to assure the validity of assessments conducted via telehealth technologies. It is critical for practitioners to be cognizant of assessment measure limitations and to appropriately disclose and document them in their practice.

(p. 32)

Cost
Just as it is for test instrumentation, the cost of a system that will administer, score, store, and/or report the results of psychological measures is a major consideration. In addition to any up-front equipment purchase or development costs, one must also consider the associated costs. Depending on the delivery system, additional expenses may result from software licensing fees, ongoing technical support, maintenance contracts, system upgrades, adding new measures to the system, and ongoing phone line and/or Internet provider fees, just to mention a few. These costs will vary from one primary care setting to another, depending on the resources that may already be available to the purchaser.

Availability/Accessibility
The technologies discussed earlier may not available to or accessible by all involved parties. A good example here are smartphones and tablets. Although it seems that everyone has a smartphone, this certainly is not the case. According to survey conducted in 2015, 73% of US adults own a desktop or laptop computer and 68% have a smartphone, but only 45% own a tablet computer (Pew Research Center, 2015). A primary care setting wishing to initiate an offsite smartphone or tablet-based assessment system should therefore determine what percentage of their patient population has access to a smartphone or tablet before completing the decision-making process. Faxback and IVR might be less
problematic for providers and/or patients to access, but this should not be assumed without evidence to this effect.

**Compatibility With Existing EHR Initiatives**

Electronic health record (EHR) systems’ contribution as a tool for healthcare improvement has been recognized and continues to be adopted by healthcare organizations (Kobus, Harman, Do, & Garvin, 2013; Tsai & Bond, 2008). The existence of over 2,400 EHR vendors in 2014 (Mandros, 2015) attests to the acceptance of and demand for these systems. The utility of their use with psychological test data and other behavioral health treatment data has been investigated in several studies (for example, see Gill & Dansky, 2003; Gill, Chen, Grimes, & Klinkman, 2012; Klein, Hunt, & LeBlanc, 2006; Rollman et al., 2002). Such studies show that acceptance of these data into EHRs can facilitate such things as a clinician’s quick access to screening results at the time of the patient visit, tracking patient results over time, and generation of periodic reminders for patient follow-up.

With regard to the present discussion, one needs to evaluate the degree to which the psychological testing technology being considered for the primary care practice is compatible with the current or planned EHR and other information technology (IT) systems that support the practice. It is unlikely that the choice of an EHR or other IT for a primary care practice will be heavily based on the psychologist’s preferred psychological testing system. Thus, the key question is which mode of administration, scoring, and data entry is most compatible with the practice’s EHR or IT system. For example, although an IVR system may be appealing to a primary care setting for many reasons, the practice’s plans to develop a multifunction website would suggest that an Internet-based test delivery system would be more practical and cost effective, all other considerations being equal. For a small primary care group practice, one may find faxback, IVR, or Internet solutions to be equally viable.

**Ease of Implementation**

Regardless of who the decision-maker is, one must consider how easy it will be to implement the assessment delivery and reporting technology in daily clinical practice. How receptive will the clinical and support staff be to the selected technology? How easy will it be to train them in its use? What demands will its use make on them? Where will its use be inserted into the daily workflow? How receptive will patients be to this technology? How difficult will it be for patients to understand how to complete an assessment instrument via this technology? These all are important questions and should have a strong bearing on any decision related to the selection of technology for the purpose of test administration, scoring, and reporting. Especially in the case of outcomes assessment, the success of an outcomes program can hinge on the degree to which a workable plan of implementation—including technology-related aspects—can be developed.

**Patient Affordability**

Situations in which testing is not fully covered by an insurer make the costs incurred by the patient worthy of careful consideration. This is particularly true if there is a per-use fee for the instrument(s) being administered and a separate charge for the use of the automated system. The full or co-pay charge incurred by patients will be made more acceptable if they can see value in both the need for testing and the use of the system. If the benefits of the test delivery are not obvious to the patient, this may impact the relationship with the provider.

**Access to Measures Selected for Use**

No matter how sophisticated the technology is, it will be of limited utility if the instrumentation that the psychologist or primary care practice wishes to use is not available or cannot be developed.
(e.g., due to copyright restrictions) for use with it. Here, one must consider not only what the current needs are but also what they will likely to be in the future. Anticipating future needs for assessment information can be difficult to do. But when it is possible, one should evaluate the potential of having the desired measures available on each type of technology being considered, in light of such factors as test software publishers’ histories of developing new applications for their products, copyright issues, and changes in the technology itself.

**Logistics of Readministration of Measures**

This becomes an issue when psychological testing is part of a system of treatment monitoring or outcomes assessment. Obtaining a baseline measure of a patient’s status is not that difficult when he or she is in-office or (in the case of IVR assessment) on the phone line going through a primary care setting’s intake procedures. At intake, patients generally are more amenable to completing psychological test instruments. Either later during treatment or after the episode of care has been completed, the interest or willingness to complete the same psychological measure one or more times again is usually on the wane.

A provider’s monitoring of patient progress while the treatment is ongoing is facilitated by the fact the patient can be retested while he or she is in the office. After treatment has been completed, the ability to conduct a follow-up remeasurement for outcomes management purposes becomes more difficult because the primary care practice is no longer having ongoing regular or scheduled contact with the patient. Thus, there must be a feasible means of conducting a “long-distance” follow-up assessment of the patient. This would include being able to cost-effectively contact the patient for retesting, provide him or her with a means of easily completing the measure(s) in question, and having the data returned for analysis. For some, this problem may be solved through the use of an IVR system. As was described earlier, the primary care practice would send the patient a letter at a specific time following the termination of treatment, asking him or her to retake the baseline measure. Included in the letter would be a toll-free phone-in number and personal identification number to access the system. One might increase his or her response rate by also enclosing a written copy of the instrument and giving the patient the option of using the IVR system or completing the paper form and mailing it back. A similar method might be used if the primary care setting’s Internet assessment system is accessible to patients.

**Flexibility**

Flexibility refers to the degree to which any number of test administration and reporting options are available through the technology. For example, faxback and IVR technologies only allow for data input via one means, that is, through either completion and faxing of a paper-and-pencil form, or through telephone keypad entry, respectively. Internet-based testing (along with in-office PC-based assessment) allows for online administration and data entry, as well as key entry, or in some cases, optical scanning of data gathered from a scannable pencil-and-paper test form. Flexibility will not be an important consideration in many cases. However, in those cases where maximization of data collection or meeting a minimum quota of tested individuals is critical (e.g., surveying service satisfaction of a specified minimal number of members from a given health plan), being able to gather data by more than one means can help increase the chances of success in the endeavor.

**Clinical Utility**

In Gray’s (1999) discussion of factors to consider in selecting a clinical outcomes system, he addresses clinical utility. In this context, Gray refers to clinical utility as “the ability to provide the clinician with real-time feedback” (p. 9). Thus, one should ascertain how quickly assessment information is
accessible once the patient completes the measure. This can be an extremely important consideration in evaluating systems that will be used to support screening, treatment monitoring, and other clinical activities involving real-time clinical decision-making. Information that is not immediately available to the clinician may prove to be of little or no value.

Although important, Gray’s (1999) view of clinical utility is quite limited. One may extrapolate from earlier discussions in this section and those of Fischer and Corcoran’s (1994) and Newman et al.’s (1994) recommendations pertaining to the evaluation of the clinical utility of psychological instruments in order to consider other important factors when evaluating technologies and modalities to use to conduct psychological assessments. Included here are factors such as patient acceptance, ease of use by all parties involved (e.g., patients, clinicians, support staff), compatibility with the practice’s EHR, and the technology’s ability to support multiple assessment activities (e.g., screening, treatment monitoring, outcomes assessment). Finally, one must also consider the technology/modality’s cost-effectiveness and its overall advantages over other options when judging its clinical utility.

SUMMARY

Choosing the right measures and technology for psychological assessment conducted in a primary care setting can be a difficult task. This author has tried to convey various criteria, features, issues, and concerns that he and other recognized experts feel are important to consider in selecting psychological measures and the technology to support their use in primary care settings. In reality, there is no one set of criteria that either instrumentation or technology must or should meet. Instead, decision-makers must decide what to measure, why they want to measure it, and how they plan to use the resulting information within one’s specific primary care practice. Based on the answers to these questions, determining the importance and necessity (if any) of each of the considerations presented in this chapter as it pertains to the primary care setting’s goals or the psychologist’s needs will be a relatively simple task.

AUTHOR’S NOTE

Portions of this chapter are adapted from M. E. Maruish (2002), Psychological testing in the age of managed behavioral health care. Mahwah, NJ: Lawrence Erlbaum Associates. Adapted with permission.

REFERENCES


As Weissman and her colleagues (2011) have recently pointed out, the objective of reliably determining empirical community-based rates of psychiatric disorder has largely been achieved during the past few decades. The data from several comprehensive epidemiologic studies emphatically underscore the fact that psychiatric disorders are highly prevalent in contemporary society. They are consistent with the benchmark NIMH Epidemiologic Catchment Area (ECA) Study (Meyers et al., 1984; Regier, Robert et al., 1988; Robins et al., 1984), which showed conclusively that psychiatric disorder is pervasive throughout the United States. Similar high rates of disorder have been documented throughout other parts of the world, such as in Kessler’s study (2008) conducted in 27 countries. In the context of primary care, which has come to be treated as “the de facto mental health services system in the United States” (Regier, Goldberg, & Taub, 1978; Burns & Burke, 1985), the prevalence of psychiatric conditions is measurably higher, ranging between 20% and 30% (Barrett, Barrett, Oxman, & Gerber, 1988; Derogatis & Wise, 1989). These rates have been additionally confirmed in other studies (Hansson, Nettlebladt, Borgquist, & Nordstrom, 1994; Olsson et al., 1995; Ustun & Sartorius, 1995). More recently, Sansone and Sansone (2010) in their review identify prevalence rates of psychiatric disorder of 25% to 30% across multiple studies reported in the United States and Europe.

The most prevalent of these psychiatric conditions are represented by anxiety and depressive disorders (Derogatis et al., 1983; Von Korff, Dworkin, Le Resche, & Kruger, 1988), which are often obscured in their presentation and are difficult to detect. They may emerge as independent disorders, but are frequently comorbid. American Medical News (2002) reported that an estimated 19 million Americans will experience an unrecognized depressive disorder this year, in spite of the fact that many of these individuals have regular contact with a physician. It is estimated that by the year 2020 depression will become the second leading cause of disability in the United States after heart disease (Murray & Lopez, 1997), and its economic significance to society is underscored by estimated costs of over $43 billion annually (Penninx et al., 1998).

These facts have led the American Medical Association to strongly endorse National Depression Screening Day and encourage their members to use the event as an impetus to investigate and incorporate screening methodology into their practices. In addition, the US Preventive Services Task Force (USPSTF) has recommended that every primary care provider screen all adult patients for depression (Pignone et al., 2002). The USPSTF has subsequently restated this recommendation (O’Connor, Whitlock, Beil, & Gaynes, 2009) based on a meta-analysis of nine randomized controlled clinical trials (RCTs). However, more recently, Thombs and his colleagues have pointed out (Thombs, Ziegelstein, Roseman, Kloda, & Ianonidis, 2014) that more data may be required to enable a valid assessment of the clinical impact of such screening. This judgment is the result of a number of factors, in part due to the fact that resolution of depressive symptoms takes many trajectory pathways (Chin, Choi, & Wan, 2016) requiring that some patients be followed more closely, but also due to the fact that a significant number of individuals are healthcare avoiders and are reluctant to see healthcare providers at all (Ye, Shim, & Rust, 2012).
Because the majority of screening for psychiatric disorders occurs in primary care, the current review will be presented primarily from a primary care perspective. This strategy is not intended to nullify the importance of community-based screening programs, but rather to facilitate an appreciation of screening methods by presenting them in a familiar and realistic context.

Psychiatric disorders represent a significant public health problem with high associated healthcare liability. This liability is magnified by the presence of comorbid medical disorders in the same individual. Comorbid conditions have a long list of undesirable features associated with them, including high utilization rates, lack of adherence to treatment regimens, atypical treatment responses, high adverse event rates, longer hospital stays, and significantly higher costs (Allison et al., 1995; Katon et al., 1990; Saravay et al., 1996).

This being the case, there are multiple incentives to develop and systematically implement effective psychiatric screening systems in community and primary care settings, not the least of which involves decreasing the morbidity and mortality associated with undetected psychiatric disorders (Hawton, 1981; Kamerow, Pincus, & MacDonald, 1986; Regier et al., 1988). It is well established that the majority of individuals with psychiatric disorders, if not wholly untreated, are treated by healthcare workers without comprehensive mental health training. These individuals are rarely if ever seen by a certified mental health professional (Dohrenwend & Dohrenwend, 1982; Regier, Goldberg, & Taube, 1978; Weissman, Myers, & Thompson, 1981; Yopenic, Clark, & Aneshensel, 1983). In addition, although there is a significant comorbidity between physical illness and psychiatric disorder (Barrett, Barrett, Oxman, & Gerber, 1988; Fulop & Strain, 1991; Rosenthal et al., 1991), rates of detection by primary care clinicians of the most prevalent psychiatric disorders (i.e., anxiety and depressive disorders) has been consistently disappointing (Linn & Yager, 1984; Nielson & Williams, 1980). Historically, it has not been surprising to find recognition rates in medical cohorts falling below 50%. Given the documented increases in rates of healthcare utilization and healthcare costs associated with undetected cases of psychiatric disorder, the efficient implementation of psychiatric screening systems in primary care settings would almost certainly result in significant healthcare savings. Finally, the implementation of effective screening systems could significantly aid those individuals in our communities who are afflicted with psychiatric disorders whose conditions currently go undetected. They could be identified earlier in the course of their illnesses, which hopefully would enable earlier initiation of treatment, and avert the cumulative morbidity associated with their conditions.

PROBLEMS IN THE RECOGNITION OF PSYCHIATRIC DISORDERS

The typical psychiatric disorder presents with few clear-cut clinical or laboratory signs and symptoms, and a pathophysiology and etiology that are obscure. For these reasons, unique problems arise in detecting psychiatric disorders. A prominent source of confusion arises from the fact that the highly prevalent anxiety and depressive disorders often present with multiple associated somatic symptoms, manifestations that are difficult to differentiate from those arising from verifiable medical causes. Shidhaye, Mendenhall, Sumathipala, Sumathipala, and Patel (2013), in their review of 21 studies of women presenting in primary care with anxiety and depression symptoms, indicate a very close relationship in many cases with a variety of somatoform disorders. Schurman, Kramer, and Mitchell (1985) indicate that 72% of visits to primary care doctors resulting in a psychiatric diagnosis presented with somatic symptoms as the primary complaint. Katon et al. (1990) and Bridges and Goldberg (1984) both specify that presentation with somatic symptoms is a key reason for misdiagnosis of psychiatric disorders in primary care. Consistent with this observation, Katon et al. (1990) report that
Table 6.1 Factors That Often Underlie the Failure to Detect Anxiety and Depressive Disorders in Primary Care Patients

- Affective syndrome masked by predominant somatic symptoms.
- Affective syndrome judged to be an associated demoralization reaction.
- Affective syndrome missed because of incomplete diagnostic workup.
- Affective syndrome minimized relative to medical disorder.
- Affective syndrome misdiagnosed as dementia in elderly patients.
- Affective syndrome misunderstood as a negative attitude.

Over a period of several decades, a substantial number of studies have documented both the magnitude and nature of undetected psychiatric disorder in primary care (Davis, Nathan, Crough, & Bairn-tree, 1987; Jones, Badger, Ficken, Leeper, & Anderson, 1987; Kessler, Amick, & Thompson, 1985; Schulberg et al., 1985). These studies reveal a wide range of rates of accurate diagnosis of psychiatric conditions, ranging from a low of 8% (Linn & Yager, 1984) to a high of 53% observed by Shapiro et al. (1987). In additional primary care studies, Gerber et al. (1989) found physicians correctly identified 57% of patients who were independently diagnosed with depression, while Simon and Von Korff (1995) reported a 64% accurate diagnosis of major depression. Yelin et al. (1996) found that 44% of a cohort of 2,000 primary care patients who screened positive for clinical anxiety on the SCL-90-R had been independently assigned a mental health diagnosis. While an improvement over rates
observed in earlier studies, these data also underscore the fact that significant proportions of patients with psychiatric morbidity went undiagnosed. Although the methodology and precision of studies of this phenomenon continue to improve (Anderson & Harthorn, 1989; Rand, Badger, & Coggins, 1988), rates of accurate diagnosis have remained for the most part lower than we would like them to be. A summary of some of these investigations along with their characteristics and accurate detection rates appears in Table 6.2.

### Aided Clinician Recognition

The data from these investigations and others strongly suggest a need to proactively work to facilitate the accurate recognition of psychiatric conditions among healthcare professionals. This is particularly the case in the light of the knowledge that in the future, primary care physicians will probably continue to play a greater role in this regard. If these health professionals cannot correctly identify psychiatric conditions, they can neither adequately treat them personally nor refer them to appropriate mental health professionals. Such a situation will ultimately serve to degrade the quality of our healthcare systems further, confounding effective treatment for medical conditions and preventing effective treatment for the unrecognized psychiatric disorders.

Evidence suggests that many primary care physicians have an accurate appreciation of the nature and the prevalence of psychiatric disorders. They estimate prevalence to be between 20% and 25% in their patient populations, and perceive anxiety and depressive disorders to be the most prevalent conditions they encounter (Fauman, 1983; Orleans, George, Houpt, & Brodie, 1985). In an effort to identify and overcome deficiencies in psychiatric diagnosis in primary care, a number of investigators have studied the effects of introducing diagnostic aids in the primary care setting, in the form of psychological screening tests. Although not unanimous, the studies completed thus far have concluded that applied appropriately, screening tests can improve detection of psychiatric conditions in primary care.

Linn and Yager (1984) used the Zung Self-Rating Depression Scale (SDS) and found an increase of 8% to 25% correct diagnosis in a cohort of 150 general medical patients. Similarly, Zung, Magill, Moore, and George (1983) reported an increase in correct identification rising from 15% to 68% in

### Table 6.2 Research on Rates of Accurate Identification of Psychiatric Morbidity in Primary Care

<table>
<thead>
<tr>
<th>Investigator</th>
<th>Study Sample</th>
<th>Criterion</th>
<th>Correct Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersen &amp; Harthorn, 1989</td>
<td>120 physicians primary care</td>
<td>DKI</td>
<td>33%-Affective disorder 48%-Anxiety disorder</td>
</tr>
<tr>
<td>Davis, et al., 1987</td>
<td>377 family practice patients</td>
<td>Zung</td>
<td>15%-Mild symptoms</td>
</tr>
<tr>
<td>Gerber, et al., 1989</td>
<td>1,068 rural primary care patients</td>
<td>HSCL</td>
<td>57%</td>
</tr>
<tr>
<td>Jones, et al., 1987</td>
<td>20 family physicians/51 patients</td>
<td>DIS</td>
<td>21%</td>
</tr>
<tr>
<td>Rand, et al., 1988</td>
<td>36 family practice residents/520 patients</td>
<td>GHQ</td>
<td>16%</td>
</tr>
<tr>
<td>Kessler, et al., 1985</td>
<td>1,452 primary care patients</td>
<td>GHQ</td>
<td>19.7%</td>
</tr>
<tr>
<td>Linn &amp; Yager 1980</td>
<td>150 patients in a general medicine clinic</td>
<td>Zung</td>
<td>8%</td>
</tr>
<tr>
<td>Simon &amp; Von Korf (1995)</td>
<td>1,952 primary care patients</td>
<td>GHQ</td>
<td>64%</td>
</tr>
<tr>
<td>Schulberg, et al., 1985</td>
<td>294 primary care patients</td>
<td>DIS</td>
<td>44%</td>
</tr>
<tr>
<td>Shapiro, et al., 1987</td>
<td>1,242 patients at a university internal medicine clinic</td>
<td>GHQ</td>
<td>53%</td>
</tr>
<tr>
<td>Zung, et al., 1983</td>
<td>41 family medicine patients</td>
<td>Zung</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SDS</td>
<td></td>
</tr>
</tbody>
</table>
family medicine outpatients with depression using a similar intervention. Also, Moore, Silimperi, and Bobula (1978) observed an increase in correct diagnostic identification from 22% to 56% working with family practice residents, and Magruder-Habib, Zung, and Feussner (1990) reported a threefold increase in accurate identification in a VA sample of 800 primary care patients using feedback from the Zung SDS. Williams, Hitchcock, Cordes, Ramirez, and Pignone (2002) also showed a 10% increase in accurate identification of depression (39% versus 29%) among physicians receiving feedback from the Center for Epidemiologic Studies Depression Scales (CES-Ds) administered to their patients.

Not all studies have shown such dramatic improvements in diagnostic accuracy, however. Hoeper, Nyczi, and Cleary (1979) found essentially no improvement in diagnosis associated with making General Health Questionnaire (GHQ) results available to doctors, and Shapiro et al. (1987) reported only a 7% increase in accuracy when GHQ scores were made accessible. Reifler and his colleagues (1996) evaluated 358 primary care patients using the Symptom Driven Diagnostic System for Primary Care (SDDSP) and showed no difference in outcomes at nine months between patients of physicians who received feedback and those who did not. The issue of aided recognition of psychiatric disorders is a complicated one, with numerous patient and clinician variables affecting results. Nonetheless, as suggested by Anderson and Harthorn (1990), the results of studies on aided recognition do offer some promise.

**OVERVIEW OF SCREENING**

**The Concept of Screening**

Screening has been defined traditionally as “the presumptive identification of unrecognized disease or defect by the application of tests, examinations or other procedures which can be applied rapidly to sort out apparently well persons who probably have a disease from those who probably do not” (Commission on Chronic Illness, 1957). Screening is an operation conducted in an ostensibly well population in order to identify occult instances of the disease or disorder in question. Some authorities make a distinction between screening and case finding, which is specified as the ascertainment of disease in populations comprised of patients with other disorders. Under such a distinction the detection of psychiatric disorders among primary care or other medical patients would more precisely fit the criteria for case finding than screening. In actual implementation there are few operational differences between the two processes, so in the current review I have chosen to use the term screening for both operations.

Regardless of its specific manifestation, the screening process represents a relatively unrefined sieve, designed to segregate the cohort under assessment into “positives” who presumptively have the condition, and “negatives,” who are ostensibly free of the disorder. Screening is not a diagnostic procedure per se. Rather, it represents a preliminary filtering operation that identifies those individuals with the highest probability of having the disorder in question for subsequent specific diagnostic evaluation. Those screened “negative” are usually not subjected to further evaluation.

The conceptual underpinning for screening rests on the premise that the early detection of unrecognized disease in apparently healthy individuals carries with it a discernable advantage in achieving effective treatment and/or cure of the condition. Although logical on its face, this assumption is not always justified. In certain conditions, early detection does not measurably improve our capacity to alter morbidity or mortality, for a variety of reasons, such as unreliable diagnostic procedures or because of ineffective treatments for the condition. Also, sometimes evidence of screening efficacy (or its absence) is flawed because research studies upon which conclusions are based are not sufficiently
rigorous to provide a valid test of the hypothesis (Thombs, Ziegelstein, Roseman, Kloda, & Ioannidis, 2014).

In an attempt to facilitate a better appreciation of the particular health problems that lend themselves to effective screening systems, the World Health Organization (WHO) has published guidelines for effective health screening programs (Wilson & Junger, 1968). A version of these criteria is listed here:

1. The condition should represent an important health problem that carries with it notable morbidity and mortality.
2. Screening programs must be cost-effective, i.e., the incidence/significance of the disorder must be sufficient to justify the costs of screening.
3. Effective methods of treatment must be available for the disorder.
4. The test(s) for the disorder should be reliable and valid, so that detection errors (i.e., false positives or false negatives) are minimized.
5. The test(s) should have high cost-benefit, i.e., the time, effort and personal inconvenience to the patient associated with taking the test should be substantially outweighed by its potential benefits.
6. The condition should be characterized by an asymptomatic or benign period, during which detection will significantly reduce morbidity and/or mortality.
7. Treatment administered during the asymptomatic phase should demonstrate significantly greater efficacy than that dispensed during the symptomatic phase.

Some authorities are not convinced that psychiatric disorders, and the screening systems designed to detect them, conclusively meet all of the above criteria. For example, the efficacy of our treatments for certain psychiatric conditions (e.g., schizophrenia) is arguable, and it has not been definitively demonstrated for some conditions that treatments initiated during asymptomatic phases (e.g., “maintenance” antidepressant treatment) are more efficacious than treatment initiated during acute episodes of manifest symptoms. Nevertheless, it is generally understood that psychiatric conditions and the screening paradigms designed to identify them do meet the WHO criteria in the large majority of instances, and that the consistent implementation of screening systems in primary care populations can substantially improve the quality and cost-efficiency of health care.

**The Epidemiologic Screening Model**

Because many readers do not possess detailed familiarity with screening models, we will briefly review the basic epidemiologic screening model will be briefly reviewed. Essentially, a cohort of individuals who are apparently well, or in the instance of case finding, present with a condition distinct from the index disorder, are evaluated by a “test” to determine if they are at high risk for a particular disorder or disease. As outlined above, the disorder must have sufficient incidence or consequence to be considered a serious public health problem, and be characterized by a distinct early or asymptomatic phase during which it is anticipated that detection will substantially improve the results of treatment.

The screening test itself (e.g., pap smear, Western blot) should be both reliable, that is, consistent in its performance from one administration to the next, and valid, that is, capable of identifying those with the index disorder, and eliminating individuals who do not have the condition. In psychometric terms this form of validity has been traditionally referred to as “predictive” or “discriminative” validity. In epidemiologic models, the predictive validity of the test is apportioned into two distinct partitions: the degree to which the test correctly identifies those individuals who actually have the disorder, termed its sensitivity, and the extent to which those free of the condition are correctly
identified as such, its *specificity*. Correctly identified individuals with the index disorder are referred to as *true positives*, while those accurately identified as being free of the disorder are termed *true negatives*. Misidentifications of healthy individuals as affected are labeled *false positives*, and affected individuals missed by the test are referred to as *false negatives*. It should be noted that each type of prediction error carries with it a socially determined value or significance, termed its *utility*, and that utilities need not be equal. The basic fourfold epidemiologic table, as well as the algebraic definitions of each of these validity indices are given in Table 6.3.

Sensitivity and specificity are a screening test’s most fundamental validity indices; however, other parameters can markedly affect a test’s performance. In particular, the *prevalence* or *base rate* of the disorder in the population under evaluation can have a powerful effect on the results of screening. Two other indicators of test performance, *predictive value of a positive*, and *predictive value of a negative* reflect the interaction between test validity and prevalence. These indices are also defined in Table 6.3, although their detailed discussion is postponed until a later section.

### PSYCHOMETRIC PRINCIPLES IN SCREENING FOR PSYCHIATRIC DISORDER

A realistic appreciation of the psychometric basis for psychiatric screening rests on the realization that they have their basis in *psychological measurement*. Basically, the general principles underlying psychological measurement are no different from those that govern any other form of scientific measurement; however, a major characteristic that distinguishes psychological measurement from other forms resides in the object of measurement: it is usually a *hypothetical construct*. By contrast, measurement in the physical sciences usually involves tangible entities, which are measured via ratio scales with true zeros and equal intervals and ratios throughout the scale continuum (e.g., weight, distance, velocity). In quantifying hypothetical constructs (e.g., anxiety, depression, quality of life), measurement occurs on ordinal-approaching interval scales, which by their nature are less sophisticated, and have substantially larger errors of measurement (Luce & Narens, 1987). Psychological measurement is no less scientific due to this distinction; it is simply less precise than physical measurement.

### Reliability

All scientific measurement is based on consistency or replicability; reliability concerns the degree of replicability inherent in measurement. To what degree would a screening measure provide the same results upon readministration one week later? To what extent do two clinicians making judgments

<table>
<thead>
<tr>
<th>Table 6.3 Epidemiologic Screening Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Actual</strong></td>
</tr>
<tr>
<td><strong>Screening Test</strong></td>
</tr>
<tr>
<td><strong>Cases</strong></td>
</tr>
<tr>
<td><strong>Noncases</strong></td>
</tr>
<tr>
<td>Test Positive</td>
</tr>
<tr>
<td>Test Negative</td>
</tr>
</tbody>
</table>

Sensitivity (Se) = a/(a + c).
False negative rate (1 − Se) = c/(a + c).
Specificity (Sp) = d/(b + d).
False positive rate (1 − Sp) = b/(b + d).
Positive predictive value (PPV) = a/(a + b).
Negative predictive value (NPV) = d/(c + d)
on a psychiatric rating scale agree? Conceived differently, reliability can be thought of as the converse of measurement error. It represents that proportion of variation in measurement that is due to true variation in the attribute under study, as opposed to random or systematic error variance. Reliability can be conceptualized as the ratio of true score variation to the total measurement variance. It specifies the precision of measurement and thereby sets the theoretical limit for measurement validity.

**Validity**

Just as reliability indicates the consistency of measurement, validity reflects the essence of measurement, that is, the degree to which an instrument measures what it is designed to measure. It specifies how well an instrument measures a given attribute or characteristic of interest, and communicates the extent to which a test is capable of discriminating those individuals who possess the characteristic from those in whom it is absent. Establishing the validity of a screening instrument is more complex and programmatic than determining its reliability, and rests upon more elaborate theory. Although the validation process involves many types of validity experiments, the most central to the screening process is predictive validity.

The predictive validity of an assessment device hinges on its degree of correlation or agreement with an external reference criterion—a “gold standard” of some sort. With screening tests, the external criterion usually takes the form of a comprehensive laboratory and/or clinical diagnostic evaluation that definitively establishes the presence or absence of the index condition. Critical to a genuine understanding of predictive validity is the realization that it is highly specific in nature. To say that a particular screening test is “valid” in general has little or no scientific meaning; tests are valid only for specific predictive purposes. Psychological tests employed in screening for psychiatric disorder(s) must be validated specifically in terms of the diagnostic conditions they are designed to predict or identify. As an example, a specific test for depression should be validated in terms of its ability to accurately predict clinical depressions; it should be of little value in screening for other psychiatric disorders except by virtue of the high comorbidity of certain other conditions with depression (Maser & Cloninger, 1990).

**Generalizability**

Like reliability and validity, generalizability is a fundamental psychometric characteristic of test instruments used for psychiatric screening. Many clinical conditions and manifestations are systematically altered as a function of parameters such as age, sex, race, and the presence or absence of comorbid medical illnesses. Once validity coefficients (i.e., sensitivity and specificity) are established for a particular test relative to a specific diagnostic condition, they may vary considerably in a new sample if the demographic and health characteristics of the new cohort are significantly different from the cohort upon which the instrument was originally validated.

To cite examples, it is well established that men are more constrained than women in reporting emotional distress. Well-constructed tests measuring symptomatic distress should provide distinct norms for the two genders (i.e., gender-keyed) to reflect this effect (Nunnally, 1978). Another illustration resides in the change of the phenomenologic characteristics of depression across age: Depression in the very young tends toward less dramatic affective display, and progresses through the classic clinical picture of young and middle adult years, to the geriatric depressions of the elderly, which are more likely to be characterized by cognitive dysfunctions. Any single test is unlikely to perform with the same degree of validity across shifts in relevant parameters; therefore, generalizability must
be established empirically and cannot merely be assumed. Consistent with this fact, an essay by Messick (1995) on modernizing our conceptualization of validity in psychological assessment explicitly identifies generalizability, along with external criterion validity, as one of six discernable aspects of construct validity (Messick, 1995).

**METHODOLOGICAL ISSUES IN SCREENING FOR PSYCHIATRIC DISORDERS**

**Assessment Modality: Self-Report Versus Clinician Rating**

Although advocates and adherents argue the differential merits of self-report versus clinician ratings, a great deal of evidence suggests that the two techniques have strengths and weaknesses of roughly the same general magnitude. Neither approach can be said to function more effectively overall in screening for psychiatric disorder. The circumstances of each screening situation must be independently assessed and weighed, to determine which modality is best suited for a particular screening implementation.

Traditionally, self-report inventories have been more frequently used as screening tests than clinical rating scales. This is probably so because the self-report modality of measurement fits very well with the task of screening. Self-report measures tend to be brief, inexpensive, and well-tolerated by individuals being screened. These features contribute to the cost-efficiency and cost-benefit of self-report. Self-report scales are also transportable; they may be used in a variety of settings, in which they tend to minimize professional time and effort. In addition, their administration, scoring, and evaluation require little or no professional input. The advent of personal computers has also very much facilitated the adoption of self-report approaches in this regard, because computer, administration, and scoring markedly reduces both the professional and technical time involved. Finally, perhaps the greatest advantage of self-report resides in the fact that the test is being completed by the only person experiencing the phenomena—the patient himself or herself. A clinician, no matter how skilled, can never know the actual experience of the respondent, but rather must be satisfied with an approximate, deduced representation of the phenomena.

This last feature of self-report measures also represents one of their greatest potential sources of error, that is, patient bias in reporting. Because the test respondent is providing the data, the opportunity exists to consciously or unconsciously distort the responses given. Although patient bias does represent a potential difficulty for self-report, empirical studies have indicated that such distortions represent a problem only in situations where there is obvious personal gain associated with response distortions. Otherwise, this problem usually does not represent a major source of bias (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974a). There is also the possibility that response sets such as acquiescence or attempts at impression management may result in systematic response distortions, but such effects tend to add little error variance in most realistic clinical screening situations.

Probably the greatest limitation of self-report arises from the inflexibility of the format: a line of questioning cannot be altered or modified depending on how the individual responds to previous questions. In addition, only denotative responses can be appreciated; facial expressions, tones of voice, attitudes and postures, and cognitive/emotional status of the respondent are not integral aspects of the test data. This inflexibility extends to the fact that the respondent must also be literate in order to read and understand the questions.

The brief psychiatric rating scale or clinical interview is a viable alternative to self-report instruments in designing a screening paradigm. The clinical rating scale introduces professional judgment
into the screening process, and is intrinsically more flexible than self-report. The clinician has both the expertise and freedom to delve in more detail into any area of history, thought, or behavior that will deliver relevant information on the respondent’s mental status. The clinician also retains the capacity to clarify ambiguous answers and probe areas of apparent contradiction. In addition, because of his/her sophistication in psychopathology and human behavior, there is the theoretical possibility that more complex and sophisticated instrument design may be utilized in developing psychiatric rating scales.

On the negative side, just as self-report is subject to patient bias, clinical rating scales are subject to equally powerful interviewer biases. Training sessions and videotaped interviews may be utilized in an attempt to reduce systematic errors of this type; however, interviewer bias can never be completely eliminated. Furthermore, the very fact that a professional clinician is required to make the ratings significantly increases the costs of screening. Lay interviewers have been trained to do the evaluations in some cases, but they are rarely as skilled as professionals, and the costs of their training and participation must be weighed into the equation as well. Finally, the more flexibility designed into the interview, the more time it is likely to take. Unless judgment is exercised, the procedure can cease to resemble a screening instrument, and begin to take on the characteristics of a comprehensive diagnostic interview.

Both self-report and brief clinical interview can be designed to define the respondent’s status to facilitate an evaluation of his or her “caseness.” Both approaches lend themselves to quantitative methods, which allow for a normative framework to be established. Most importantly, both approaches “work.” It depends upon the nature of the screening task, the resources at hand, and the experience of the clinicians or investigators involved to determine which method will work best in any particular situation.

The Problem of Low Base Rates

Sixty years ago, Meehl and Rosen (1955) published a report in the psychological literature that sensitized psychologists to the dramatic impact of low base rates on the predictive validity of psychological tests. The authors showed that attempts to predict rare attributes or events, even with highly valid tests, would result in substantially more misclassifications than correct classifications with a sufficiently low prevalence. Knowledge and understanding of this important but little known relationship remained limited to a few specialists at that time. Eleven years later, Vecchio (1966) published a report in the medical literature dealing with essentially the same phenomenon. In Vecchio’s report, because the substantive aspects of the report dealt with screening tests in medicine, the information reached a much wider audience. As a result, knowledge of the special relationship between low base rates and the predictive validity of screening tests has since become well established.

To be precise, low prevalence does not equally affect all aspects of a test’s validity; its impact is felt only in the validity partition that deals with correctly classifying positives or “cases.” Predictive validity concerning negatives, or “noncases,” is minimally affected because with extremely low prevalence, even a test with moderate validity will perform the task of identifying negatives adequately. This relationship is summarized by the data in Table 6.4, which represents a synopsis of data originally provided by Vecchio (1966).

In the example developed by Vecchio (1966), the sensitivity and specificity of the screening test are given as .95, values that are higher than those associated with most psychological screening tests. Table 6.5 provides a more realistic example of the relationship between prevalence and positive predictive value, based on a hypothetical cohort (N = 1,000), with validity coefficients (i.e., sensitivity and specificity) more consistent with those that might be genuinely anticipated for such tests.
The data of Tables 6.4 and 6.5 make it clear that as prevalence drops below 10% the predictive value of a positive experiences a precipitous decline. In the first example, when prevalence reaches 1%, the predictive value of a positive is only 16%, which means in practical terms that in such situations 5 out of 6 positives will be false positives. The predictive value of a negative remains extremely high throughout the range of base rates given, and is essentially unaffected by low prevalence situations. The example from Table 6.5 is more realistic in that the validity coefficients are more analogous to those commonly reported for psychological screening tests. In the screening situation depicted here, the predictive value of a positive drops from 77% when prevalence = 30% (e.g., the approximate rate of psychiatric disorders among specialized medical patients) to 7.5% when prevalence falls to 1%. In the latter instance, 12 out of 13 positives would be false positives.

**Sequential Screening: A Technique for Low Base Rates**

Although screening for psychiatric disorders in general is not usually affected by problems of low base rates, there are specific mental health phenomena (e.g., suicide) and diagnostic categories (e.g., panic disorder) which reveal prevalences that are low. In addition, as Baldessarini, Finklestein, and Arana (1983) have noted, the nature of the population being screened can markedly affect the quality of screening outcomes. A good example of this distinction is provided by the dexamethasone suppression test (DST) when it was used as a screen for major depressive disorder (MDD). The DST functioned relatively effectively as a screen for MDD on inpatient affective disorders units where the prevalence of the disorder is quite high. In general medical practice, however, where the prevalence of MDD is estimated
to be about 5–6%, the DST results in unacceptable rates of misclassification. The validity of the DST is insufficient to support effective screening performance in populations with low base rates of MDD.

A method designed to help overcome low base rate problems is commonly referred to as sequential screening. In a sequential screening paradigm, there are two phases to screening and two screening tests. Phase I involves a less refined screen, whose primary purpose is to correctly identify individuals without the condition and eliminate them from consideration in Phase II evaluation. The initial screening also has the important effect of raising the prevalence of the index condition in the remaining cohort. In Phase II a separate test of ideally superior sensitivity is then utilized. Because the base rate of the index condition has been significantly raised by Phase I screening, the performance of the Phase II screen will involve much lower levels of false positive misclassification. A hypothetical example of sequential screening is given in Table 6.6.

In Phase I of the hypothetical screening, a highly valid instrument with sensitivity and specificity equal to .90 is used in a large population cohort (N = 10,000) with a prevalence of 4% for the index condition. Because of the low base rate, the predictive value of a positive is only 27.2%, meaning essentially that less than 1 out of every 3 positives will be true positives. The 1,320 individuals screened positive from the original cohort of 10,000 subsequently become the cohort for Phase II screening. With an equally valid, independent test (sensitivity and specificity = .90) and a base rate of 27.2%, the predictive value of a positive in Phase II rises to 77%, representing a substantial increase in the level of screening performance.

Sequential screening essentially zeros in on a high-risk subgroup of the population of interest by virtue of a series of consecutive sieves. These have the effect of eliminating from consideration individuals with low likelihood of having the disorder, and simultaneously raising the base rate of the condition in the remaining sample. Sequential screening can become expensive because of the increased number of screening tests that must be administered. However, in certain situations where prevalence is low (e.g., HIV screening in the general population) and the validity of the screening test is already close to maximum, it may be the only method available to minimize errors in classification.

### Receiver Operating Characteristic Analysis

Although some screening tests operate in a qualitative fashion, depending on the presence or absence of a key indicator, psychological screening tests function, as do many others, on a quantitative continuum. The individual being screened must obtain a “score” above some criterion threshold or
“cutoff” to be considered a “positive” or a “case.” The cutoff value is usually determined to be that value that will maximize correct classification and minimize misclassification relative to the index disorder. If the relative consequences of one type of error are considered more costly than the other (i.e., the consequences have dramatically different utilities; e.g., false negative = missed fatal but potentially curable disease), the cutoff value will often be adjusted to take this differential utility into account. Although quantitative methods exist to estimate optimal threshold values (e.g., Weinstein et al., 1980), traditionally, they have been selected by simple inspection of cutoff tables and their associated sensitivities and specificities.

The selection of a cutoff value automatically determines both the sensitivity and specificity of the test because it defines the rates of correct identification and misclassification. Actually, an entire distribution of cutoffs is possible, with corresponding sensitivities and specificities. Further, as touched upon in the previous section, test performance (i.e., the errors associated with a particular cutoff value) is highly affected by the prevalence or base rate of the disorder under study. Viewed from this perspective, a test should not be characterized by a sensitivity and specificity; rather it should be seen as possessing distributions of sensitivities and specificities associated with the distribution of possible threshold values and the distribution of sample prevalences possible.

Receiver operating characteristic (ROC) analysis is a logistic regression method that enables the visualization of the entire distribution of sensitivity/specificity combinations for all possible cutoff values and prevalences. As such, it enables the selection of a criterion threshold based on substantially more information, and represents a much more sophisticated clinical decision process. ROC analysis was first developed by Swets (1964) in the context of signal detection paradigms in psychophysics. Subsequently, applications of the technique were developed in the areas of radiology and medical imaging (Hanley & McNeil, 1982; Metz, 1978; Swets, 1979). Madri and Williams (1986) and Murphy et al. (1987) introduced and applied ROC analysis to the task of screening for psychiatric disorders. Somoza and his colleagues (Somoza, 1996; Somoza & Mossman, 1990, 1991; Somoza, Steer, Beck, & Clark, 1994) have also published an extensive series of in-depth reports integrating ROC analysis with information theory to optimize the performance of diagnostic and screening tests. In their series, these investigators reviewed the topics of construction of tests (Somoza & Mossman, 1990), the effects of prevalence (Mossman & Somoza, 1991), optimizing information yield (Somoza & Mossman, 1992b), and maximizing expected utility (Mossman & Somoza, 1992), among others.

Typically, a ROC curve is developed by plotting corresponding values of a test’s sensitivity (true positive rate) on the vertical axis, against the complement of its specificity (false positive rate) on the horizontal axis, for the entire range of possible cutting scores from lowest to highest. A number of computer programs (e.g., Somoza, 1991; AROC, etc.) are available to generate and plot ROC curves. The ROC curve demonstrates the discriminative capacity of the test at each possible definition of threshold (cutoff score) for psychiatric disorder. If the discriminative capacity of the test is no better than chance, the curve will follow a diagonal straight line from the origin of the graph (lower left) to its uppermost right corner. This line is termed the “line of no information.” The ROC curve rises from the origin (point 0, 0) to its termination point (1, 1) on a plane defined earlier. To the extent that a test has discriminative ability, the curve will bow in a convex manner toward the upper left corner of the graph. The greater the deviation toward the upper left corner, the greater discriminative ability the test has for the particular application at hand. Generally, the cutoff score corresponding to the upper leftmost point on the ROC curve is the empirically determined best cutoff that optimizes the balance between the true positive rate (sensitivity) and the false positive rate (1 minus the specificity). In other words, this is the cutoff score that yields the best discriminative efficiency for the test in the current sample.
An ROC summary statistic describing the overall discriminative capacity of a test is termed the *area under the curve* (AUC). The AUC may be thought of as a probability estimate that at each cutoff score a randomly chosen positive (or “case”) will demonstrate a higher score than a randomly chosen negative. When the ROC curve follows the line of no information, the AUC = .50. In the situation of theoretically optimal discrimination, the ROC curve would follow the outline of the ordinate of the graph from point (0, 0) to point (1, 0), and then move at right angles to point (1, 1). In this situation the AUC would equal 1.0.

ROC analysis was introduced to the area of screening for psychiatric disorders approximately three decades ago; however, investigators have found numerous applications for the technique. In addition to simply describing the distribution of validity coefficients for a single test, ROC analysis has been used to compare various screening tests (Somoza, Steer, Beck, & Clark, 1994; Weinstein, Berwick, Goldman, Murphy, & Barsky, 1989), aid in the validation of new tests, compare different scoring methods for a particular test (Birchnell, Evans, Deahl, & Master, 1989), contrast the screening performance of a test in different populations (Burnam, Wells, Leake, & Lanverk, 1988; Hughson, Cooper, McArdle, & Smith, 1988), and assist in validating a foreign language version of a standard test (Chong & Wilkinson, 1989). ROC analysis has also been effectively integrated with paradigms from information theory to maximize information yield in screening (Somoza & Mossman, 1992b), and with decision-making models to optimize expected utilities of screening outcomes (Mossman & Somoza, 1992). Although ROC analysis does not represent a definitive solution for the complex problems of psychiatric screening, it does significantly increase the information available to the decision-maker and provides a relatively precise and sophisticated methodology for making decisions.

**SCREENING TESTS FOR PSYCHIATRIC DISORDERS**

The predecessors of modern psychological screening instruments date back to the late nineteenth and early twentieth centuries. Sir Francis Galton (1883) created the prototype psychological questionnaire as part of an exposition for the World's Fair. The first self-report symptom inventory, the *Personal Data Sheet*, was developed by Robert Woodworth (1918) as part of the effort to screen American soldiers entering World War I for psychiatric disorders. At approximately the same time, the psychiatrist Adolph Meyer constructed the first psychiatric rating scale, the *Phipps Behavior Chart*, at Johns Hopkins (Kemp, 1914).

Since these pioneering efforts, many hundreds of analogous tests and rating scales have been developed and published. A number of them have become well validated and widely used in primary care and other medical and behavioral health settings. Many of the most commonly used of these instruments are discussed extensively later in other chapters of this book (Chapters 9–23) and thus will not be discussed here. For further reviews of screening concepts and tests for psychological and other conditions, the reader is referred to Maxim, Niebo, and Utell (2014), Rush, First, and Blacker (2008), Sederer and Dickey (1996), Spilker (1996), and Zalaquett and Wood (1997).

**PSYCHIATRIC SCREENING IN MEDICAL SETTINGS**

In medical populations, prevalence estimates of psychiatric disorder are substantially increased over community rates. This is particularly true of anxiety and depressive disorders, which account by far for the majority of psychiatric diagnoses assigned to medical patients (Barrett et al., 1988; Derogatis et al., 1983; Von Korff et al., 1988). In reviews of psychiatric prevalence in medical populations, Barrett et al. (1988) observed prevalence rates of 25% to 30%, while Derogatis and Wise (1989) reported...
prevalence estimates for a broad range of medical cohorts that varied from 22% to 33%. These authors concluded, “In general, it appears that up to one-third of medical inpatients reveal symptoms of depression, while 20 to 25% manifest more substantial depressive symptoms” (p. 101). Concerning anxiety, Kedward and Cooper (1966) observed a prevalence rate of 27% in their study of a London general practice, while Schulberg and his colleagues (1985) observed a combined rate of 8.5% for phobic and panic disorders among American primary care patients. In another review, Wise and Taylor (1990) concluded that 5% to 20% of medical inpatients suffer the symptoms of anxiety, while 6% receive formal anxiety diagnoses. They further determined that depressive phenomena are even more prevalent among medical patients, citing reported rates of depressive syndromes of 11% to 26% for inpatient samples. To further reinforce this notion, Yano and her colleagues in the VA system (2011) recently evaluated practice-wide, survey-based depression screening, and identified twice as many positive screens as were demonstrated through chart-based VA performance measures.

With prevalence rates such as these, and the acknowledged escalations in morbidity and mortality associated with psychiatric disorders, there is little doubt that screening programs for psychiatric disorders in medical populations could achieve impressive utility. Potential therapeutic gains associated with psychiatric screening would be further enhanced and magnified by the fact that attendant related problems such as substance abuse, chronicity (Weissman et al., 2010), and high utilization of healthcare services also would be reduced. Particularly dramatic gains could be realized in specialty areas where estimated prevalence rates are over 50% (e.g., HIV, Lyketsos, Hutton, Fishman, Schwarz, & Trishman [1996], or obesity/weight reduction, Goldstein, Goldsmith, Anger & Leon [1996]). In general, early and accurate identification of occult mental disorders in individuals with primary medical conditions would lead to a significant improvement in their well-being. It would also help to both relieve the fiscal and logistic strain on the healthcare system and reduce the risk of recurrence and chronicity (Rizvi et al., 2014; Unützer & Park, 2012).

Having said this, care must be taken not to substitute one’s desire or expectancy for empirically supported facts. A number of people in the field argue that while anticipating that the introduction of systematic screening programs will bring with it significant improvement in the status of psychiatric disorders in primary care, there is currently not a lot of evidence that this is the case. As an example, in a well-designed study by Romera et al. (2013), a group of 35 primary care practitioners who received depression screening training routinely for 6 months were contrasted with an equivalent-sized group of PCPs who managed depression in the standard manner in their practices. The recognition rates between the two were not significantly different, although there were some potentially mitigating circumstances.

**SCREENING FOR SPECIAL PROBLEMS**

Screening for a specific type of psychiatric/psychological disorder or problem in the primary care setting presents its own set of unique issues and concerns. Indeed, one of the primary purposes of this book is to address the assessment of each of several types of disorders that are frequently seen in patients seeking services from primary care providers. Given this, these matters will not be dealt with here. However, a few examples of the matters that should be considered when screening or assessing particular types of problems and populations are worth mentioning in this chapter.

**Screening for Alcohol Abuse/Alcoholism**

Kessler et al. (1994) found a lifetime prevalence of alcoholism in the general adult population of 14%. In primary care and community hospital settings where less severe forms of alcohol use (i.e.,
hazardous drinking) are included, the prevalence of problem drinking has been reported to be as high as 20% (Bradley, 1992; Muller, 1996; Saunders, Aasland, Amundsen, & Grant, 1993a; Schorling, Klas, Willems, & Everett, 1994). Additionally, alcohol use and abuse is strongly related to violent crime, motor vehicle accidents, and lost productivity (National Institute on Alcohol Abuse and Alcoholism [NIAAA], 2001). Clearly, alcohol-related problems present a major public health concern. Routine screening for alcoholism and hazardous drinking enabling earlier effective interventions could help lessen the burden of alcohol-related consequences. However, alcoholism is a disease that many physicians feel unprepared to diagnose and treat. Israel et al. (1996) noted that 30% to 40% of physicians feel unprepared to diagnose alcoholism, whereas Chang (1997) reported that only 19% of physicians felt that they had adequate preparation for treating alcohol-related disorders. Physicians often lack knowledge about the disorder and its symptoms, which can impede treatment and may contribute to further medical problems. For instance, Hopkins, Zarro, and McCarter (1994) observed that physicians accurately identified only 37% of problem drinkers. Other barriers to the routine screening for problematic alcohol use and dependence include unfamiliarity of physicians with available screening instruments (Bradley, Curry, Koepsell, & Larson, 1995; Wenrich, Paauw, Carline, Curtis, & Ramsey, 1995) and time constraints (Stange, Flocke, & Goodwin, 1998). For example, Ford, Klag, Whelton, Goldsmith, and Levine (1994) noted that although 45% of physicians had heard of the CAGE screen, only 14% could recite all four questions needed. To summarize, practitioners have several valid instruments and effective strategies available to screen for alcohol abuse/dependence. The review by Fiellin, Reid, and O’Connor (2000) provides an excellent summary of the most commonly used screening instruments and the factors known to influence their accuracy and utility. Some of these factors include the target condition the screen is designed to detect, the criterion standard used to compare the screening instrument, and the study sample/population (e.g., gender, ethnicity, and setting—primary care vs. emergency department). All of these factors influence the performance and accuracy of the instruments and thus will determine the utility of a given screen. Also, the recommendations of Fleming (2001) and Gordon et al. (2001) concerning the sequential use of an initial abbreviated screen followed by a more comprehensive screen for initial positives, shows promise as ultimately representing an effective approach. Moreover, a recent study by Geneste et al. (2012) has demonstrated good discriminate capacity for the CAGE, AUDIT, and Rapid Alcohol Problem Screen (RAPS-4 and RAPS-4QF), but emphasized the need to take gender into account when setting normative thresholds for all of the scales. Interested readers are referred to this review as well as to Chapter 15 of this book for more detailed discussion of the assessment of substance abuse in primary care settings.

Screening for Suicidal Behavior

“Suicidal behavior” is a phrase that strongly affects most physicians, psychologists, and other healthcare professionals. Suicidal behavior has always been a perplexing subject for members of the healthcare community because of its apparent unpredictability and life-threatening nature. Chiles and Strosahl (1995) define suicidal behavior as a “broad spectrum of thoughts, communications, and acts . . . ranging from the least common, completed suicide . . . to the more frequent, suicidal communications . . . and the most frequent, suicidal ideation and verbalizations.” At the time, Chiles and Strosahl reported that the rate of suicide has remained stable in the United States at approximately 12.7 deaths per 100,000 over the preceding several decades, and ranked as the eighth leading cause of death in the general population. Suicide ranked as the third leading cause of death for individuals 18–24 years old, while the suicide rate in the elderly (i.e., over 65) is approximately double the rate of suicide in the 18- to 24-year-old population.
Suicidal behavior is generally broken down into three categories. The first type is suicidal ideation, which involves thoughts about suicide. Not a lot is known about the predictive value of suicidal ideation in general; however, Nock and his associates (2009) have recently reported, based on a World Health Organization Mental Health Survey, that primarily disorders characterized by high anxiety and poor impulse control were predictive of which individuals would act on these thoughts. The second type of suicidal behavior concerns suicide attempts, which tend to be more common in females and younger individuals. Chiles and Strosahl (1995) indicate that approximately 50% of those who attempt suicide have no formal mental health diagnosis. The last category of suicidal behavior is completed suicide, which is more common in males and older individuals, many of whom have formal psychiatric diagnoses. Evidence also suggests that whites and divorced or separated persons are at increased risk for suicide, as are individuals with diagnoses of depression, drug abuse, and other psychiatric disorders (Lish et al., 1996). Other risk factors include loss of a spouse (increases risk for up to four years), unemployment, physical illness, bereavement, and physical abuse. Some personality traits associated with suicide include poor problem-solving, dichotomous thinking, and feelings of helplessness (Chiles & Strosahl, 1995).

Lish et al. (1996) note that 82% of the people who commit suicide have visited primary care physicians within the past six months, 53% within one month, and 40% one week prior to the suicide. This situation makes it vital that primary care physicians are able to screen for and recognize the risk factors involved in suicidal behavior. Because medical illness is itself a risk factor for suicide, primary care physicians are more likely to see cases of suicidal behavior in their initial stages compared to mental health professionals such as psychologists or psychiatrists. Because they are usually not as well trained in the identification of mental disorders, these practitioners are more likely to miss the risk factors associated with suicidal behavior. Relevant to this issue, Schain (2007) has provided an excellent review of risk factors, including interviewing techniques that help identify mood disorders, with a focus on adolescents in PC populations.

A major problem that affects screening for suicidal behavior is the phenomena of low base rates, discussed previously. In this instance, the problem of low base rates derives from the low prevalence of suicide in the general population to be screened. Chiles and Strosahl (1995) report a lifetime prevalence of suicide between 1% and 12%, and Lish et al. (1996) report a 7.5% prevalence of suicidal behavior in a VA hospital sample. As pointed out earlier, with base rates this low even the most valid screening tests will produce an unacceptably high number of false positives for every true positive identified.

One of the most common techniques used to estimate or predict suicidal behavior is the profile of risk factors. As has already been mentioned, age (younger and older are at higher risk for suicidal behavior) and race (whites, Hispanics, and Asians are two times more likely to attempt suicide than African Americans [Lish et al., 1996]) have significant predictive value. In addition, those with a mental health diagnosis are 12 times more likely to attempt suicide, those who have had previous mental health treatment are seven times more likely to attempt suicide, and those in poor physical health are four times more likely to attempt suicide. Even with these ratios, Chiles and Strosahl (1995) note that profiling is not sufficiently powerful to accurately predict suicide in individuals. Rather, it is better suited to documenting differential rates of occurrence of suicidal behavior across groups. Some key risk factors addressed by these authors as part of an overall evaluation of suicidal behavior include positive feelings toward suicidal behavior, low ability to tolerate emotional pain, intense feelings of hopelessness, a sense of inescapability, and poor survival and coping beliefs.

Although a clinical interview with a detailed history of treatment and previous suicidal behaviors appears to be the most effective predictor of current suicidal behavior, it can often be very time-consuming and is not cost-effective or practical for screening purposes. Several brief instruments have
been found to be useful in predicting suicidal behavior, including the Beck Hopelessness Scale (BHS; Beck, Kovacs, & Weissman, 1975) and the BDI (Beck & Steer, 1993). Westefeld and Liddel (1994) note that the 21-item BDI may be particularly useful for screening for suicidal behavior in college students. Derogatis and Derogatis (1996) have documented the utility of the SCL-90-R and the BSI in screening for suicidal behavior. A number of investigators have reported the primary symptom dimensions and the global scores of the SCL-90-R/BSI capable of discriminating suicidal behavior in individuals diagnosed with depression and panic disorder (Bulik et al., 1990; Noyes, Chrisiansen, Clancy, Gravey, & Suelzer, 1991). Similarly, Swedo et al. (1991) found that all SCL-90-R subscales successfully distinguished suicide attempters from controls in an adolescent population, and the majority of subscales were effective in discriminating “attempters” from an intermediate “at-risk” group. Adolescents and adults who attempted suicide tended to perceive themselves as more distressed and hopeless on the SCL-90-R than the at-risk group, a finding confirmed by Cohen, Test, and Brown (1990) using the BSI.

Two other screening approaches to suicidal behavior are worth brief mention. Just as young patients and adolescents present unique problems for detection of suicidal ideation, the same is true of older adults who have relatively high rates of suicide. Recently, Heisel and colleagues (2010) pioneered an attempt with the Geriatric Depression Scale (GDS; discussed in detail in Chapter 10 of this book, and its 5-item subscale, the GDS-SI, to screen for suicidal ideation in a primary care population of 626 adults over 65 years of age. Results proved encouraging: cut scores developed through ROC analysis maximized sensitivity at .754 and corresponding specificity at .815. The authors conclude that while both versions of the instrument show promise, the GDS-SI might prove preferable, both because it is shorter and also showed less gender difference.

A somewhat unique approach to screening for suicide in primary care, in contradistinction to the traditional psychometric approaches, is the “asQ’em” methodology reported recently by Horowitz et al. (2013). These authors have piloted a 2-item set of questions focused on current suicidal thought and past behavior in an inpatient hospital (NIH Clinical Center) population, to establish the feasibility and acceptability of the method. The questions were administered by nursing staff and were well received by both patients and nurses. In a sample of 331 patients, 13 (4%) screened positive, but no acute interventions were required. The authors indicate that the feasibility of the technique was clearly established but that more extensive controlled clinical trials are required.

An extensive discussion of assessing for suicidal risk in primary care settings is presented in Chapter 11 of this book.

Screening for Cognitive Impairment

Screening for cognitive impairment, especially when dealing with geriatric populations, is extremely important because it is estimated that up to 70% of patients with an organically based mental disorder go undetected (Strain et al. 1988). Because some of these disorders are reversible if discovered early enough, screening programs in high-risk populations can have a very high utility. Even in conditions found to be irreversible, early detection and diagnosis can help in the development of a treatment plan and the education of family members.

There are several instruments available that provide quick and efficient screening of cognitive functioning. Most of these address the general categories of cognitive functioning covered in the standard mental status examination, including attention, concentration, intelligence, judgment, learning ability, memory, orientation, perception, problem-solving, psychomotor ability, reaction time, and social intactness (McDougall, 1990). However, not all instruments include items from all of these categories.
These general instruments can be contrasted with another class of cognitive screening measures characterized by a more specific focus. Previously, specific types of measures tended to be less common, owing to their limited range of applicability. More recently, specific scales have become more popular as they have been used in conjunction with general measures. One example is the Stroke Unit Mental Status Examination (SUMSE), which was designed specifically to identify cognitive deficits and plan rehabilitation programs for stroke patients (Hajek, Rutman, & Scher, 1989). Another example of a screening instrument with a specific focus is the Dementia of Alzheimer’s Type Inventory (DAT), designed to distinguish Alzheimer’s disease from other dementias (Cummings & Benson, 1986).

Unlike other screening tests, the great majority of cognitive impairment scales are administered by an examiner. There are no pencil-and-paper inventories that can be completed by the respondent alone. Instead, these screening measures are designed to be administered by a professional and require a combination of oral and written responses. Most of the tests are highly transportable, however, and can be administered by a wide variety of healthcare workers.

Cognitive Screening in Primary Care Settings
As mentioned earlier, many cases of cognitive impairment go undetected. This may be due to the fact that the early stages of cognitive dysfunction are often quite subtle, and that many of these cases first present to primary care physicians (Mungas, 1991) who frequently have their principal focus on other systems. Also, many are unfamiliar with the available procedures for detecting cognitive impairment, while others are reluctant to add a formal cognitive screening to their schedule of procedures. Although brief, the 10–30 minutes required of most cognitive screening instruments remains a formidable requirement, considering the fact that on average a family practice physician spends 7–10 minutes with each patient. Because cognitive screening techniques are highly transportable and actuarial in nature, and may be administered by a broad range of healthcare professionals, the solution to introducing such screening in primary care may be to train nurses or physician’s assistants to conduct screening. Such an approach would not add to the burden of physicians, and would at least effect an initiation of such programs so that their utility can be realistically evaluated.

Cognitive Screening in Geriatric Populations
As alluded to earlier, an important consideration in any screening paradigm concerns the prevalence of the index disorder in the population under investigation. The prevalence of cognitive disorders is relatively dramatic in elderly populations. Furher and Ritchie (1993) note a 6% prevalence rate for dementia in the general patient population, which may rise to as high as 14% to 18% in the elderly (Jagger et al., 1992). In studying delirium, Hart et al. (1995) note a prevalence of 10% to 13% in the general patient population, which they estimate might rise to as high as 15% to 30% in elderly patients.

Screening the geriatric patient can often be a challenging enterprise for a number of diverse reasons. First, these patients often present with sensory, perceptual, and motor problems that seriously constrain the use of standardized tests. Poor vision, diminished hearing, and other physical handicaps can undermine the appropriateness of tests that are dependent upon these skills. Similarly, required medications can cause drowsiness or inalertness, or in other ways interfere with optimal cognitive functioning. Illnesses such as heart disease and hypertension, common in the elderly, have also been shown to affect cognitive functioning (Libow, 1977). These limitations require screening instruments that are flexible enough to be adapted to the patient with handicaps or illnesses, and yet be sufficiently standardized to allow normative comparisons.

Another difficulty with this population involves distinguishing cognitive impairment from aging-associated memory loss, and from characteristics of normal aging. This distinction requires a sensitive
screening instrument, as the differences between these conditions are often subtle. Normal aging and dementia can be differentiated through their different effects on such functions as language, memory, perception, attention, information processing speed, and intelligence (Bayles & Kaszniak, 1987). The Global Deterioration Scale is a screening test designed for this specific purpose. It has been shown to describe the magnitude of cognitive decline, and to predict functional ability (Reisberg, Ferris, deLeon, & Crook, 1988).

A final problem encountered when screening in geriatric populations is the comorbidity of depression. Depression is one of several disorders in the elderly that may imitate dementia, resulting in a syndrome known as pseudodementia. These patients have no discernable organic impairment, and the symptoms of dementia will usually remit when the underlying affective disorder is treated. Variability of task performance can distinguish these patients from truly demented patients, who tend to have an overall lowered performance level on all tasks (Wells, 1979). If depression is suspected, it should be the focus of a distinct diagnostic workup.

Further, a number of instruments have been developed that help address these problems. One of these is the Cognitive Test for Delirium (CTD; Hart et al., 1995). The CTD is a 9-item, examiner-administered assessment that evaluates orientation, attention span, memory, comprehension, and vigilance. The CTD is completely nonverbal and requires only 10–15 minutes administration time. Through the application of ROC analysis the authors were able to establish an optimal cutoff score of less than 19 to discriminate delirium from other disorders (Hart et al., 1995). The authors also report that the CTD correlates highly with the Mini-Mental Status Examination (MMSE) in delirium and dementia patients, and that it achieved a sensitivity and specificity of 100% and 95%, respectively, in an implementation with dementia in ICU patients.

Assessment of cognitive impairment in primary care settings is the focus of Chapter 14 of this book.

CONCLUSION

Currently, there is little doubt that if not all, many psychiatric disorders meet the WHO criteria for conditions appropriate for the development of effective health screening programs (Wilson & Junger, 1968). These conditions represent a health problem of substantial magnitude, and the morbidity, mortality, and costs associated with these conditions is imposing. There currently exist valid, cost-efficient, psychological screening tests to effectively identify these conditions in medical and community settings, and the efficacy of our treatment regimens is consistently improving. Although confirmation of the incremental treatment advantage of early detection remains somewhat equivocal, evidence is compelling that left to their natural courses, such conditions will result in chronic, compound morbidities of both a physical and psychological nature (Derogatis & Wise, 1989; Katon et al., 1990; Regier, Robert et al., 1988; Weissman et al., 2010). Paradoxically, however, it will be of little ultimate consequence to develop effective systems of treatment planning and outcomes assessment for these conditions if the majority of individuals who would benefit from their utilization are lost to the system.

In large measure this undesirable reality has to do with the fact that a substantial majority of patients with psychiatric conditions are never seen by mental health professionals, and up to 20% are never seen by any healthcare professional. Even when engaged by the healthcare system, the majority of individuals with psychiatric morbidity are seen by professionals who have been insufficiently trained to recognize or effectively treat these conditions. A substantial plurality of these cases go unrecognized, and of those in whom a correct diagnosis is made, only a minority are referred to
mental health professionals for treatment. Often, primary care physicians prefer to treat these cases personally, even though many indicate that they do not feel confident in doing so.

The major implication of this realization for treatment planning and outcomes assessment is that although there are now efficient methods to identify individuals suffering from psychological problems, these techniques are essentially going unused, or addressed to the “tip of the iceberg.”

Unützer and Park (2012) conclude specifically regarding depression,

> Effective management of depression in the primary care setting requires a systematic population-based approach which entails systematic case-finding and diagnosis, patient engagement and education, use of evidence-based treatments including medications and/or psychotherapy, close follow-up . . . and a commitment to keep adjusting treatments or consult with mental health specialists until depression is significantly improved.

(p. 415)

This prescription is equally appropriate for all psychiatric disorders seen in primary care.

In this era of integrated primary care, it seems imperative that psychologists introduce and adopt effective methods to facilitate primary care physicians’ diagnostic and treatment decisions concerning psychiatric disorders. Available psychological screening techniques can substantially contribute to delivering valid, cost-effective identification of these conditions currently. Given their numerous benefits and demonstrated cost savings, it seems obvious that such systems should continue to be widely implemented as the anticipated growth and desirable advances in this field are made.

REFERENCES


As mental health care continues to benefit from parity and the American healthcare system undergoes the most significant reforms of the last half century, mental healthcare delivery has been moving to primary care and continues to be redefined as a component of general health care. This marks positive progress for the mental health field, as the paradigm shift may help reduce stigma and increase acceptance of mental health care as a fundamental component of high-quality care—not an ancillary benefit that payors can opt in or out of funding. Integrated primary care is central to this new paradigm as it is the venue in which the entire population can be reached for mental health assessment, treatment, prevention, and outreach. Of all the services that can be provided in integrated primary care, this chapter will focus on psychological assessment. While other chapters in this edited book address the assessment of specific conditions and the selection of various instruments for each, the following chapter addresses issues and matters related to monitoring patient progress during a course of treatment in a primary care setting and identifying the outcomes at the conclusion of that treatment using psychological assessment instruments. Also discussed is the importance of the resulting information and information from other sources in the evaluation of behavioral health services within primary care settings.

GENERAL CONSIDERATIONS

One important strategic goal for improving the entire US healthcare system is the Quadruple Aim, which reflects four goals of our current healthcare revolution: lower costs, better patient experience, better health, and wellness (low burnout) among medical providers (Bodenheimer & Sinsky, 2014). The Quadruple Aim builds on the Triple Aim (Berwick, Nolan, & Wittington, 2008) by adding a fourth goal: reducing staff burnout. Yet, as some have pointed out, formal cost and treatment analyses cannot be done without having specific measures for specific health conditions (Zivin et al., 2008), and outcome measures should involve data that can be generated as a normal consequence of clinical documentation (Ader et al., 2015).

Aside from measuring patient outcomes during clinical documentation, measurement of population health at the program and system levels serves the purpose of helping healthcare systems pursue the Quadruple Aim. Yet, there are other reasons to measure at the program and system levels (Ader et al., 2015; Zivin et al., 2008). For example, without adequate measurement efforts, one cannot compare the effectiveness between programs. This means patients could receive an inferior treatment regimen across two practices within the same organization (e.g., in Accountable Care Organizations [ACOs]). On a practical level, without measuring a program’s results, few
administrators and healthcare leaders will continuously support a particular program. In today’s rapidly evolving healthcare landscape cost and quality of care are central topics. These terms might be combined and referred to as the value of the care delivered. If all assessment in primary care—at the patient, program, and system levels—is not considered in terms of the Quadruple Aim, those conducting the assessments render themselves irrelevant to the priorities of today’s primary care environment.

Focus on the Quadruple Aim and its relevance to behavioral health personnel (BHPs) working in integrated care is unavoidable. This means one must think and operate with population health paradigms in mind. Likewise, one may be required to provide or simply be exposed to business models and plans that focus on provision of services to minimize cost and maximize clinical outcomes. The more clearly BHPs document, the easier it is to measure their contribution to the Quadruple Aim.

A final remark about this is that value-based models of care are developing across the United States, with initiatives like Colorado Access (Colorado Innovation Model) and Oregon Primary Care Transformation Initiative leading the way. This means that in the future, some BHPs may not rely on billing codes for reimbursement. The value of their contribution will be judged exclusively on the clinical outcomes and the clinical services they perform. Thus, assessing outcomes, documenting the model of care delivered, and formatting this information in the most succinct and mineable form is a preeminent goal in integrated behavioral health care. Collecting these data cannot be ancillary procedures or solely for research purposes, which are often connected to particular interests of payors or research groups (Nelson et al., 2015).

Broadly speaking, psychological assessment in integrated primary care may occur at the individual level (i.e., patient treatment outcomes) and at the group or population level (i.e., program, practice, or medical system). This chapter will discuss these with particular attention to how BHPs can apply this knowledge to their practice in integrated primary care. If BHPs understand how to employ assessment strategies to measure clinical progress, that’s an important start. Demonstrating the quality, cost, and sustainability of one’s IBH program is even more vital since IBH is increasingly being required by the business and healthcare operations leadership across the healthcare field.

**Nomenclature**

Working in primary care settings requires the BHP to adopt cultural terms, work habits, and language that improves the quality of teamwork when working together with non-mental health providers. The field has been clear that shared language and shared training are integral to ensuring high-quality team-based care (Corso & Gage, 2016; Peek & the NIAC, 2013). BHPs practicing in primary care are strongly encouraged to adopt the *Lexicon for Mental Health and Primary Care Integration* (Peek & the NIAC, 2013). This resource helps all members of the primary care team understand the different implications of using terms like “mental health,” “behavioral health,” and “substance misuse treatment.” Furthermore, when all members of the primary care team can differentiate all members of their team (registered dietician, social worker, behavioral health consultant, psychiatry liaison, physical therapist, nurse educator, case manager, care coordinator, health coach, etc.), they can effectively use all these resources to the patients’ benefit. For example, in this chapter the term “behavioral health” will be used to connote a number of services including mental health and behavioral medicine (Peek & the NIAC, 2013).

In the age of patient-centered and team-based primary care, one must also consider how care providers impact patients. Without clear language, patients will be confused. If one intends to operate in a patient-centered manner and deliver high-quality outcomes (per the Quadruple Aim), the language
must be standard and clear so that one knows exactly what to expect from each team member (Corso, Hunter, Dahl, Kallenberg, & Manson, 2016). Without shared interprofessional language, effective teamwork is unlikely to transpire (Corso & Gage, 2016).

Because it is preferable in integrated care settings that BHPs also share the same systems the primary care team uses, it is worth explaining a few terms. For example, *electronic medical records* (EMRs) equate to digital versions of paper charts and enable clinicians to document medical and treatment information, track data easily, and monitor baseline information (e.g., quality metrics). A term that is sometimes used erroneously and interchangeably with EMR is *electronic health record* (EHR). These systems transcend simple charting and advance medical practices, improving communication and quality metrics while increasing the patient’s role in the delivery of their health care.

Other terms to be aware of are listed in Table 7.1. These are important because practice managers, administration, nursing and medical leadership, or business healthcare professionals may ask you to collect these data or use these data to demonstrate quality (i.e., clinical outcomes), profitability (i.e., how much direct revenue it can generate), scalability (i.e., how can you replicate it to the larger system or expand it within the same clinic), or sustainability (i.e., the extent to which your program can be effectively maintained and managed with the same or fewer resources longitudinally).

**Table 7.1 Important Terms for Those Conducting Psychological Assessment in Primary Care**

<table>
<thead>
<tr>
<th>Term</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral Health Provider (BHP)</td>
<td>These are licensed independent practitioners who are hired and integrated into the medical home in an integrated behavioral healthcare (IBH) role. Typically excludes nonlicensed independent practitioners such as nurses, peer or health coaches.</td>
</tr>
<tr>
<td>Healthcare Effectiveness Data and Information Set (HEDIS)</td>
<td>A “set of process metrics used by more than 90% of America’s health plans to measure performance on important dimensions of care and service. Altogether, HEDIS consists of 81 measures across 5 domains of care. Because so many plans collect HEDIS data, and because the measures are so specifically defined, HEDIS makes it possible to compare the performance of health plans on an ‘apples-to-apples’ basis.” (<a href="http://www.ncqa.org/HEDISQualityMeasurement.aspx#sthash.Sk98CsqA.dpuf">http://www.ncqa.org/HEDISQualityMeasurement.aspx#sthash.Sk98CsqA.dpuf</a>)</td>
</tr>
<tr>
<td>Health Information Technology for Economic and Clinical Health (HITECH) Act</td>
<td>This was enacted as part of the American Recovery and Reinvestment Act of 2009 to promote the adoption and meaningful use of health information technology (including financial incentives for use). Subtitle D of the HITECH Act addresses the privacy and security concerns associated with the electronic transmission of health information, in part, through several provisions that strengthen the civil and criminal enforcement of the HIPAA rules.</td>
</tr>
<tr>
<td>Electronic Health Record (EHR)</td>
<td>This goes beyond basic charting and moves practices into improved communication and quality care while improving patients’ abilities to self-manage and be active members of their health care. EHRs expand practices to reach beyond a provider’s medical care to encompass other health professionals, interprofessional healthcare communication, and shared information systems with hospitals, specialists, and laboratories. They also are designed to be accessed by patients to improve their education and interaction in their own health care, meeting stages of meaningful use.</td>
</tr>
<tr>
<td>Electronic Medical Record (EMR)</td>
<td>Digital versions of paper charts that allow providers to document medical and treatment history; track data easily, and monitor baseline quality of care.</td>
</tr>
<tr>
<td>Implementation Metrics</td>
<td>Implementation metrics help gauge readiness to collect all other metrics. These are often a measure of how established or functional an IBH program or service is.</td>
</tr>
</tbody>
</table>

*(Continued)*
Term | Meaning
--- | ---
Operational Metrics | Operational metrics are the practical factors and metrics that may affect outcome metrics or the bottom line results that interest patients, healthcare staff, administrators, and payers. For example, the most common patient diagnoses treated by the BHP are operational metrics.
Outcome Metrics | Clinical improvement as assessed by objective, self-report or biometric measures for both mental (e.g., PHQ-9s, GAD-7s, functional status) and general health conditions and health behaviors (e.g., Hemoglobin A1c, cholesterol levels, smoking status, alcohol use, blood pressure, body mass index).
Process Metrics | These metrics help one understand the extent to which one is reaching your operational and outcome metrics. They reflect evidence of a variety of processes a practice may have in place to deliver care and produce the desired outcomes. For example, the percentage of patients who have been referred to the BHP is a process metric.
Physician Quality Reporting System (PQRS) | A quality reporting program encouraging individual eligible professionals (EPs) and group practices to report information on the quality of care to Medicare. Psychologists are considered EPs. PQRS gives participating EPs and group practices the opportunity to assess the quality of care they provide to their patients, helping to ensure that patients get the right care at the right time. By reporting on PQRS quality measures, individual EPs and group practices can also quantify how often they are meeting a particular quality metric.
Practice Management System (PMS) | Assists with managing, scheduling, billing, and capturing patient data while complying with the Health Information Technology for Economic and Clinical Health (HITECH) Act.
Pro Formas | Describes a presentation of data, in financial terms, where the data reflect the world on an “as if” basis. These models are often used to project the financial outcomes associated with any particular clinical initiative or program.
Quality Metrics | This refers to the clinical outcomes or “health” of a patient as defined by the particular clinical inventory (e.g., PHQ-9) or biometric (e.g., blood pressure). In discussing quality metrics, healthcare systems often refer to how healthy their patients rate on any number of indices as a function of the care the organization provides. It is similar to outcomes/clinical metrics, but this terminology is usually used by healthcare systems to encompass broader components such as patient satisfaction, how care is delivered, and the patients’ experience of their care.

**Model-Specific Differences in Selection of Psychological Measures**

It is worthwhile to mention that the particular model of integrated behavioral health one is practicing may lead to differences in assessment instruments used. This is because a model like co-located specialty mental health, by definition, delivers a highly specialized service, whereas the primary care behavioral health model functions as the primary level of behavioral health care, comparable to primary care itself. BHPs are cautioned not to allow the time allocated to determine the level of assessment they provide. This is tantamount to an internist providing two different levels of care for a particular presenting problem based on how much time he or she happens to have for that patient. Ethically, it is important to provide the care that is clinically indicated. If the patient’s problem requires a level higher than primary care, then the patient needs to be referred to that higher level. Some BHPs may also provide different levels of care based on the patient’s severity. This may be a risky practice. Without following one clear standard of care reflective of one’s role, the service a BHP provides can be misleading to patients and...
confusing for providers. Ultimately, it may unnecessarily complicate analyzing the program-level outcomes. If one provides a specialty level of care, then the instruments used should be highly specialized. By contrast, a BHP who provides a primary level of care are advised to use instruments that the primary care providers are able to use (e.g., ones that don’t require additional training in psychometry and that do not exhaustively evaluate the “psychiatric system” of the body).

**Documentation**

Acknowledging that local laws (which continue to evolve) may not support documenting mental health or substance use information in the same medical record as primary care, behavioral health practitioners working in primary care must document in accordance with their role. Specifically, one benefit of working in an integrated setting is the sharing of information between primary care providers and behavioral health clinicians. This means behavioral health assessment information about patients must be entered into the shared EMR wherever local law permits it. Critically, BHPs must document with the primary care team as the audience. They are the most important audience. This means that unlike documenting assessment results in traditional mental health settings, providers must synthesize the findings into brief summaries and recommendations that apply to case managers, primary care providers (PCPs), nurses, and other members of the healthcare team.

While tailoring their documentation to the primary care audience, BHPs also must ensure their documentation clearly distinguishes the model of integrated behavioral health care they are providing. This is important for a couple of reasons. First, unlike in traditional settings, these notes may be viewed by many people, who will make other assessment, referral, and treatment decisions based upon the BHP’s notes. If the primary care team believes the patient has received extensive psychological assessment for a condition (i.e., co-located specialty mental health) but has actually received a brief assessment and intervention (primary care behavioral health model), the primary care team may draw incorrect conclusions about the severity of the problem, the prognosis, and the implications of the assessment results. This underscores the importance of ensuring your primary care staff appreciate the strengths and limitations of whichever model of integrated care a BHP delivers.

Second, depending on the organization and setting (accountable care organization [ACO], federally qualified healthcare center [FQHC], private group primary care practice, hospital, etc.), information gleaned from EMRs are often used within the larger EHR to make calculations, estimations, and other conclusions about the system such as financial sustainability, population health, value of the services, and effectiveness of the program. If healthcare administration and informatics personnel believe the program accomplishes “x” and costs “y,” but due to unclear documentation it actually accomplishes “w,” then problems may arise in funding, resourcing, scaling, or sustaining the program. In short, it leaves the system vulnerable to drawing false conclusions. This means it’s important to have this data entered in an objective format that is easy to “mine.” It also means documentation must illustrate the type of care the BHP provides. A grave mistake would be to unintentionally lead one’s employer or payor to believe that a co-located specialty assessment (e.g., for chronic pain) is equivalent in time, quality, and level of specialty to a brief chronic pain evaluation performed in primary care behavioral health.

Consider that many of the medical systems that hire mental health providers for integrated positions may not understand the implications of these various integrated care models. For example, suppose the organization hires a behavioral health consultant, and then appreciates their productivity and contributions, believes they can be used in many other business models it has derived to help meet its population’s health needs, recognizes them as a psychologist, and therefore decides to hire another psychologist. This does not mean that the new psychologist will provide the same model of
care that the first psychologist has documented and delivered. This is because in most cases, practicing in primary care is not simply a matter of practicing in a different location. Rather, it usually means the psychologist must obtain additional training to develop the skills needed to function within a primary care team. Thus, if the system hires the new psychologist predicated on the false assumption that all psychologists have the same training, the integrated behavioral health service, the patients, and the organization all suffer. In summary, documentation itself affects program outcome measurement, and participating in how the organization chooses to measure these program and system outcomes is imperative for accurate analysis and subsequent decision-making.

Screening

Screening should be considered a part of treatment and outcome monitoring process because the primary care field employs a population health approach—something typically unfamiliar to specialty mental health providers and settings. This brief section helps BHPs understand the primary care frame of reference regarding screening, assessing, and treating conditions (i.e., through clinical pathways) and why this is so important to pursuing the Quadruple Aim. In short, primary care leadership and staff select specific conditions to manage within their patient population, and the first step to managing the conditions is to initiate screening. Those conducting psychological assessment in primary care may be asked to assist in selecting target conditions for the clinic, screening instruments, and the treatment outcome instruments. BHPs must be prepared to understand the priorities of the primary care field in order to make a valuable contribution to the process of managing population health, which requires outcome monitoring at the patient and system levels.

Placing the discussion of screening in a broader context, consider that the United States spends more than any other comparable nation on health care when calculated as a percentage of our country’s gross domestic product (GDP). Our costs are 32% higher than the country ranking second in such spending and the average of the other 20 comparable nations is 9.3% (Organization for Economic Cooperation and Development [OECD], 2015). Other estimates of US healthcare spending find that the cost per capita has been rising over the last 50+ years and in 2013 was $9,523 per person (OECD, 2015). These figures are illustrated in Figures 7.1 and 7.2.
This pattern of high costs has not resulted in Americans being healthy. In fact, America is among the worst third of OECD (Organization for Economic Cooperation and Development) nations with the highest incidence of diabetes mellitus, overweight children and adults; and avoidable hospital admissions for adult asthma, COPD, and diabetes mellitus (OECD, 2015). While there are many problematic aspects to the current healthcare system, it has been known for years that comorbid mental health problems are costly (Table 7.2), and resolving these could mean one giant step of progress toward meeting the Quadruple Aim.

Simply detecting comorbid mental health disorders is not the only way to improve health and decrease costs. BHPs can address a variety of mental health and physical health concerns patients present to their primary care provider (Strosahl, 2000). In order for a BHP to efficiently find the patients who may benefit from these services, especially within a formal care management model (University of Washington AIMS Center, IMPACT Tools: http://impact-uw.org/tools/), screening can be helpful.

Logically, one might ask, “Who should be screened?” According to the under the Agency for Healthcare Research and Quality (AHRQ), screening should only be done if (1) the condition is common within the population, (2) there are well-supported methods for applying behavioral health expertise to treat the condition, (3) it leads to early intervention and therefore more effective treatment, (4) the instrument used has high specificity and sensitivity, and (5) the instrument is feasible (i.e., easily completed and acceptable to patients, costs little relative to allowing the conditions to persist unaddressed; Talen, Baumer, & Mann, 2013).

### Table 7.2 The Cost of Co-occurring Mental Health Disorders and Chronic Conditions

<table>
<thead>
<tr>
<th>Condition</th>
<th>PPPY Cost—Those Without MH Condition</th>
<th>PPPY Cost—Those With MH Condition</th>
<th>The Cost of Co-occurring Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart Condition</td>
<td>$4,697</td>
<td>$6,919</td>
<td>$2,222</td>
</tr>
<tr>
<td>High Blood Pressure</td>
<td>$3,481</td>
<td>$5,492</td>
<td>$2,011</td>
</tr>
<tr>
<td>Asthma</td>
<td>$2,908</td>
<td>$4,028</td>
<td>$1,120</td>
</tr>
<tr>
<td>Diabetes</td>
<td>$4,172</td>
<td>$5,559</td>
<td>$1,387</td>
</tr>
</tbody>
</table>

The US Preventive Services Task Force (USPSTF) under the AHRQ has published several behavioral health screening recommendations for primary care (2010). The USPSTF recommends (Grade B) or highly recommends (Grade A) screening for the following conditions: smoking screening and tobacco cessation; obesity screening and counseling for children and adults; depression screening for adults; alcohol misuse screening and counseling for adults and pregnant women (Talen et al., 2013). Therefore, if one is interested in lowering healthcare costs and improving a population’s health by proactively screening the entire population of a practice, consider screening for any of the aforementioned conditions. Also, one should consider working with their organization’s business and administration in order to identify conditions that are worthwhile for screening in their patient population. Fortunately, within most integrated primary care settings, team-based models of care are the vehicle. Thus, if one develops a screening program well, it will garner conceptual buy-in and manpower support for the program.

Within primary care, integrating screening with the patients’ routine check-in and vital signs processes are the most advisable way to implement these processes. The nurse or medical assistant will often screen patients with the instruments the clinic decides to use. Asking a brief set of questions can be rapidly completed while taking one’s vital signs. Alternatively, some clinics hire numerous BHPs so that one of these personnel can provide a brief screening before every clinic appointment (usually during vital signs).

What conditions the clinic chooses to screen for is an entirely different issue, as it often involves collaboration with the business side of the practice. This means it is often equally as contingent on business cases as it is on patient needs. This may be a drastic shift for clinicians accustomed to practicing in traditional mental health settings. The primary care setting requires team efforts and team-based solutions for addressing the population’s health. Discovering which needs one’s population possesses requires a fair amount of practice data analysis. Before launching any screening, it would be beneficial to consult with those who can help you mine data from the EHR or practice management system.

No screening instrument or other outcome assessment tool should be used if it does not meet psychometric standards. While in-depth discussions of the psychometrics of various instruments for various conditions is beyond the scope of this chapter, consider that using an instrument that has inadequate specificity, sensitivity, construct validity, or face validity will result in false positives and negatives, and may also have a negative effect on the patient experience of care in the clinic. If the screening tool will also be used regularly, there must be appropriate test-retest reliability. The screening instrument must also align with any criteria needed for diagnosis (including DSM-5 and ICD-10).

It is particularly important that BHPs use screening methods after gaining the support of the organization’s leadership and the clinic stakeholders (i.e., nursing, providers, and management). Without such collaboration, one ends up launching a mini-mental health clinic within a primary care clinic instead of a well-integrated program that services all patients in the clinic. When the BHP makes himself/herself extremely relevant to the priorities of the clinic, then they gain support for initiatives such as screening. Finally, considerations should be made about how frequently to administer such screenings so that it is not inconvenient or redundant for patients. Imagine if a patient is screened for smoking and depression at every appointment. If the patient made an appointment multiple times in one year or is routinely following up per the provider’s instructions, the screener may have these unintended effects.

**TREATMENT MONITORING AND OUTCOMES ASSESSMENT**

Treatment monitoring in primary care often involves using official measures of one’s symptoms, functioning (e.g., occupational, social, emotional, physical), or some biological metric. Most BHPs
working in primary care will use self-report measures to monitor treatment progress, although biological metrics will be used when providing behavioral medicine services. For example, a patient’s blood pressure, hemoglobin A1c levels, and body mass index are often monitored by the BHP during treatment in order to gauge the patient’s response to a specific intervention or series of interventions. By contrast, outcomes assessment is the final measure of the treatment outcome. Because primary care involves short courses of care for many different clinical problems, the culture is fast-paced and sometimes chaotic, and appointments usually last 15–30 minutes, BHPs will often use the same self-report instrument or biological metric for monitoring as they will for outcomes assessment. However, in some cases outcome assessment will be a more lengthy process that includes an examination of whether or not the patient still meets diagnostic criteria for a clinical problem.

**Why Bother?**

The American Psychological Association (APA) has emphasized the importance of assessing outcomes (APA, 2006; Kazdin & Blase, 2011; Newnham & Page, 2010). In addition, research has shown that when patients report their outcomes via *patient-reported outcome measures* (PROMs), it helps improve the quality of care (Nelson et al., 2015). Also, this aligns the clinician’s and patient’s views of what is happening clinically so that treatment can be tailored to the patient’s preferences and needs (Barry & Edgman-Levitan, 2012). PROMs also lead to better communication and decision-making between doctors and patients, improved patient satisfaction with care, and better patient-reported outcomes throughout the course of treatment. In addition, measurement feedback systems, that is, providing clinicians with feedback about the clinical impact of their treatments, have led to faster improvement for patients (Bickman, Kelley, Breda, de Andrade, & Riemer, 2011). Thus, clinicians and patients alike benefit from discussing the results from the treatment outcome measures.

In addition, there is a difference between objective and subjective measurement. Clinicians who historically practiced primary and mental health in nonintegrated settings may have used subjective outcome measurement (e.g., “How are you feeling today compared to the last time we met?”). The problem with this type of measure is that it is not a standard method of assessing progress. It is also vulnerable to an array of biases (e.g., response bias). Objective self-report measures (e.g., PHQ-9, GAD-7, SF-12v2) are more beneficial as they are psychometrically sound and also eliminate some, but not all, biases. Other objective physical variables (e.g., body mass index, blood pressure, hemoglobin A1c) are perhaps the most helpful types of variables when they respond to a particular treatment in an expected manner. These are how the primary care practices measure progress and therefore bear high relevance and acceptability by the rest of the primary care team. Additionally, these measures are not subject to the self-report biases. The problem is that many of these biometrics do not change as a result of treating common mental health problems (e.g., anxiety, depression). Therefore, it is advisable to select instruments that measure the patient’s functioning and the presence of symptoms. For clinical conditions that contain subjective diagnostic criteria (e.g., pain is a subjective report of sensations; insomnia is the subjective report of inadequate sleep), measuring functioning versus symptoms may be the best approach to measuring as objectively.

While assessing outcomes may be time-consuming, it is feasible to implement if appropriately adapted to the primary care environment. After all, primary care itself seeks to help patients self-manage their symptoms and chronic conditions and improve their functioning. With the team-based models of care delivery spreading across integrated primary care settings, it is feasible to secure the assistance of other team members who can help administer these outcome instruments.
What Clinical Outcomes Do I Measure?

Generally, BHPs are advised to always consider the Quadruple Aim when selecting instruments for screening or assessment. If the instruments are too long, the patients may experience their care negatively. Similarly, if the procedures or instruments are too complex, technical, or cumbersome, it may place undue burden on the primary care providers. Many BHPs have encountered significant problems trying to infuse a mental health culture (including all its traditional cultural and clinical practices) into primary care. These problems include being fired, not having patients referred to them, or deciding to leave the job due to not adapting to the culture in which they work. Unfortunately, many of the outpatient mental health practices are too lengthy and specialized for delivery in primary care. BHPs must adapt the functional aspects of their specialty to the modern day pace, milieu, and purpose of primary care. Thus, they must address acute conditions and manage chronic conditions between “flare-ups” in order to prevent disease, empower patients, and meet the Quadruple Aim.

Because the primary care team helps administer them, providers who routinely use assessment tools in the context of clinical practice will need brief instruments that are implementable in primary care and easily understandable by the rest of the primary care team. Note that “easily understandable” means it doesn’t require one to read the instrument’s instruction manual, and the implications of its results are comprehensible in less than a minute with no extensive explanation of psychometrics or psychological jargon. This is primarily because PCPs and other team members have little time in the chaotic primary care environment to read extensive notes or learn about every instrument a BHP uses. Moreover, without face validity, primary care teams will be reticent to administer these, practice leaders will be hesitant to require it, and patients will refuse to complete them.

There is benefit to the BHPs using a general measure of mental health and physical health symptoms and functioning (i.e., academic, occupational, interpersonal, emotional, social, physical). This helps communicate to the primary care teams a snapshot of how the patient is doing. The more quickly and effectively BHPs communicate or illustrate this (through EMR documentation), the higher value they have within the delivery of integrated care. And if the primary care teams cannot understand the BHP’s contribution to the primary care mission, they will be unlikely to utilize the services and the BHP’s impact will be limited. A recent RAND study of the 56 integrated care SAMHSA grantees corroborates this notion—that poor utilization of integrated behavioral healthcare services is common, as is poor fidelity of the particular model of implementation (Scharf et al., 2014). If primary care teams don’t know how to use the BHP or don’t understand how the BHP’s assessment information is relevant, integration fails. The service simply will not be used. Therefore, an underlying goal for all assessment in primary care is to make one’s psychological findings relevant and useful to the primary care team, thereby increasing the value of the service the BHP provides.

Of course, there are limitations to using a general measure of health and health functioning. The results may be too vague to drive some treatment decisions. Or constructs measured by the instrument (e.g., social functioning) may be derived from one item. Therefore, when feasible BHPs are advised to use the symptoms specific outcome instruments. Common ones are listed in Table 7.3, although the other chapters in this book culminate a more exhaustive set of resources.

Lastly, when involving other staff members in the administration of the general or condition-specific instruments, the BHP is advised to train these staff members and maintain involvement as to how these are implemented clinic-wide. What BHPs usually recognize, but the other primary care staff may not fully appreciate, is the test-retest reliability of an instrument. Therefore, the BHP plays an integral role for the clinic in ensuring that the instruments used within the behavioral health services and by the other primary care members are not used so frequently (e.g., every time the patient attends a primary care appointment) that the instrument becomes unreliable over time. There is
one final note about the nature of outcomes assessment in IBH. While the instrument a BHP selects may be very practical and fitting for the particular IBH model and integrated primary care setting, because the case is never closed in primary care, BHPs may see patients at different frequencies and for different durations than in specialty mental health settings. In the latter settings, patients are referred to the mental health provider for a specific problem that is severe enough to require mental health assessment and treatment. Treatment is therefore contained to a specific period in which the mental health disorder is present. By contrast, in primary care, patients having even minor clinical problems can be referred. Also, the same patient can be referred multiple times throughout the same year for various ongoing conditions—mental health (e.g., depression or anxiety that are in remission and for which the primary care team has requested a “check-in”) and general health problems (e.g., ongoing problems managing diabetes or obesity). This means that for some patients the BHP must select instruments that can be readministered every week, if needed, during the course of treatment for a given condition. Additionally, the longitudinal relationship between BHP and patients may be multigenerational, as is the case for many primary care providers because in most cases the BHP sees whomever the primary care team sees. For example, a BHP may assist the primary care team in assessing and treating a child for behavioral problems, then a few years later the BHP may assist with education about sexually transmitted infections, weight management, and avoiding alcohol and drugs. Once that same patient reaches adulthood, the BHP may assist the primary care team in treating that same patient for postpartum depression and eventually begin seeing the patient’s child as time continues. In short, when practicing in primary care, the services are ongoing and treat the entire person (mind and body). It is not as time-limited as specialty mental health. As such, BHPs should be prepared to decrease the frequency of measuring these outcomes when each course of treatment is completed. Subsequently, there may be a benefit to readministering the screening or outcome instrument regularly thereafter (e.g., annually). This would mirror the primary care practices for mammography, colonoscopy, and other assessments that are conducted at regular intervals in the absence of the related condition (i.e., breast and colon cancer in these particular cases). Similar to how the dental model works, the primary care team may ask a BHP to provide annual check-ups for behavioral medicine and mental health problems of all the patients enrolled to the clinic. Because payment models for integrated care are evolving, this type of care delivery may be common in the coming years.

### MEASURING CHANGE WITHIN THE CONTEXT OF TREATMENT

In the primary care environment, there are many moving parts, particularly within team-based care settings. Because time is limited, a high volume of patients are seen, and primary care is responsible for identifying and managing acute conditions, while also managing ongoing chronic ones. Thus,
providers must work efficiently. The work requires evaluating and treating numerous patients, with limited information available for each, frequent communication with other staff, and coordination of care. Therefore, it is critical that BHPs select measurement tools that are appropriate and practical for this setting. They must administer these instruments regularly and quickly, and in a way that accurately measures clinical change. Finally, they must effectively communicate changes in patients’ symptoms or functioning to the primary care team.

**General Considerations**

It is important to monitor every patient at every encounter (perhaps except telephone consults) in order to ascertain evidence of the treatment progress and outcome. Shorter instruments offer the most promise, providing they are equally as valid and reliable. Due to the team nature of many contemporary primary care clinics, the front desk staff may be able to ask every patient who arrives to complete the global symptoms/functioning instrument you select. If this is not feasible, the team member who takes the patient’s vitals is often selected to administer the measure for the BHP, particularly if the patient is being referred same-day to you as a “warm handoff.”

Lastly, it is advantageous to have the patients complete the tool at the beginning of the appointment rather than at the middle or end. This way the results are contingent on how the patient has been since the last appointment, not how the patient is feeling due to something that occurred during the appointment (e.g., instilling hope in the patient). It is well documented that hope influences patients’ self-reported outcomes (Di Blasi, Crawford, Bradley, & Kleijnen, 2005; Eaves, Ritenbaugh, Nichter, Hopkins, & Sherman, 2014; Kaptchuk et al., 2009). By contrast, bodily sensations may also influence patients’ self-reported progress, with particularly negative sensations being correlated with negative self-reports (Petersen, Van Staeyen, Vogele, von Leupolt, & Van den Bergh, 2015). For example, if a patient feels a nonspecific pain in her abdomen, she may also report higher anxiety, poorer health, or the worsening of her health. Taken together, this means dynamic factors influence how patients report their symptoms at each appointment. One should assess before introducing more variables (e.g., through interventions and the course of assessment and treatment) and consistently assess at the same juncture during each appointment to minimize biased reporting.

**Common Statistical Measures of Change**

One of the challenges faced by BHPs in measuring outcomes is the interpretation of the numeric score changes on a clinical instrument over time. While some instruments have published data informing users about the statistical measures of change, other instruments do not. For example, for the PHQ-9 and GAD-7 it is well known that a 5-point change signifies statistically significant change (Titov et al., 2013). This means if a patient’s score changes by 5 points, the BHP can conclude that this change is not spurious and suggests it may be clinically meaningful. When instruments do not have published measures of change statistics, these can be calculated.

Two types of statistics one may consider calculating are the **reliable change index** (RCI; Jacobson & Truax, 1991) and the **minimal clinically important difference** (MCID; McGlothlin & Lewis, 2014). The RCI reflects a ratio in which the numerator is the observed difference between two scores and the denominator is the standard error of measurement of the difference. This method of measuring change is maybe less meaningful to patients, as it exclusively captures the nature of the change in a numeric score on an outcome measure. Yet, it also may be statistically robust.
The MCID, by definition, is a measure of change that is patient-centered, as it captures a magnitude of the change as well as the value patients place on that change. MCIDs are more difficult to calculate, partly because small changes may be meaningful to patients but difficult to demonstrate statistically. Thus, statistical significance is linked to sample size (McGlothlin & Lewis, 2014). There are also multiple ways to determine MCIDs, which makes it challenging.

There are distribution-based and anchor-based methods, which may be feasible statistically, but the patient’s perspective may also be lost (Copay, Subach, Glassman, Polly, & Schuler, 2007). For practical use, BHPs are advised to select instruments that have published RCIs. If one wishes to use MCIDs, they may need to be uniquely calculated for the outcome measure BHPs used, as each method for calculating the MCID gives different values. Furthermore, the change measured on the instrument varies based on the patients’ baseline scores. In other words, change is measured in relative terms to the patient’s other scores, rather than each score being compared to a standard scale of possible scores. While arbitrarily defined ranges have been recommended to help overcome this challenge, this method is difficult and carries with it limitations (Copay et al., 2007). Of course, neither the RCI nor the MCID account for the cost of delivering care, making these important, but not comprehensive measurement of change statistics where the Quadruple Aim is concerned (i.e., when measuring change at the program or system level).

**Monitoring Programs and Assessing System Outcomes**

As stated earlier, in the current healthcare landscape BHPs may be charged to calculate the value of their services. This may be defined any number of ways: actual dollars spent on the IBH service versus dollars collected for the IBH service, or some other calculation. A BHP who is reasonably familiar with program outcome evaluation may be best able to rise to this challenge. Comprehensively determining the return on investment (ROI) for one’s program, including cost savings and offset, is beyond the scope of this chapter and have been published elsewhere (e.g., Corso et al., 2016). But what is well within the current discussion is how to develop and analyze the various types of program and system metrics to measure program and system progress (i.e., pursue the Quadruple Aim).

Methods for collecting these metrics vary depending on the specific practice’s administrative and clinical systems. Perhaps the richest sources of data in a practice or system is the EHR and the practice management system, as these contain data that can be used to calculate the cost of care, health outcomes at a population level, and so forth. In systems where these information technology tools are scarce or rudimentary, measuring program and system outcomes becomes much more labor-intensive.

Program monitoring and outcomes assessment should begin by measuring implementation metrics. These are the variables that provide information about how developed, established, or operational the IBH program or service is. These include things like mission, leadership, funding, materials, space, staff, training, administrative policies, standard operating procedures, and business rules. In many cases, these are simply tracked in terms of what is done and what is left to complete, with specified expected dates for completion.

Next, there are process and outcome metrics. These help one understand the ultimate outcome measures without necessarily serving as the ultimate measures. See Tables 7.4 and 7.5 for examples of these. Selection of metrics from among these and other metrics can be an important preliminary step to ensuring your IBH service produces the desired outcomes of the organization. Collecting these metrics may be onerous due to the information technology capabilities of one’s practice. The stronger relationship a BHP develops with the personnel who conduct administration, finance and EMR improvements, the more easily he or she will be able to measure program and system variables.
<table>
<thead>
<tr>
<th>Quadruple Aim Goal</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience of Care</td>
<td>• Percentage of patients asked to sign a release of information consent to allow agencies to exchange information</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who sign a release of information consent</td>
</tr>
<tr>
<td></td>
<td>• Frequency of contacting patients’ other providers to coordinate care</td>
</tr>
<tr>
<td></td>
<td>• Referral “hit rate,” or the number of PCP-referred patients who actually accept the referral or warm handoff</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who were asked to complete a healthcare satisfaction measure</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who completed a healthcare satisfaction measure</td>
</tr>
<tr>
<td></td>
<td>• Reasons patients did not complete satisfaction measure</td>
</tr>
<tr>
<td></td>
<td>• Amount of time (e.g., same day, 3 days out) to next available IBH appointment</td>
</tr>
<tr>
<td></td>
<td>• Patient educational materials are culturally appropriate and written in a language and at a level that best meets the patients’ needs</td>
</tr>
<tr>
<td></td>
<td>• Availability of staff who speak the same language as the population being served</td>
</tr>
<tr>
<td>Population Health</td>
<td>• Number of patients seen by the BHP in a week/month/quarter</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients seen in IBH who were screened for a given problem (e.g., depression)</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who screen positive for a problem</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who screened positive who were referred to the BHP for further assessment or intervention</td>
</tr>
<tr>
<td></td>
<td>• Reason patients who screen positive were not referred to BHP for further assessment or intervention</td>
</tr>
<tr>
<td></td>
<td>• Average number of clinic visits per patient per quarter (are those who need to be seen to ensure ongoing good health being seen regularly?)</td>
</tr>
<tr>
<td></td>
<td>• Reasons patients with a given problem (e.g., diabetes) are not attending clinic appointments per recommended guidelines</td>
</tr>
<tr>
<td></td>
<td>• Percentage of PCP patients who have been referred for IBH (IBH service penetration rate)</td>
</tr>
<tr>
<td></td>
<td>• Percentage of PCP patients who have been treated by BHP</td>
</tr>
<tr>
<td></td>
<td>• Percentage of PCP patients—by diagnosis or panel—who should have been referred for IBH (are patients receiving appropriate evidence-based care?)</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients with a clearly documented integrated treatment plan</td>
</tr>
<tr>
<td>Cost</td>
<td>• Percentage of patients who were referred to the BHP who kept the appointment</td>
</tr>
<tr>
<td></td>
<td>(patients with poor follow-up may have worse health, therefore demanding a higher overall treatment cost from the payer)</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who kept initial BHP appointment that were seen more than once</td>
</tr>
<tr>
<td>Provider Wellness</td>
<td>• Percentage of patients who were referred for a BH appointment outside of the primary care clinic</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients who were referred that kept the BH appointment outside of the primary care clinic</td>
</tr>
<tr>
<td></td>
<td>• Type and duration of IBH treatment</td>
</tr>
<tr>
<td></td>
<td>• Percent of patients newly prescribed psychotropic medication</td>
</tr>
<tr>
<td></td>
<td>• Percentage of patients prescribed psychotropic medications who fill their prescription</td>
</tr>
<tr>
<td></td>
<td>• Number and type of diagnosis for patients who have high emergency room utilization</td>
</tr>
<tr>
<td></td>
<td>• Number and type of diagnosis for patients who have higher hospital readmissions</td>
</tr>
<tr>
<td></td>
<td>• Percentage of primary care staff vacation days used versus unused</td>
</tr>
<tr>
<td></td>
<td>• Average duration of lunch and breaks of primary care staff</td>
</tr>
<tr>
<td></td>
<td>• Average number of days in a given period that primary care staff take lunch and breaks</td>
</tr>
</tbody>
</table>
Quadruple Aim Goal | Metric
--- | ---
• Average number of days in a given period that primary care staff take breaks
• Average number of hours primary care staff spend doing clinic-related work after the normal clinical hours (including documentation or correspondence completed at home, etc.)
• Percentage of telephone consults, secure messages, and other communication/documentation that occur during nonclinical work hours among primary care staff
• The development of policies that make primary care staff wellness a priority (implementation metric)


### Table 7.5 Outcome Metrics That Help BHPs Pursue the Quadruple Aim

<table>
<thead>
<tr>
<th>Quadruple Aim Goal</th>
<th>Metric</th>
</tr>
</thead>
</table>
| **Experience of Care** | • Level of patient satisfaction with access to general health services
• Level of patient satisfaction with accessibility to IBH services
• Level of patient satisfaction with effectiveness of physical health services
• Level of patient satisfaction with effectiveness of IBH services
• Level of primary care provider satisfaction with delivery of IBH services
• Level of primary care staff knowledge and comfort level in IBH service provision
| **Population Health** | • Patient quality of life functioning (e.g., score on a quality of life measure)
• Patient mental health functioning (e.g., score on a mental health measure)
• Patient general health status
• Patient general health indicators (e.g., body mass index, waist girth, weight, blood pressure, blood glucose levels, lipid levels, pain level, alcohol use, physical activity, tobacco use)
• Percentage of improvement of number of enrollees in a given measure (e.g., body mass index or tobacco use) compared to previous year
| **Cost** | • Annual percentage increase in per capita costs
• Emergency room visits per 100 enrollees per year for any reason
• Emergency room visits per 100 enrollees per year for mental health presentation alone
• Frequency of psychiatric hospital admissions
• Frequency of hospital admissions
• Number and severity of general health and BH relapses
• Rate of appropriate psychotropic prescription
• Decreased overall specialty healthcare use
| **Provider Wellness** | • Primary care staff burnout ratings (see Robinson, Gould, & Strosahl, 2011)
• Primary care staff job satisfaction ratings (e.g., Likert scales)
• Number of primary care staff who quit over a particular time period
• Number of primary care staff who were fired over a particular time period


### SUMMARY

While the mental health field continues to penetrate the primary care domain, the practice of psychological assessment must maintain its standards and adapt its delivery and instrument selection to the primary care culture. Essentially, this means using outcome measures that can easily be taught to and administered by non-BHPs. It also means ensuring that documentation accurately reflects the degree
of specialization of the BHP’s IBH model. Documentation and outcome measure selection must also enable facile data mining processes and require additional manpower to accomplish data collection, entry, and retrieval. Finally, as the primary care community welcomes BHPs into this domain, the business and healthcare administration within that community are empowering BHPs to measure patient, program, or system outcomes. While this presents an excellent opportunity, it also carries challenges. BHCS may not possess the expertise and experience to measure metrics such as return on investment, cost savings, RCI, MCID, and physician burnout. Yet as they pursue the Quadruple Aim with their medical colleagues, BHPs face an opportunity for professional growth where their contribution may be vital to this future of US health care and the field of integrated primary care.

REFERENCES


At its most basic level, an interview is “a conversation directed to a definite purpose other than the satisfaction in the conversation itself” (Bingham and Moore, 1959, p. 3). In the case of the clinical interview, this purpose is to gather sufficient information to correctly identify a problem and develop a plan to address that problem. At its core, the clinical interview in any treatment setting is the engagement of a help-seeking individual by a knowledgeable expert who has the formidable task of understanding the individual in her or his uniqueness in the span of 5 minutes to approximately an hour (MacKinnon, Michaels, & Buckley, 2015; Sullivan, 1970).

THE CLINICAL INTERVIEW

The clinical interview is perhaps the most important interaction that exists between patient and provider. It is the foundation upon which all future care rests, both interpersonally and diagnostically. This is the case no matter the clinical field or provider. Even in the case of the specialist physician, the clinical interview impacts care quality more than the physical exam or diagnostic testing, in no small part because the clinical interview guides the application and accuracy of both (Hampton, Harrison, Mitchell, Prichard, & Seymour, 1975; Sandler, 1980; Vroomen, De Krom, Wilmink, Kester, & Knottnerus, 2002). Although the clinical interview is equally important regardless of the clinical field of practice, the manner in which the clinical interview is conducted may vary greatly depending on the clinical field and setting. This is, of course, a product of the available diagnostic methods, the environmental constraints of the setting in which care is provided, and the ability of the patient to effectively engage with the clinician in the interchange.

It is also important to remember that the clinical interview is important not just for the ostensible goal of an interview (i.e., to gather a verbal narrative of the conditions of concern to date) but also for its role in the establishment of the patient–provider relationship. The clinical interview is the foundation of rapport between provider and patient. The patient collects “data” about the provider almost as much as the provider does the patient. Therefore, the initial clinical interview represents the critical first impression the provider makes with the patient and fosters trust, which, consequentially, affects the quality and accuracy of the information gathered. This trust also affects clinical outcomes beyond data accuracy. The level of rapport the clinician can establish predicts a patient’s adherence to provider’s recommendations and his or her engagement in their medical care (DiMatteo, 1993).

The broad goals of the interview are the same whether conducted by a physician in a primary care setting or by a mental health professional: engage the patient in the therapeutic alliance, collect valid information about the presenting concern and about the patient personally so the presenting concern can be understood in context; foster trust and rapport with the patient; develop a tentative diagnosis and treatment plan; lessen the patient’s apprehension and worry about both the presenting...
The Mental Health Clinical Interview

Although the clinical interview in any treatment setting is an important tool to better understand the patient’s problem and effectively develop a plan for treatment, in the mental health setting the clinical interview is the primary tool the clinician has to conduct an assessment. However, the clinical interview is both a rich and a limited assessment tool. The interview provides an opportunity to gather a rich depth of information; however, this information is largely the subjective report of the patient. Although the patient’s report of their experience is critically important information, its subjective nature creates limitations for its use. For example, when a patient provides an account of certain events, the interviewer cannot know objectively what occurred, but only how the patient perceived the situation, which has likely been influenced by the meaning the patient made from the experience. Of course, this is assuming that the patient is relating the truth as they know it. In order for the patient to relate events truthfully, trust must be established, which usually requires time to develop. It is for this reason that, historically, the mental health interview has been conducted over multiple sessions, particularly in psychodynamic practices (Sullivan, 1970). However, in modern mental health treatment, few have the luxury of several hours to devote to the establishment of trust with the patient in order to gather an open and complete history, and there are few settings in which this is more salient than while conducting interviews in integrated primary care. As discussed later in this chapter, the ability to rapidly establish rapport with patients is an essential skill in the practice of integrated care, and integrated behavioral health providers (IBHP) must utilize all resources at their disposal to do so.

With the exception of those trained in programs that use facilics (Shea & Barney, 2007) for formalized supervision, many mental health clinicians do not receive formal training in the clinical interview; rather, they are instructed in the practical and philosophical goals of the interview. Utilizing a conversational style with which to build rapport is prioritized over structure and systematic information gathering. Students are informed that certain information must be gathered to conduct a thorough assessment; however, this list of required information is often presented in broad strokes (for example, educational and occupational history), without specific items to inquire about. This method of training creates a “trial-and-error” learning process wherein the learner begins to develop a gestalt of what information is needed to arrive at an accurate diagnosis (Deutsch & Murphy, 1955). MacKinnon, Michaels, and Buckley (2015) assert that in the case of mental health providers, the clinical interview cannot be supervised in the same way as training of medical students is conducted and therefore it may take years for a provider to become skilled. However, given our experience in training medical students in interviewing, we assert that such intensive training is possible with adequate resources (which may be the real barrier). It is possible that the decreased time demands (i.e., a 50-minute appointment) in comparison to limited time allotted for the medical appointment accommodates this trial-and-error learning because the efficiency demands are less.

In mental health interviewing, establishing rapport is the prioritized goal. MacKinnon, Michaels, and Buckley (2015) state, “An interview that is centered on understanding the patient provides more valuable diagnostic information than one that seeks to elicit psychopathology” (p 4). Because diagnoses can be determined over time and treatment methods are often quite similar for different diagnoses...
(e.g., cognitive behavioral therapy, or CBT, for both depression and generalized anxiety addresses distorted cognitions), complete and accurate diagnosis need not come at the expense of rapport. This is, of course, not always the case, particularly in the case of prescribing medication. However, much like a broad spectrum antibiotic may be prescribed empirically before the specific bacterial infection is identified (Van der Eerden et al., 2005), treatments in mental health often address a broad spectrum of problems and can be implemented without a definitive diagnosis.

Structure of the Mental Health Clinical Interview
There may be nearly as many ways of structuring and organizing the mental health clinical interview as there are practicing providers. No one format has been agreed upon. MacKinnon, Michaels, and Buckley (2015) assert that “the interview in most effectively organized around the clues provided by the patient and not around the outline for psychiatric examination” (p. 26). Some experts suggest that it may be helpful to follow a general outline of different content domains (APA, 2006; Carlat, 2005; Morrison, 1995; Othmer & Othmer, 2002). However, when sequential phases of an interview are described, experts typically identify four or five phases (Foley & Sharf, 1981; Shea, 1998; Sullivan, 1970), and most who write on the subject preface any discussion of the subject by stating that the stages of the interview are only suggested to assist learners in conceptualizing a structure of their own.

Shea (1998) suggests five stages of the clinical interview: (1) introduction, (2) opening, (3) body, (4) closing, and (5) termination. On the other hand, Sullivan (1970) suggests “the formal inception,” which includes the initial greeting, orientation to the therapy process, and agenda setting; “the reconnaissance,” in which gathering a detailed history and “getting to know” the patient is the focus; “the detailed inquiry,” in which the provider examines the patient’s “self-system” in order to explore the presenting problem, including a detailed psychosocial history that includes such topics as “childhood disorders of toileting habits . . . attitudes toward competition and compromise . . . the pre-adolescent chum”; and finally “the termination,” in which the provider summarizes the interaction, discusses next steps, and enacts a “formal leave taking between the interviewer and interviewee.” Thus, unstructured mental health interviewing may range from a free-form gathering of information to an interaction of phases with separate but interrelated goals. At the same time, there is a clear distinction between these and the formal semi-structured diagnostic interviews, which may also be used as diagnostic tools but typically lack consideration of rapport development and aren’t developed with the goal of a strong therapeutic relationship in mind.

Strengths of the Mental Health Clinical Interview
Establishing rapport quickly with patients is a critical skill in mental health treatment. The mental health interview therefore emphasizes interpersonal warmth and a conversational tone that is intended to put the patient at ease. Although many structured and semi-structured interviews exist to assist clinicians in assessment and diagnosis, the majority of clinicians still rely on the unstructured interview as the foundation of the clinical relationship. In fact, MacKinnon, Michaels, and Buckley (2015) suggest that

questionnaires may save time and be useful in clinics and other places where professional resources are limited, but that efficiency is obtained at a significant price; it deprives the clinician and patient the opportunity to explore feelings elicited while answering the questions.

(p. 38)

Of course integrated care is a clinic setting in which efficiency may have elevated priority, but rapport is still essential.
The nonstructured format of most mental health interviews allows for individuality and more naturalistic expression. Shea (1998) asserts that a skilled interviewer is “remarkably flexible” and in integrated care, this flexibility could not be more essential. The fact that the unstructured interview is by far the most common approach to assessment (Craig, 2003; MacKinnon, Michaels, & Buckley, 1998; Miller, 2003) despite evidence that suggests more directive approaches result in more accurate diagnoses (Basco, 2003; Cox, Hopkinson, & Rutter, 1981; Helzer, 1981) speaks to the utility and value of flexibility. IBHPs must be flexible in their interview structure, goals, and format. In integrated settings, the initial clinical encounter can range from an appointment that very much resembles the typical mental health clinical interview to a 5-minute greeting, statement of the patient's presenting concern, and a brief description of anticipated treatment frequency and modality.

The strengths of rapport and flexibility in the mental health clinical interview create a patient-centered approach to care. Unstructured interviews conducted by a clinician with the twin goals of establishing rapport and collecting data to inform an understanding of the patient foster patient-centeredness. In an unstructured interview, the patient guides the format with only occasional direction from the clinician, usually to assure all information necessary for a complete intake assessment are available. As one learns the value of and work to enhance patient engagement in their care (Barello, Graffigna, & Vegni, 2012; Meyer et al., 2002; Olfson et al., 2012), the patient-centeredness of an unstructured assessment begins the therapeutic relationship with the patient as an essential and active participant in their care from the very outset and ultimately sets the stage for a collaborative treatment process likely to reduce disengagement from care and enhance treatment outcomes (Laugharne & Priebe, 2006; Olfson et al., 2009).

The Medical Clinical Interview

The medical interview is the most important universal diagnostic tool at a physician’s disposal. In many cases, the medical interview is equally as or even more important than the physical exam for identifying an accurate diagnosis (Hampton et al., 1975; Rich, Crowson, & Harris, 1987). However, the medical interview must focus on data collection in order to create a comprehensive differential for the diagnosis. Further assessment measures (lab work, imaging, etc.) and subsequent treatment all hinge on developing a comprehensive and informed differential diagnosis. An incorrect differential diagnosis can delay treatment by weeks or even months, extending the patient’s suffering or causing harm through applying the wrong medical intervention. Because the differential is so critical, the gathering of complete information about the chief concern and the patient’s history are essential to reaching the correct diagnosis. In order to assure that sufficient information is gathered, the medical interview is typically taught with a specific structure that is practiced with standardized acting patients until such time as the student demonstrates competency in collecting a complete history (in the limited time allotted) and is able to do so in a reasonably conversational manner. Although rapport is acknowledged as an important part of the physician–patient relationship, data gathering is prioritized over rapport building.

The Structure of the Medical Clinical Interview

The clinical interview in medical settings is taught in a highly structured format that students are later encouraged to personalize into a more conversational tone. Nearly all medical schools provide formal instruction in clinical interviewing skills (Novack, Volk, Drossman, & Lipkin, 1993; Weihs & Chapa-dos, 1986) although some studies have shown that these skills can deteriorate over time (Kraan et al., 1990), and therefore practicing providers may deviate from the prescribed format as they gain experience and comfort. Medical students are tested on these specific elements through standardized acting patient encounters throughout their academic career. Although students are also given points for a
natural conversational style, they are praised for an organized structure. The separate elements are so emphasized that many students develop mnemonics in order to avoid forgetting a particular piece of the history.

The structure of the medical clinical interview is as follows:

The opening includes greeting the patient, explaining the process for the visit, and the interviewer’s role in that process. Finally, the opening elicits the chief concern.

The elements of the history of the present illness (or chief concern), often referred to as the Cardinal 7, are as follows:

Location often refers to the part of the body where the symptoms occur.

Quality is a general description of the symptoms. For example, if the patient’s chief concern is pain, the interviewer should ask the patient to describe the pain (sharp, throbbing, aching, gnawing, pressure).

Severity is most often assessed using a Likert scale reflecting discomfort or a narrative description of any alteration in functioning.

Context describes the environment in which the symptoms first occurred. It can be a physical location, a time of day, a stressor, or another impactful circumstance.

Timing refers to the duration, frequency, and temporality of the problem. Any pattern that might be noticed in the timing of symptoms can rule diagnoses in or out.

Modifying factors include any symptom triggers and any treatment the patient may already have tried (and failed or has insufficient treatment response) to ameliorate the chief concern.

Associated signs and symptoms are other changes in sensation or behavior that exhibit a similar pattern to that of the chief concern.

After the Cardinal 7, more general information is gathered either in a targeted way asking about history pertinent to the chief concern or in a comprehensive way as is common for an annual complete physical exam. The remaining sections of the medical interview are as follows:

Attributions reflect the patient’s hypothesis regarding what might be causing his or her symptoms, including what he or she is most worried about.

Review of systems is an inquiry about symptoms common to particular body systems intended to screen for any as yet undiscovered disease process related to the chief concern or past medical history. This section of the interview is often asked as a series of yes or no questions.

Past medical history includes the patient’s history of previous illnesses, injuries, and hospitalizations, as well as medications, lifestyle (sexual health, nutrition, exercise, and substance use), allergies, and illness prevention.

Family history includes a list of the major medical conditions of the patient’s first-degree relatives (parents, siblings, and children).

Social history describes the patient’s current living situation, marital status, where he or she was born and raised, educational attainment, current occupation, life stressors, and general quality of life.

Strengths of the Medical Interview

As mentioned earlier, the medical clinical interview is designed for efficiency. The 10 minutes allowed for a medical appointment (occasionally more time is allowed for an annual physical) necessitates efficient communication. Efficiency doesn’t have to come at the cost of a rapport building conversational tone, but the interviewer must be well practiced to balance the two (Billings & Stoeckle, 1999;
Lichstein, 1990). Most texts on medical interviewing emphasize the use of open-ended questions and the importance of building rapport. Engel (1998) suggested that physicians must balance the patient’s need to be understood with the need to understand as much as possible about the presenting concern. In mental health settings, rapport often is the intervention, whereas in medical settings, rapport building enhances the quality of the information gathered during the interview and provides a foundation for trust that engages the patient in treatment.

In mental health settings, the patient may only interact with the mental health provider for care. However, care in medical setting is team-based. When multiple providers are involved in the treatment of a patient, effective communication is essential. Because nearly all physicians are involved in the treatment of a patient, effective communication is essential. Because nearly all physicians are trained using the same structured interview (and during medical school are trained to present the information gathered during the interview the same way), communication becomes standardized and is therefore simpler to communicate. The benefits go beyond verbal communication and even apply to documentation in the medical chart.

**CLINICAL INTERVIEWING IN PRIMARY CARE**

The clinical interview in integrated primary care settings is unique. It is the melding of the medical and mental health approaches to care. In order to meet the unique demands of providing behavioral health in primary care settings, the IBHP needs to enhance the strengths of the mental health clinical interview with the structure and efficiency of the medical interview. Typically, the time available to conduct an initial assessment is halved to 30 minutes in comparison to traditional mental health settings. Integrated care is team-based care with higher communication demands in addition to the time demands. In fact, in integrated care, the initial interview begins not with the IBHP, but with the primary care physician (PCP), both because the PCP has gathered considerable information throughout his or her relationship with the patient, and also because the PCP typically initiates the first contact between the patient and IBHP. As part of a treatment team, the IBHP has the benefit of the treatment team’s knowledge, but also the responsibility of collecting information that informs both his or her own treatment of the patient as well as the treatment provided by other members of the team. The IBHP must also work with the PCP and other team members to establish collaborative treatment goals.

**Goals of the Clinical Interview in Primary Care**

The goals of the clinical interview can vary and change based on the context in which the IBHP is brought into the patient’s care. Goals pertain to the patient, the IBHP, the clinic, and the PCP. For IBHPs new to integrated care, the PCP’s goals are often a new component. Assessment in integrated care includes assessing the needs and goals of the PCP, which may, at times, be relatively straightforward, (e.g., “I’m not sure if this patient is having true panic attacks, or heightened anxiety”) whereas at other times the IBHP may need to ask questions to clarify challenges the PCP is having and how to help. Interpersonal and communication challenges with “difficult patients,” patients diagnosed with personality disorders, or patients experiencing heightened emotions are common, less straightforward, and vague goals of the PCP. The PCP may not explicitly state, “This patient’s behavior angers me and I don’t know how to deal with him/her,” but the IBHP may help to clarify and/or intuit this need and then follow up with recommendations to the PCP on helpful communication strategies with certain patient behaviors.

Clinic goals may include addressing certain patient behaviors in the clinic that are problematic, such as repeated calls to the clinic to fill prescriptions after hours. Increasingly, payment structures are based on a clinic’s management of patient health as measured by process measures, such as screening for certain conditions and medication adherence measured via claims for prescription refills, and
outcome measures, such as blood pressure and blood glucose control. The IBHP is particularly adept at identifying barriers to behavior change and patient engagement. For example, a patient with diabetes may regularly forget to take his or her insulin, resulting in out-of-control blood glucose measures. The IBHP can identify practical barriers such as accessibility and adequate knowledge of blood glucose meters and insulin pens, or emotional barriers such as a sense of shame associated with out-of-control blood sugars that results in avoidance of self-care.

The IBHP’s goals are typically focused on gathering information to learn about the nature and severity of the problem or symptom, establish rapport, provide useful recommendations, and tailor the interview and intervention to the needs of the patient. Patients may be ambivalent about engaging in behavioral health interventions, or other health behaviors. An IBHP’s goal may shift to addressing the patient’s ambivalence toward behavior change or toward therapy itself using skills such as motivational interviewing (MI). MI is quickly becoming a foundational skill across disciplines in health care. Maintaining a healthy lifestyle and managing chronic health conditions undoubtedly involve behavior change, and people are often ambivalent about making certain changes in their life. MI is a skilled approach to conversations about behavior change and is rooted in a provider or healthcare team member’s ability to demonstrate empathy and patient-autonomy while working to evoke reasons for change from the patient’s perspective (Butler, Rollnick, & Miller, 2007).

**Referrals to Behavioral Health in Primary Care**

Numerous situations can prompt a PCP to engage an IBHP to assist in improving the health of primary care patients. If a patient’s problem has a stress and/or behavioral component an IBHP can add value to the patient’s health needs. The context in which IBHPs interact with patients and PCPs in primary care tends to be dependent on the structure or model of Primary Care Behavioral Health (PCBH) the clinic adopts. Over 25 years ago, innovators in the field of “integrated behavioral health” and “primary care psychology” began theorizing, practicing, and writing on the topic of changing our nation’s fragmented and siloed physical and mental health care. In the past decade, a number of healthcare providers, governmental agencies, insurance companies, and policy makers have acknowledged the impact the PCBH model has on improving several dimensions of healthcare delivery and health outcomes. Several models of PCBH exist, typically falling into the categories of integrated behavioral health, colocated behavioral health or coordinated behavioral health. An abundance of literature exists outlining optimal frameworks for the role and responsibilities of IBHPs in primary care (Collins, Levis Hewson, Munger, & Wade, 2010; Kwan & Nease, 2013; Strosahl, 1998). This chapter will consider the clinical interview as it applies to PCBH models that tend to have the largest array of opportunities for the IBHP to interact with patients, PCPs, and other primary care team members. This will allow an illustration of the complexities of interviewing patients in a fast-paced, collaborative primary care environment. Lastly, Robinson and Reiter’s (2007) Behavioral Health Consultant model will be highlighted, emphasizing the IBHP’s role as a consultant to the PCP in addition to meeting the needs of patients.

The broad range of referral questions in primary care leads to variable approaches to the clinical interview. There are two routine contexts in which an IBHP meets with a patient. The first is the *initial introduction* (also known as the “warm handoff”). The initial introduction follows the PCP’s identification of a behavioral health need. The PCP communicates to the patient that in order to comprehensively address their health, their primary care team would like to recruit another team member who specializes in one or more of the patient’s presenting concerns. The initial introduction may be spontaneous or preplanned, perhaps in the morning during the primary care team huddle or via email the previous day. In some settings, patient encounters become joint appointment or “dual visits” with both the PCP and IBHP in the room, assessing and intervening collaboratively.
The initial introduction may involve a brief meet and greet with a subsequent follow-up visit or may lead to a same-day initial visit or a scheduled appointment within the next one to two weeks. This depends on the severity of the presenting problem (e.g., risk assessment), the IBHP’s schedule, the consult question from the PCP (i.e., does the IBHP need to assess and make a recommendation back to the PCP same day to inform treatment), as well as the preferences and needs of the patient. A key, yet challenging skill for the IBHP is to maintain flexibility in the approach to the initial interview so that it complements the needs of the patient and PCP, in addition to gathering information that is clinically indicated for the IBHP to make appropriate recommendations. For example, an on-the-spot consult may be called for when the IBHP is asked to assess for a history of manic episodes prior to the PCP starting the patient on an SSRI. In situations when a treatment relationship is being established, such as in the case of treatment for newly developed panic disorder, a scheduled initial session in the IBHP’s clinic may be appropriate. However, even in the case of a consultation that is intended for assessment and brief treatment, an initial introduction, even as short as five minutes, is recommended. Research has shown that when patients are introduced to the IBHP at the time the PCP makes a recommendation to engage in behavioral health care, they are more likely to show up to subsequent follow-up visits with the IBHP (Bartels et al., 2004; Guck, Guck, Brack, & Frey, 2007; Pomerantz, Cole, Watts, & Weeks, 2008).

**Reasons for IBHP Engagement in Patient Care**

The following is a brief discussion of common situations in primary care settings that trigger a referral from the PCP to the IBHP. While this topic could certainly take on a chapter of its own, subsequent content focuses on a broad review of common types of referrals that have unique considerations for the clinical interview.

Several pathways exist that help guide the direction and purpose of the clinical interview. More than one pathway may be emphasized during the visit. The first pathway involves gathering data in order to inform the decision-making and treatment planning of the PCP. This includes interviews to help inform psychotropic medication initiation or selection, to clarify diagnoses such as attention deficit hyperactivity disorder, or to conduct a behavioral health assessment prior to initiating a patient on opioid therapy. The second pathway involves situations in which emotional health symptoms such as anxiety and depression are the reason for referral. In these instances, the IBHP must gather information regarding functioning and symptom severity in order to determine appropriateness for behavioral health intervention in primary care versus referral to outpatient mental health treatment, social service supports in the community, care management, or higher levels of care (e.g., inpatient and partial hospitalization programs). The third pathway involves gathering data to provide recommendations for patients and/or family members. An example of this includes interviewing along with brief screening for cognitive impairment in older adults and locating support services for patients with family members who suffer from addiction or severe and persistent mental illness. The fourth pathway involves gathering data to inform the patient’s behavioral goals for continuing behavioral health intervention in primary care. The fifth and final pathway pertains to an awareness of when gathering data needs to cease, with the IBHP transitioning to reflective listening, crisis interventions, and/or motivational interviewing.

Having an awareness of the pathway and direction of the interview helps the IBHP set initial goals for the encounter, manage time appropriately, and triage the most important questions that pertain to the goals of the encounter. As emphasized in the fifth pathway, it is equally important to remain flexible and shift course based on the patient’s needs and the IBHP’s clinical judgment.

**Risk Assessment and Crisis Interventions**

Risk assessments and crisis interventions are common in primary care. Primary care is often the first stop for patients experiencing a mental health crisis or other crisis involving threats to safety and
basic needs such as housing, food, finances, and domestic violence. Working knowledge of community resources and state programs available to assist PCPs and patients in accessing resources is imperative for the IBHP in primary care.

The IBHP has access to multiple sources of information in the primary care setting that assist clinical judgment during risk assessments, which are rarely as immediately accessible in traditional mental health settings. The IBHP might be meeting the patient for the first time when there is a concern for risk of harm to self or others, but they have invaluable access to collateral information from the shared medical record and PCP. Oftentimes, nurses and medical assistants have a wealth of knowledge about a patient’s baseline and history.

**Mental Health Symptoms and Disorders**

Mental health concerns are often the most common trigger for engaging the IBHP. Primary care patients present with mental health concerns on a continuum. This can range from mental health symptoms resulting in changes in a person’s emotions and behavior to mental health symptoms that are impairing functioning and meet criteria for a mental health disorder. Assessing where the patient falls on this continuum often becomes the first objective. On the less severe side of the continuum, patients present with decreased coping resources and confidence to manage daily stress that may or may not be leading to or exacerbating physical health symptoms (e.g., stress affecting hypertension, gastrointestinal issues, headaches). Some patients present with increased stress and emotional symptoms secondary to recent major life stress or changes, with minimal to no effect on daily functioning (i.e., nonpathological reactions to stress/change). Brief interventions with these patients may prevent more severe consequences of stress such as chronic health problems and/or the onset of a mental health disorder.

Then there are patients who present with symptoms meeting criteria for an active mental health disorder with mild to moderate symptoms, and those patients with active diagnoses with severe and worsening symptoms and notable functional impairment. Approaching the clinical interview for addressing mental health symptoms in primary care often starts with differentiating clinically severe symptoms caused by a mental health disorder from common reactions to stress, change, or loss. The IBHP also assists the PCP by clarifying diagnosis and identifying treatment plans that are both evidence-based and patient-centered.

**Substance Use Disorders**

Primary care providers have important roles to play in screening, early intervention, and referral to specialty care for patients with substance use disorders. Determining the severity of a patient’s substance use can be a challenging task that many PCPs are not formally trained to evaluate. IBHPs assist in determining severity of use, risk for harm due to substance use, and to recommend appropriate levels of care to the patient and PCP. The IBHP’s skills and expertise in assessing the patient’s readiness to change and motivation to seek treatment are crucial in these encounters. IBHPs make recommendations to the PCP and other team members on how to work more effectively with patients who are engaging in problematic substance use but who are not ready to change or engage in treatment. Assessing the patient’s receptiveness to discussions of harm-reduction strategies often becomes the priority in these situations.

**Chronic Disease Management and Health Behavior Change**

Research has identified several evidence-based behavioral interventions for chronic conditions such as diabetes, hypertension, headaches, obesity, irritable bowel syndrome, chronic pain, and cardiovascular disease. Referrals for behavioral health assessment and intervention may be incorporated as part of a primary care clinic’s treatment protocol for certain conditions. An example of this would be screening, assessment, and intervention for patients newly diagnosed with type II diabetes. Health
behavior change referrals also include issues such as adherence to treatment for obstructive sleep apnea, smoking cessation, strategies to increase blood sugar monitoring, problems with taking or remembering to take medications, or identifying lifestyle changes for a patient newly diagnosed with high cholesterol who wants to avoid medication. In these situations, interviews focus on identifying underlying mechanisms reinforcing certain behaviors or avoidance of behaviors, along with assessing the patient’s stage of change in order to inform goal setting.

**Patient Engagement**

Patient engagement pertains to a patient’s active role in managing their emotional and physical health and hinges on knowledge, understanding, skill acquisition, motivation, and confidence as it relates to managing medical conditions and overall health. It is common that referrals to behavioral health involve aspects of patient engagement. Low patient engagement is not just simply low motivation. Patient engagement is a multidimensional concept that includes motivation, but also expands to other possible factors that may be impeding the patient’s engagement. Examples include knowledge and understanding affected by health literacy, the actual skill and ability to engage in health behavior (e.g., a patient who was never shown how to check her blood sugar) as well as success in other areas of health that foster self-confidence and self-efficacy. While enhancing patient engagement is the responsibility of several primary care team members, IBHPs often have specialized training in the behavioral and emotional complexities of managing and improving one’s health. The identification of factors that may impede patient engagement is a valuable skill during the clinical interview.

**CONSIDERATIONS FOR THE CLINICAL INTERVIEW IN PRIMARY CARE**

The nuances of providing behavioral health assessment and intervention in primary care settings can be observed across several points of interaction between the IBHP, patient, and PCP. This section reviews specific domains of the clinical interview that tend to deviate from traditional approaches. Strategies are offered to help IBHPs adapt their approaches to the clinical interview in order to complement the distinctive demands of primary care. Topics covered in this section include modifications to the informed consent process, such as timing and breadth of limits of confidentiality and use of shared medical records. Approaches to optimizing use of data sources such as screeners, collateral information from the primary care team, and the electronic medical record will be provided. The balancing act of establishing rapport while eliciting the most critical and useful information in a time-limited encounter can certainly be a challenge. Important considerations and practical suggestions for navigating this challenge are reviewed.

**Informed Consent**

Aspects of the informed consent process are unique to primary care. A more thorough explanation of informed consent and limits of confidentiality often depends on the type of interaction (e.g., a same day 10–20 minute interaction versus a 30-minute scheduled follow-up) and the presenting problem of the patient. In traditional mental health settings, it is common for this to be reviewed prior to the clinical interview. However, in primary care, particularly in the case of an initial introduction, exhaustive informed consent beyond an explanation of the role of the IBHP might not be completed until it has been established that the patient will return for behavioral health treatment with the IBHP.
Information about treatment in primary care and the potential for referral to the community are important details to consider incorporating into the informed consent. During the initial visit when the severity of mental health symptoms are being triaged for possible referral to a traditional mental health setting, it is important to alert the patient at the onset that based on information from the assessment, treatment in the community may be recommended. Comparing this to the PCP’s role can be a helpful explanation, such that when patients are seen in primary care for a medical condition that requires specialty care, PCPs make referrals to specialists in the community. In other situations, the IBHP might begin treatment with a patient, but with more clinical data gathered over time and/or limited symptom improvement from treatment provided in primary care, the IBHP may recommend more frequent treatment in outpatient mental health setting. Clarifying this at the onset can prevent subsequent confusion and upset for the patient.

Informed consent may also include a description of the unique delivery of behavioral health care in primary care, such as the typical 30-minute, biweekly follow-up visits with a shorter-duration, solution-focused approach. This explanation tends to be more prudent if the patient has a history of mental health care in the outpatient setting, as they may expect the weekly 50–60 minutes they were accustomed to in that setting. While the 30-minute, biweekly structure allows for greater access to behavioral health care for the patient population served by the primary care clinic, there are cases when such time-limited therapy isn’t possible given case-specific nuances. Again, flexibility and adaptability are paramount to providing patient-centered care. Consider a patient seen in the clinic and diagnosed with acute stress disorder secondary to a car accident two weeks prior, who also has a history of a PTSD diagnosis 10 years prior. She hasn’t been able to leave her house in the past two weeks, aside from making it to her PCP visits. This patient will likely require a referral to a behavioral health provider in the community who specializes in the treatment of PTSD. As is oftentimes the case, the IBHP learns that the waitlist for a trauma/PTSD specialist is 6 weeks. It is reasonable in this situation that the IBHP schedules slightly longer, weekly visits in order to provide acute bridge care while the patient is on the waitlist.

Another consideration is to alert the patient to possible interruptions for warm handoffs when it has been determined that the patient will begin treatment with IBHP. Behavioral health providers new to the integrated care setting commonly express concerns about how interruptions and stepping out of patient visits will adversely impact the patient. This is very easily remedied with communication at the onset. It is recommended that IBHPs circle back to the time they were initially introduced to the patient. The IBHP can state, “Remember when I first met you after Dr. Smith introduced me and we spoke about ways I work with patients to help cope with uncontrollable worry? I make myself available to the PCPs at all times so that they can introduce me to a patient who might benefit from behavioral health, much like I did with you, because we know it increases the chances of someone coming back to meet with me.” Patients connect this with their own experience and appreciation of being able to meet the IBHP.

Clarifying limits to confidentiality with regards to ongoing consultation and communication between the IBHP and PCP can be efficiently covered in the IBHP’s initial introduction.

Example: “I’m a behavioral health provider who works specifically with PCPs and patients in primary care so that we can better care for your physical and emotional health. What that means is it’s an open line of communication between myself and your PCP and we work as a team to help you meet your goals.”

It’s also important to inform the patient of the IBHP’s use of a “shared medical record” that allows the IBHP access to the patient’s medical record and the availability of the IBHP’s progress notes to
any medical provider involved in the patient’s care within the clinical system or to providers who request access to the patient medical record. This may be limited to a small number of providers in an independent, private primary care practice to an entire system of providers in an academic medical center setting.

The timing of communicating limits of confidentiality related to risk to self or others and mandated reporting for abuse or neglect of children, elderly, or dependents is important to consider. For IBHPs transitioning to primary care from traditional mental health settings, the reflex to state this to all patients immediately can be challenging to change. PCPs have the same level of responsibility to protect patient safety and are mandated reporters, and a statement that the IBHP is bound by the same confidentiality and reporting laws as the patient’s PCP may be more appropriate depending on the situation. The key here is to be flexible. Something to consider in the primary care setting are the nuances of communicating this to people with no history of behavioral health treatment and who may be referred for nonpsychiatric reasons such as behavioral sleep management. Stating limits of confidentiality and mandated reporting right away upon meeting a new patient may make certain patients uncomfortable and confused. Once it is established that the patient is appropriate to continue treatment in primary care and plans to schedule a follow-up, details regarding informed consent and limits of confidentiality can be emphasized at the closing of the first or beginning of the follow-up appointment.

After hearing contents of the informed consent process (particularly shared medical records and ongoing communication with the PCP), a small number of patients may elect not to receive behavioral health care in an integrated primary care setting. This occurrence reinforces the importance of these discussions with patients. We recommend periodically revisiting these discussions and ensuring clear documentation of the patient’s ongoing consent to the integrated care model and use of shared medical records. This can also be done indirectly by having the patient’s medical record visible during the clinical interview as well as opening up previous behavioral health progress notes for the patient and IBHP to review together to help recall goals/plans from a previous encounter. It is much more common that patients express preference and relief that their providers communicate in between visits, as patients know firsthand the errors, misunderstandings, and headaches caused by inadequate information sharing across providers and clinical systems.

### Time Constraints

In traditional mental health settings, it is common that initial interviews are scheduled between 60 and 120 minutes. This allows ample time to cover several specific domains across a person’s lifespan, which may inform conceptualization, diagnosis, and treatment planning and provide insight into the historical origins of problematic emotional and behavioral patterns. In primary care, the IBHP may have 30 minutes or less during the initial patient encounter. One of the challenging shifts for IBHPs trained in the traditional 60- to 120-minute initial interview model is “letting go” of the impulse to gather a certain quantity of psychosocial data. Triaging what information is necessary based on the presenting concern and goals of the PCP, patient, and IBHP is a skill that is learned and honed over time. The other critical perspective here is flexibility, in that important data may be collected over subsequent follow-up visits with a patient.

The following are time-efficient strategies for gathering data:

1. Optimize use of screeners;
2. Circle back to the PCP;
3. Medical record;
4. Use targeted, open-ended questions that elicit pertinent details from the past (e.g., “Are there any significant events from your childhood that still bother you?”);
5. Focus on what matters most to the patient (e.g., “What are the top three things in your life that are creating stress that you want to change?”);
6. Provide instructional and educational handouts.

**Use of Standardized Instruments in PCBH**

Because time is more limited in integrated care settings, techniques to collect data outside the actual face-to-face interview are especially important. Specialists of all types, including mental health clinicians, may use standardized forms to collect information about patients. In integrated settings, pre-appointment forms and screening instruments can be invaluable tools both for the PCP and for the IBHP.

**The Role of Screening in Referral**

For PCPs, clinic screening protocols serve to promote population health management and alert clinicians that a patient might be experiencing behavioral health difficulties. Most clinics use a standardized screening interval, not unlike those used for cancer and chronic illness screening tests. Because many instruments have suggested cutoff score ranges, each clinic must decide the cutoff they use to trigger further intervention. In fact, in the authors’ clinics, these mutually agreed upon cutoff scores are printed in small font underneath the instrument for PCPs to reference.

Typically patients are screened with the instruments of the clinic’s choice on a yearly basis. In addition, screening instruments can be made available in exam rooms for as-needed situations. For example, if a patient is reporting feelings of depression or anxiety, a screening instrument can be administered as a way to assess risk for mental health difficulties. In either case, the positive screening results first trigger a discussion with the patient to ascertain the impact of the reported symptoms on the patient’s day-to-day life and to rule out alternative diagnoses. Then the PCP can discuss treatment options with the patient, which will typically include an offer of some form of interaction with the IBHP (warm handoff, telephone outreach, scheduled appointment).

**Screening as Preinterview Data Gathering**

For the IBHP, these same instruments serve to prepare them for conducting the clinical interview prior to the initial meeting for the likely focus of the interview. When these instruments are readministered upon the initial IBHP meeting, it can suggest the clinical course of the problem and assist with triaging the patient. For example, if a patient scored in the moderate clinical range for the first time on a regularly administered screen but then scored in the mild to subclinical range at a later date for an initial appointment with the IBHP, the clinician can feel more confident that the problem is either resolving independently or that any intervention enacted by the PCP is both effective and possibly sufficient depending on the severity of the second score. Conversely, if the second score was worse than that of the PCP visit, it should raise the IBHP’s sense of urgency as the symptoms may continue to progressively worsen.

Finally, and perhaps most usefully, given time constraints, it is common to forget to ask about a particular symptom during the initial interview. Screening instruments often serve as a backup measure in these instances, allowing the IBHP to comprehensively document the diagnosis. This also applies to certain demographic and historical questions such as living situation, religious tradition, mental health treatment history, or family history. In addition, having items already answered can save time and allow the IBHP to focus on pertinent positives in the patient’s demographics and
history. Screening instruments can also be used to assess safety issues, which, depending on the presenting symptoms, may seem out of place in the clinical interaction with the patient. For example, a patient who presents with ADHD symptoms may be surprised to be asked about suicidal ideation and such a question could potentially damage rapport by making the patient feel unheard (Kaplan et al., 1994). On the other hand, standardized paperwork is so common in primary care that answering these questions on a pre-appointment form that was provided to the patient by nursing or check-in staff may be less off-putting.

Information Resources Available in Integrated Primary Care

Despite the challenges of practicing in a primary care setting, strong advantages exist as well and if used effectively, these advantages can mitigate the impact of such challenges as limited time. For most mental health clinicians, the only source of information they have about the patient is what the patient says in session during that first interview. The IBHP has multiple sources of information that can span across several years and multiple domains.

In most cases, a significant amount of patient history has already been gathered during previous medical encounters and is documented in the chart. Utilizing the already gathered information can be another valuable time-saving tool. It is important to confirm all pertinent information, particularly information that may change over time. However, given the shorter interaction, certain information can be filled in from the chart if the clinician doesn’t have the opportunity to discuss it at length. For example, the medical history can simply be updated by asking “Has anything changed about your health since you last saw your PCP? Are there any health conditions that you have that you feel are relevant to your current concern?”

The Primary Care Provider

The treatment team, including the PCP, nurse, medical assistant, and even office staff are wonderful resources for the IBHP. Prior to or during a warm handoff, the IBHP is typically provided with a brief history by the PCP. The PCP may even know the patient’s family and environmental situation in addition to pertinent symptom history. The ability to learn a patient’s history from the PCP is one of the most valuable tools for effective patient care. For example, in one family medicine clinic, a mother and her children were all patients of the same PCP. That PCP had also treated this woman’s other family members, including the woman’s mother, in the past. The mother was struggling with her children. She had faced much criticism for her actions and she was at risk for having her children removed from the home for neglect. Prior to meeting with the patient, the PCP explained to the IBHP that the mother received almost no parenting as a child herself and so it was made clear that the mother wasn’t uncaring, but rather, she didn’t know how to help her children with their emotional difficulties and her best efforts had sadly served to exacerbate the problem. Without this knowledge, the IBHP would have likely concluded as most others involved in the case had: that the mom just couldn’t be bothered to do the difficult task of consistently managing her children’s behavior. Instead, the support that was provided to the mother established a trusting relationship that allowed the IBHP to make suggestions regarding how to better help the children, and within six months the case with the Department of Children and Families was closed.

Clinic Personnel

Other team members serve to round out the IBHP’s initial conceptualization of the patient. Front desk staff can alert the provider if the patient is particularly upset on the day of the appointment, or they may know that the patient is struggling financially or that he or she just isn’t acting like his or her
normal self. Medical assistants will know if the patient has recently gained or lost weight, has elevated blood pressure, or is due for preventative care. And nursing staff may have spoken to the patient on the phone since their PCP appointment and have an update for the IBHP.

The Medical Record

The patient’s medical record is the second most valuable source of information for behavioral health providers working in primary care settings. In most traditional mental health treatment settings the provider is lucky if they receive the patient’s records within a few weeks of the initial visit. IBHPs have the medical record available for every patient on their schedule as soon as the appointment is scheduled. Basic information from the patient’s problem list can provide clues to both the patient’s medical and emotional health. Diagnoses like esophageal and gastrointestinal reflux (Jansson et al., 2007), migraine (Juang, Wang, Fuh, Lu, & Su, 2000), irritable bowel disease (Lydiard & Falsetti, 1999), and fibromyalgia (Weir et al., 2006) are often associated with chronic stress, anxiety, and posttraumatic stress disorder. The problem list may include a diagnosis of a highly stigmatizing illness such as current or past substance use disorder, which the patient may or may not have felt secure enough to reveal to the IBHP during the initial clinical interview. The problem list also may include diagnoses of chronic pain disorders that are amenable to behavioral interventions and if left untreated, could exacerbate the stress response or interfere with the patient’s ability to engage in exercise or other coping strategies. In addition, the patient’s current and past medication list can both indicate the patient’s mental health history and the course of previous illness. Lab work available in the medical record can eliminate the concern that depressed mood might be caused by hypothyroid.

Finally, PCP and other provider notes can provide a wealth of history. Taking the time to review a patient’s history when the concern is opiate pain medicine dependency can reveal patterns the PCP may not have noticed over a period of several months, but which becomes clear upon review. Issues such as a pattern of early refills, slowly escalating dosage, and initial effectiveness followed by reports that the medication “does nothing” can all lead to identification of an escalating substance use problem.

Gathering information from the sources listed earlier can save considerable time. Pertinent information can be quickly confirmed with the patient. Some providers have expressed concern that patients might be unsettled by this type of knowledge; however, most patients are comfortable with the IBHP knowing information about them prior to the clinical interview and, in fact, many fully expect it (MacKinnon, Michaels, & Buckley, 2015).

Collecting all of the information also serves a purpose that is fairly unique to integrated settings. Most patients who present for an initial assessment with a mental health provider have done so because they were motivated to seek treatment for either a particular personal concern or because they have been mandated for legal reasons. However, in integrated care, meeting with an IBHP is often suggested to the patient by the PCP and an appointment is made at the time of checkout. The patient may not understand why his or her PCP suggested an appointment with the IBHP; at times, the patient may not even fully understand the role of the IBHP. The following exchange is not unusual in integrated settings, particularly early on in the process of integration, when the PCPs are still learning how to present integrated care to patients.

IBHP: “What brings you in today?”
Patient: “I don’t know, my doctor told me to.”
IBHP: “Do you remember what you were talking about when your doctor suggested that scheduling an appointment with me might be a good idea?”
Patient: “Not really.”
IBHP: “Well, looking at your doctor’s last note, it looks like you were discussing ____ and she/he thought I might be able to help you with _____. Is _____ something that has been an issue for you lately?”

Patient: “I don’t know, not really.”

IBHP: “Often, the doctors here ask me to meet with patients when they’re dealing with a lot of stress. I also help patients if they’re trying to make a change in their life like losing weight or quitting smoking. I think your doctor was a little concerned that you have so much going on right now and thought I might be able to help.”

Patient: “Well, I guess I have been ______. . . .”

Balancing Rapport With Gathering the Necessary Information

Establishing rapport with clients is an essential part of the initial clinical interview. However, given the limited time, there is a real tension between maintaining an open and conversational tone and gathering the information necessary for an informed differential diagnosis and the necessary documentation. As mentioned earlier, the use of standardized questionnaires can be particularly helpful for this. Use of the additional sources of contextual and historical information via consultation with the PCP and review of the electronic medical record is another valuable strategy. Borrowing a more structured format from medical interviewing facilitates gathering the essential information and aids in effective documentation such that the information is readily accessible to PCPs and other members of the interdisciplinary care team.

When having the first interview with the patient, the IBHP has to decide whether to conduct a full intake assessment, an assessment focused only on the presenting concern, or simply a rapport building interview where the patient takes the lead. During the full intake assessment, the interviewer gathers a complete patient mental health history, in addition to fully exploring the patient’s presenting concern. The goal is to collect enough information to identify any mental health diagnosis that could be present, even if the symptoms aren’t currently of primary concern to the patient. Conducting a thorough assessment in this way gathers information that may be missing from the medical record. By gathering a complete history, the interviewer can be more confident that he or she hasn’t missed any critical elements such as a family history of bipolar disorder or a personal history of suicide attempt or trauma. This information could be pertinent to the clinician’s triage, assessment, and plan, or potentially for the PCP’s treatment choices in the future. It is reassuring to eliminate any surprises down the road. The trade-off, of course, is the time required to conduct such an assessment. In most integrated settings, a complete history isn’t feasible apart from a scheduled appointment for which sufficient time has been set aside. Additionally, the extensive questioning and intimate nature of the questions asked might be off-putting and damage rapport with a patient who came in for a PCP appointment not expecting a full mental health assessment.

The presenting concern-focused assessment is similar to the history of the present illness for an acute medical visit. The focus of the interview is the patient’s reason for the visit and most information discussed will be pertinent to the diagnosis and treatment of the presenting concern. The goal is not to unearth any heretofore unknown mental health problem. For the patient, the focus on the presenting concern is fitting and they feel heard. This type of encounter is more compatible with warm hand-offs as it is typically more brief and targeted. Additionally, it is this type of interview that lends itself to immediate intervention as problem-solving becomes a focus. It wouldn’t be at all unusual for this type of interview to include solution-focused or motivational interviewing techniques. The patient will leave the interaction not just with a treatment plan that describes forthcoming interventions,
but also with ideas for coping strategies or a behavior plan. While this isn’t excluded from the intake assessment, the intervention is prioritized over data gathering.

ELEMENTS OF THE CLINICAL INTERVIEW

A typical clinical interview in integrated care is built upon the core values of the traditional mental health interview; however, the principles of the mental health interview can be formed to the structure of the medical interview to enhance efficiency. The clinical interview essentially begins the moment the PCP engages the IBHP, as medical providers often provide succinct and detailed summaries of their patient’s history, presenting concerns, and the consultation question. The excerpt below provides an example.

PCP: “Mrs. Smith is a 45 year-old who has a history of major depressive disorder and panic disorder. Her husband suffered a heart attack and died 4 months ago. She is having frequent abdominal pain that is keeping her from leaving her house. We’ve worked her up for just about everything, and we can’t find an organic cause. She was in counseling in the past, following a suicide attempt in her early 20’s, but stopped after a few sessions saying ‘it wasn’t for me.’ She doesn’t want to increase her antidepressant at this time. She says she is hesitant, but willing to give counseling another try.”

Even in this brief exchange with the PCP, the IBHP already has a working formulation that will help guide a time-efficient clinical interview.

The Opening

The opening of the interview is similar across mental health and medical settings. It is the opportunity to conduct greetings and introductions. Most importantly, this is the opportunity for the provider and patient to develop first impressions. In some cases, this process can be an uphill battle for the provider, who must gain sufficient trust from the patient so that the patient feels comfortable disclosing private and often emotionally distressing information to a person he or she just met. In integrated settings, the IBHP has a distinct advantage over most other providers. In many cases, the IBHP has had an opportunity to briefly meet the patient during a warm handoff, but most importantly, the IBHP has the clear endorsement of the PCP. In many ways, the trust and rapport that the patient has developed over years with his or her PCP can be transferred to the IBHP. Several patients who have repeatedly refused the PCP’s suggestion that he or she seek therapy due to distrust of the mental health system become willing to at least meet the IBHP during a medical appointment. Typically, this initial meeting lays powerful ground work for a successful initial clinical interview. Because many patients first meet the IBHP primed for a good first impression, the IBHP does not need to devote considerable time to this “getting to know you” phase.

Orienting the Patient to Integrated Behavioral Health

In the initial interaction, the IBHP typically needs to communicate his or her role in the clinic, emphasizing a partnership with both the PCP and the patient. Best practice includes reiterating the specific questions and concerns (reported by the PCP and/or patient) that prompted consideration for BH services, followed by an explanation of how the IBHP helps patients with certain types of problems. Imagine a PCP introduces the IBHP to a 55-year-old man who reports no history of behavioral
health treatment and does not understand how an IBHP could help him with his immobility issues and hypertension. Engaging the patient and establishing rapport, reducing potential discomfort, and providing information about the role of behavioral health all serve to increase the chances of him returning and may become more important than focusing on gathering information. For example,

My name is _______, I’m a behavioral health provider who works specifically with physicians and patients in primary care. I’m often called in when patients are experiencing any challenges that are related to stress and behavior. I understand that you’ve talked to Dr. Smith about difficulty sleeping over the past month, which may be related to some new stressors in your life. Would it be alright if I asked you some more questions to better understand your situation and to see how we might help?

Patients who have never seen a therapist before may not know what to expect from an appointment with an IBHP. Conversely, patients who have received treatment from providers in traditional mental health setting in the past are likely to have a schema for a “psychotherapy appointment” based on those experiences. For example, they may expect 50-minute weekly appointments or a less directive, more supportive therapy approach. In both the former and the latter cases, the patient will need to have integrated care explained to them. Although one may strive to educate PCPs in the most effective ways to foster accurate expectations of integrated care when they introduce it as an option to the patient, even the best PCPs have times when there simply isn't enough time to provide a full explanation and/or the patient may not recall the conversation. Thus, the IBHP should always plan to have a discussion with the aim of establishing accurate expectations of primary care. A typical conversation might go like this:

IBHP: What did your PCP tell you about what my role is here?
Patient: She said that you’re a therapist and can help me with my anxiety.
IBHP: That’s true. Have you ever seen a therapist for help with your anxiety or any other life stress in the past?
Patient: Well, not for my anxiety, but I did see a therapist with my husband when our marriage was going through a tough spot.
IBHP: What was that experience like for you? What types of things did you talk about?
Patient: Well, we mostly just talked about what happened between us (my husband and I) over the past week. Then the therapist would help us talk to each other about those issues. He also had us do things like start a date night.
IBHP: Were they helpful to you?
Patient: Yeah, we’re doing much better now.
IBHP: What do you think was the most useful part of that for you?
Patient: He helped me understand my husband’s perspective better. And we learned to invest time in each other.
IBHP: Did your PCP tell you anything about how therapy works here?
Patient: Not really.
IBHP: Well, in many ways it will be similar to what you experienced in couples therapy. We’ll talk about your struggles, maybe try to find some solutions and help you get a new perspective on your struggles. There are a few differences though. The most obvious difference is the schedule here. Most therapists will see you for 50 minutes or more. Appointments here usually last 30 minutes, though often, if my next patient isn’t here yet, we can extend the length of your appointment if we seem to be making good progress. So if you need to leave at a certain time, please tell me so I can plan for that. Another big difference is that many
therapists see their patients weekly, especially at first. Weekly visits often aren’t possible here because I see the patients of all of the PCPs who work here and in order to try to fit everyone in, I’m usually only able to see patients every other week and I’m often not able to give my patients the same time slot for visits. Finally, because most therapists work in a different office from your PCP, there’s limited information shared between the therapist and PCP. Here, my office is right next to your PCP’s and we discuss all of our shared patients to make sure your care is coordinated. That way, if I think you might need something checked or changed medically, I can talk to your PCP about it.

Body or Presenting Concern and Review of Systems

IBHP may find that a more focused assessment of the presenting concern is most effective. Patients have higher tolerance for increased structure because the purpose of that structure (i.e., why the provider is asking these particular questions) is clearly tied to the problem for which they are seeking help. The Cardinal 7 (discussed earlier) can translate fairly effectively. Timing of the presenting concern is nearly identical to a medical interview. The quality of the symptoms can be assessed by asking the patient to describe how they’ve been feeling and what they’ve been thinking. A functional assessment can address severity by asking “How had this affected your day-to-day life?” Context includes stressors and social support. Modifying factors are coping mechanisms and triggers. Associated signs and symptoms might include behavioral or health challenges. Finally, the review of symptoms would include a mental status exam and might include common symptoms that the patient hasn’t mentioned and/or symptoms of common comorbidities.

Past Medical, Psychiatric, and Family History

These sections are essential, but uncomplicated. Typically the past medical history focuses on relevant medical diagnoses, including sleep difficulties, chronic pain, chronic illnesses included in the patient’s and/or PCP’s concern, and stress-related illnesses. In many cases, providing education about the role of stress in chronic pain, gastrointestinal distress, and autoimmune disorders helps motivate patients who fail to see the connection between high stress levels and their health. The past psychiatric history should include previous psychopharmacological treatments, previous experiences with psychotherapy, and any history of psychosis, suicidality, or hospitalizations. Finally, family psychiatric history may become partly evident during a discussion of stressors. It is also important to assess for other known or suspected family diagnoses, including family history of substance abuse and any effective treatments for relatives’ mental health (if known).

Social History

The social history is perhaps the most similar across medical and mental health practice, particularly in integrated care. Many IBHPs will have a tendency to collect a detailed history dating back to childhood as they may have learned during their training. Commonly, extensive information about the patient’s current and family life will be brought into the discussion as it is relevant. Gathering the information that is necessary for documentation and billing is generally sufficient to elicit the social history needed for effective diagnosis and treatment, particularly if the interviewer is alert to emotional cues such as body language and tone of voice. For example, a patient who has a strained relationship with his or her parents will generally offer that information (or provide the information in such a way as to
suggest that further inquiry might be warranted) when asked about family medical history or simply “Do you have any siblings?” Simply raising the topic of family typically brings relevant family information to light without the need to formally ask for a description of the patient’s childhood. Therefore, the social history most closely resembles that of the medical interview. It typically includes a brief family genogram (or similar documentation of family names and relationships), a statement about the patient’s highest level of education, religious tradition, current occupation, and so forth.

When engaging patients in a more focused interview, only questions pertinent to the chief concern are typically asked. Certain social history questions may not be as relevant. For example, it is often important to learn with whom the patient lives to ascertain available support or possible ancillary stressors. However, there may be times when knowing the city in which the patient grew up or the patient’s birth order isn’t essential to addressing the problem at hand. For example, this information may not be critical to assist a patient with insomnia or weight loss. In fact, there have been cases in which information such as this was not essential for treatment of recent onset depression or anxiety. Good clinical judgment should guide the interviewer in these cases.

In order to organize the information efficiently, reduce notetaking, and facilitate documentation in the chart, the authors have developed a form that is intended for use as a hard-copy interview guide (see Figure 8.1). In order to serve as a mnemonic tool, the relevant symptoms from the most common diagnoses are listed on the form and may be circled and/or crossed out as the patient endorses or denies them. Note that the interview guide also includes the patient history, family history, and social history necessary for documentation and billing.

<table>
<thead>
<tr>
<th><strong>Name:</strong></th>
<th><strong>Date:</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Why referred?</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Current Symptoms:</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Sleep:</strong> Hyper/ initial/ intermed/ terminal</td>
<td></td>
</tr>
<tr>
<td>Duration:</td>
<td></td>
</tr>
<tr>
<td>Onset:</td>
<td></td>
</tr>
<tr>
<td>Hygiene?</td>
<td></td>
</tr>
<tr>
<td>Nightmares?</td>
<td></td>
</tr>
<tr>
<td><strong>Depression:</strong> Mood/ anhedonia/ worthlessness/ irritability/ anger/ tearfulness/ guilt/ energy/ concentration/ indecisiveness/ psychomotor change/ hopelessness/ self-harm/ hyperphagia/ hypophagia</td>
<td></td>
</tr>
<tr>
<td>SI?</td>
<td></td>
</tr>
<tr>
<td>Duration:</td>
<td></td>
</tr>
<tr>
<td><strong>Mania:</strong> Grandiose/ talkative/ pressured speech/ racing thoughts/ distractible/ agitation/ increased goal-directed activity/ sleep/ pleasurable activities with potentially painful consequences</td>
<td></td>
</tr>
<tr>
<td>Duration:</td>
<td></td>
</tr>
<tr>
<td><strong>Anxiety:</strong></td>
<td></td>
</tr>
<tr>
<td>Rumination/ muscle tension/ can’t relax/ irritability/ easily fatigued/ mind going blank/ worry—About?</td>
<td></td>
</tr>
<tr>
<td>Avoidance?</td>
<td></td>
</tr>
</tbody>
</table>

⇒ Figure 8.1 Standardized Intake Guide
**Panic:** SOB/ Lt headed/ palpitations/ chest pain/ shaking/ choking/ nausea/ sweaty/ chills/ numbness/ tingling/ derealization/ fear losing control/ fear of dying

Duration?  
Freq?  
Onset:  
Triggers:  
Coping?

**Trauma History**

Child—emot /phys /sexual /neglect  
Domestic abuse/ combat/ assault/ robbery/ natural disaster/ MVA  
**SX:** Re-experiencing (memories, dreams, flashbacks, phys/emot distress w/cues)/ Avoidance (cues, memories)/ Cog & mood (inability to recall, Neg beliefs, Distorted pineiv emot, pessimism Anhedonia, Detachment/numbingpineiv emot)/ Arousal (Irritability, Recklessness, Hyper-vigilance, Startle, Concentrate, Sleep) numbing

**Psychiatric History?**

Therapy:  
RX:  
DX?  
Hospitalization?  
Psychosis?  
SI/ Attempt:  
Fam Hx: Depression/ Bipolar/ Anxiety/ ADHD/ Substance abuse/ Psychosis/ Autism/ Other

**Substance use/addictive behavior?**

Caffeine:  
Alcohol:  
Tobacco: pk/day Since?  
Illicit drugs:  
Quit attempts  
Substance abuse TX?

**ADHD: provide examples**

Distracted:  
Focus:  
Follow thru:  
Loses items:  
Organization:  
Forgetting:  
Impatience/interruption:  
Hyperactivity:

**Chronic Pain:**  
Location  
Origin  
Quality?  
Better Worse  
Cope?

**Legal History?** DUI/ incarceration/ probation/ lawsuit/ personal injury/ disability

➔ Figure 8.1 Continued
**Social History:**
Desc of family growing up

**Siblings:**

**Parents:**
Parent divorce? Pt age?

**Current living situation:**
Marital Status:

**Children:**
- **Education:** Highest grade? Grade repeated?
- Developmental/learning problems?
- **Employment:**
Problems:

**Stressors?** Medical/Work/Financial/Spousal/Kids/Oth Fam

**Strengths/Coping?**
- What do you do well?
- How handle tough situations?
- How important is change?
- How motivated to work for change?
- Spirituality:
- Recreation:
- Social support:

➔ Figure 8.1 Continued

**Documentation**

Like the interview process itself, documentation of the interview, assessment, and plan must also be efficient for the sake of easy readability and ready accessibility to other providers on the treatment team. The long narrative notes that are common in the mental health field are incompatible with integrated care. For ease of use by everyone on the treatment team, notes should be brief and structured so that other clinicians know exactly where to look to locate the information they may need. Bulleted lists, headings, and subheadings in a standardized format are particularly helpful in this regard. Often, a review of the medical providers notes with an eye to the ways in which they structure their documentation will help inform the most familiar and useful structure upon which to adapt the IBHP’s own documentation style. Figure 8.2 is a template used to organize information from the initial clinical interview. It was developed based on the structure of the medical provider notes in the authors’ clinic and it relies heavily on the structure of the medical interview, which is the format for case presentation that is most familiar to medical providers. This format also contains sections for each of the required documentation for billing for an intake assessment to be sure that the IBHP’s documentation is both practical and complete.

**Treatment Planning**

In integrated care, treatment planning is multidimensional. It includes typical mental health treatment planning, such as selecting a therapeutic modality based on the patient’s preference and symptoms, but it also includes recommendations to the PCP and finally triage to the most appropriate setting for treatment. In order to facilitate the access necessary for effective integrated care, not all consults will continue to receive treatment in the primary care setting. In most integrated settings, a significant portion of patients will require a facilitated referral to a mental health provider in the community.
Integrated Primary Care Behavioral Health Services

Ambulatory Service Record
Psychological Diagnostic Evaluation, XX minutes—90791

Chief concern:
Presenting Problem:
This is my initial appointment with XXXX. Prior to meeting XXX, I reviewed her/his medical chart, the context of which was discussed with the patient during the course of the visit. Much of the visit was spent gathering information and building rapport. XXX presented today with concerns of ________. She/He denied ________.

Symptoms:

Context:

Duration/Frequency:

Coping/Ameliorating Factors:

Exacerbators/Triggers:

Impact/Functioning:

Strengths:

Resources:

Education/Intervention

History:
Psychiatric History
Substance Use
- ETOH:
- Tobacco:
- Drugs:
- Caffeine:

Medical History
Denied significant medical history; or see progress note in medical chart dated

Psychosocial Concerns

Social History
- Family Psych History:
- Family Social History/Relationships:
Occupational:
Educational History:
Trauma History:
Religious:
Sleep:
Exercise:

Mental Status:
  General: Appropriately dressed with good hygiene (friendly and cooperative, candid, confused, indifferent, withdrawn, oppositional, evasive, guarded, shy, defensive)
  Orientation: Awake, alert, fully oriented; eye contact—(appropriate, indirect, inconsistent, darting, fixed)
  Speech: Rate—(normal, rapid, slow); rhythm—(normal, flat, excessive prosody, slurred, pressured, lively), fluency—(normal, hesitant, expansive, rambling, halting, stuttering, jerky), quantity—(normal, scant, verbose, repetitive, responds only to questions)
  Thought Processes: (linear and goal directed, tangential, circumstantial, stereotyped language, idiosyncratic, perseverative, easily redirected, indecisive); no evidence of A/V hallucinations, no evidence of delusions
  Cognitive function: (Memory, attention/concentration, insight, judgment)
  Affect: (Full range and appropriate, animated, blunted, flat, superficial, exuberant)
  Mood: (Sad, irritable, anxious, elated, elevated, despondent, optimistic, pessimistic, hopeless)
  Pain: Documented in vitals
  Risk: Denied current SI/HI ideation, intent or plan. Not determined to be an imminent risk to self or others at this time. (Or describe SI/HI assessment)

Reviewed confidentiality and its constraints. Verbally consented to collaborative care.

Formulation: XXX is a ___ yr old (single/married/divorced/widowed) woman/man who is reporting symptoms consistent with ________. She/He would likely benefit from

DSM Diagnosis:

   **Plan:** Together the patient and I developed the following plan:

   **Homework:**

** Figure 8.2 Continued

There are several important factors to consider when referring a patient to traditional mental health settings. Severity of symptoms and level of functional impairment are often heavily weighted when determining whether or not to refer out to the community for more comprehensive, longer term treatment. While this may be the IBHP’s clinical recommendation, it is not uncommon for patients to express disinterest or hesitation to pursue treatment in the community due to greater comfort with their primary care clinic, preference for the integrated care model, past unpleasant experiences in traditional mental health settings, travel/financial restrictions, wait times, and insurance limitations. Consideration of the potential success of the referral is key. In these cases, the IBHP and patient collaboratively make a decision about what is best for the patient informed by the IBHP’s clinical judgment. In the case of the unsuccessful referral, the IBHP must adjust, problem-solve, and consider harm reduction. A provider may follow up with the patient to help enhance coping while also problem-solving solutions to getting the treatment from which the patient is most likely to benefit.
In addition to patient-specific factors such as severity of illness, access to traditional mental health and willingness to engage in treatment in traditional mental health, the IBHP’s skill set and previous training may also be relevant. For example, an IBHP may not have any experience working with children under 12 years old and therefore agrees to initial evaluations with the understanding that if any level of treatment is indicated, it will be referred out to a specialist.

“Closing the Loop”: Communicating With the PCP

Gathering information collected from the clinical interview and synthesizing it into a summary that is salient to the PCP is a critical step to ensure collaborative and integrated care. Simply documenting in the chart, assuming the provider will seek out the note, often does not suffice. Several mechanisms exist for active communication between the IBHP and PCP. Electronic medical record systems often have functions embedded that allow notes to be sent to other providers or create a task or alert for another provider. The medical record or encrypted emails are simple ways to ensure the PCP has a summary of the IBHP’s assessment and recommendations. The added value of practicing in primary care is that this conversation can be had in person, providing rich discussions about concerns and next steps for patients. What is included in this summary is likely to vary, and may require the IBHP to have a discussion with the PCPs in the practice to learn about the most helpful information to include. Consider the following framework to use:

1. The referral question;
2. A summary of major emotional symptoms, functional impairments, and potential causes underlying emotional and behavioral changes (e.g., Does the patient meet criteria for a mental health diagnosis? Why or why not?);
3. The patient’s goals;
4. Information that may be relevant to the PCPs understanding of the problem and future interactions with the patient;
5. Plans for follow-up.

### CASE EXAMPLE

The following example (Table 8.1) is a typical primary care case. The patient presents for an appointment with his PCP for a medical complaint (insomnia) that is related to emotional distress. The patient hadn’t considered that behavioral health might be useful for his problem and so the IBHP and PCP must work together to explain the basis for the referral and the relationship of the PCP and IBHP.

<table>
<thead>
<tr>
<th>Table 8.1 Example Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Case Example</strong></td>
</tr>
<tr>
<td><strong>Patient</strong>: John Smith, a patient of the referring PCP for the past 10 years</td>
</tr>
<tr>
<td><strong>Context</strong>: 25-minute initial visit, following a warm handoff</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Background Information and Dialogue</strong></th>
<th><strong>Comments</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>A PCP provides the following overview of the patient prior to the IBHP and PCP entering the exam room for a warm handoff.</td>
<td>Presenting Symptoms/ Problem: Ongoing management of hypertension, high cholesterol, and prediabetes; disrupted sleep, possible grief reaction, higher</td>
</tr>
<tr>
<td>PCP: “I have a patient I’d like to meet with you, a 55-year-old man, reporting issues falling asleep at night for the past few weeks. No history of insomnia. I don’t get the sense there are severe mood or anxiety symptoms currently, but he does</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Case Example

Patient: John Smith, a patient of the referring PCP for the past 10 years

Context: 25-minute initial visit, following a warm handoff

Background Information and Dialogue

endorse feeling more stressed than usual, which I suspect is due to approaching the 7-year anniversary of his wife’s death. She died of pancreatic cancer. In the past, he’s been remarkably stoic as he approaches the anniversary of her death; he has a very tight-knit supportive family. Today was the first time I’ve seen him in rough shape in a while. He looks exhausted. He’s on quite a few medications for hypertension, hyperlipidemia, and was recently started on metformin for prediabetes. He doesn’t like the thought of adding more medication, but wanted to discuss options because he can’t function with his sleep like this. I told him that I’d like to bring in our behavioral specialist. He was on the fence to be honest, but said he’d try just about anything to be able to sleep.”

PCP: [IBHP and PCP enter the exam room.]

“Hi John, this is Dr. _____, our behavioral health provider who specializes in problems related to sleep. [PCP turns to IBHP.] Dr. _____, John and I were discussing his recent issues with trouble falling asleep and not feeling rested the next day. This is the longest span of time he’s had trouble with sleep, which is very concerning to him. We also discussed recent stressors, one of which is the anniversary of his wife’s death is next month. Also, we recently started him on metformin because his HbA1c was high and several members of his family have been diagnosed with type II diabetes. We started talking about medication options for his sleep, but we both agreed that the ideal situation wouldn’t involve adding more medications. I shared with John that some of the best strategies we have for sleep involve changing behavior. John, is there anything you want to add?”

Patient: “Nope, that just about covers it . . . I just really need to get some sleep.”

PCP: “Alright, I’m going to step out. Dr. _____ and I will touch base after and discuss next steps based on your conversation today.”

IBHP: “Hi John, it’s nice to meet you. Like Dr. _____ mentioned, I’m one of the behavioral health providers here. I’m often called in to help patients who experience problems that may be related to stress. I do want to say that Dr. _____ and I are going to work together to help you, which means it is an open line of communication between him and me. John, have you ever met with anyone in behavioral health before?”

Patient: “No, not myself. Some of my family members have, but I’ve never heard any one working with someone like you for sleep.”

A helpful way to enhance communication and patient inclusion and engagement is to reiterate the referral question to the IBHP in the room with the patient.
**Case Example**
**Patient:** John Smith, a patient of the referring PCP for the past 10 years

**Context:** 25-minute initial visit, following a warm handoff

<table>
<thead>
<tr>
<th>Background Information and Dialogue</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBHP: “I’ve heard that from quite a few of my patients. Like Dr. ____ said, one of the best strategies for addressing sleep problems is changing sleep-related behaviors. My goal is to take your expertise on yourself and combine it with some of my expertise on changing behavior and habits related to sleep, to see if we can figure out new ways to improve your sleep and avoid taking more medication.”</td>
<td>Use patient’s preferences and ideas about the most important direction to go in.</td>
</tr>
<tr>
<td>Patient: “Alright, that sounds okay.” IBHP: “Dr. ____ mentioned a few things going on including troubling sleeping and some increased stress. As I’m sure you know, these two things tend to go hand in hand, as more stress tends to affect sleep, and poor sleep causes more stress.”</td>
<td>Identify possible focus points related to the presenting problem, and present them as options for the patient. This assures that the time spent is actually on what is most important to the patient.</td>
</tr>
<tr>
<td>Patient: “That’s for sure.” IBHP: “I’m wondering, what would be most useful from your perspective for us to focus on today, talking more about the stress during the day and ways to manage this, or to focus on aspects of sleep and strategies to address this?”</td>
<td>Now the IBHP knows not to delve too deep into specific mechanisms causing stress and not to focus on stress management skills that aren’t directly related to sleep.</td>
</tr>
<tr>
<td>“Well, I’m sure both would be helpful, but to be honest, I think the sleep is making the stress during the day more unmanageable.”</td>
<td>Identify possible focus points related to the presenting problem, and present them as options for the patient. This assures that the time spent is actually on what is most important to the patient.</td>
</tr>
<tr>
<td>IBHP: “Ok, I’d like to take the next 5–10 minutes to ask you questions, so that I can really understand the sleep issues from your perspective, and so that you and I can start getting a good idea about ways to problem-solve.”</td>
<td>This can be an effective way to elicit the most information with one question. It also gets that patient to think about recent, actual events related to the problem and can elicit potential triggers.</td>
</tr>
<tr>
<td>IBHP: “I’d like you to think about the last night this past week you had trouble falling asleep. As if you’re being the commentator for a basketball game, I want you to walk me through the events of that day leading up to not being able to fall asleep.”</td>
<td></td>
</tr>
<tr>
<td>Patient: “Ok, well that would be yesterday, Tuesday. I get up for work typically around 6 a.m. Have a few cups of coffee, eat breakfast. I work as a senior project manager for a construction company, so the mornings are pretty busy. I’d say yesterday was a typical work day. Came home that evening, made dinner, watched some TV, then sat down to work on finding a new home, mortgage applications, and plans for putting my house up for sale. That’s been a little rough, I’m selling the home I lived in with my wife and raised my two boys. Then around 10 p.m. I got into bed, and my mind just wouldn’t shut off and before you know it was lying in bed wide awake for an hour and half.”</td>
<td></td>
</tr>
</tbody>
</table>
**Case Example**
**Patient:** John Smith, a patient of the referring PCP for the past 10 years
**Context:** 25-minute initial visit, following a warm handoff

<table>
<thead>
<tr>
<th>Background Information and Dialogue</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBHP: “It sounds like making the decision to sell your house was a hard one.”</td>
<td>Important data that communicates something about a healthy grief process and social support.</td>
</tr>
<tr>
<td>Patient: “Yes, it was. But, at the same time I’m excited to be moving in with my fiancé.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “How many nights per week does it take you more than 30 minutes to fall asleep?”</td>
<td></td>
</tr>
<tr>
<td>Patient: “About 3 or 4. I feel so unrested, that by the time the weekend comes, I’m sleeping in until around 11, with a few naps during the day.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “Dr. ___ mentioned this has been going on for a few weeks. How many weeks has it been that you’ve been having trouble falling asleep 3–4 nights?”</td>
<td></td>
</tr>
<tr>
<td>Patient: “I’m going on about 3 weeks now. It’s really taking a toll.”</td>
<td>Validating and normalizing patient’s experience; anyone in his position would be having difficulty.</td>
</tr>
<tr>
<td>IBHP: “I can imagine, poor sleep really starts to have a ripple effect in people’s lives and makes just getting through an already stressful day that much harder.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “Any events or changes in particular happen around the time you started having sleep issues?”</td>
<td></td>
</tr>
<tr>
<td>Patient: “Hmmm, well I guess it was around the time I decided to put my house on the market.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “What are some of your thoughts about how the two are related?”</td>
<td>It is more powerful and efficient to first see if the patient can identify the connection.</td>
</tr>
<tr>
<td>Patient: “I’m not sure . . . I mean yes it’s stressful, but lots of stressful things happen in my life, and they don’t cause this kind of sleep issue.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “You mentioned something a bit ago, when you were walking me through your day yesterday, about not being able to ‘shut your mind off’ after you went into bed. What did you notice kept running through your mind?”</td>
<td></td>
</tr>
<tr>
<td>Patient: “I guess I’m usually thinking about things related to the move. I start going through a checklist about all the stuff that needs to happen to put the house on the market and I think about a plan. I’m definitely an over-planner. I guess I also think about what my wife would say if she knew I was selling the house. I feel like it would break her heart.” [Patient starts to tear up, puts his head down.]</td>
<td></td>
</tr>
<tr>
<td>IBHP: “There are a lot of emotions and memories tied to that home, and it think it makes a lot of sense that its stirring up some conflicting emotions—it doesn’t seem like there would be any situation where this would be a simple, easy experience.”</td>
<td></td>
</tr>
</tbody>
</table>
**Case Example**

**Patient:** John Smith, a patient of the referring PCP for the past 10 years

**Context:** 25-minute initial visit, following a warm handoff

<table>
<thead>
<tr>
<th>Background Information and Dialogue</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBHP: “How do you feel you are managing your grief right now?”</td>
<td>Efficient way to directly assess whether or not the IBHP should ask more follow-up questions regarding mood symptoms and grief.</td>
</tr>
<tr>
<td>Patient: “It tends to hit me at night recently. But overall, I feel like I’m doing as well as I can. We have a family gathering planned in a few weeks to remember and celebrate her life.”</td>
<td>The IBHP glances at the PHQ-9 the patient filled out prior, and notices the only items endorsed are somatic symptoms, and a “0” on item 9, assessing suicidality. This is additional data that reinforces the decision to continue assessing sleep as primary concern.</td>
</tr>
<tr>
<td>IBHP: “That sounds wonderful.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “Does your planning and researching for putting your house on the market and finding a new home typically happen in the evening?”</td>
<td></td>
</tr>
<tr>
<td>Patient: “Yes, I’ve been trying to take care of it in the evening during the week, so I don’t have to worry about it on the weekends. You know what... as we are talking about this, I’m realizing that I don’t typically have problems falling asleep on the weekends and that’s when I’m not doing the planning.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “Sounds like you made an important connection between the time of day you plan and how much it affects your sleep.”</td>
<td></td>
</tr>
<tr>
<td>Patient: “Yeah, to be honest I never put two and two together.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “You just touched on one of the common recommendations for people with issues falling asleep, which is to make the 2–3 hours before bed as relaxing as you can. Once we get our brains running, it’s hard to slow them down.”</td>
<td>Elicit useful strategies from the past. A skill that is efficient and reveals strategies most likely to be effective.</td>
</tr>
<tr>
<td>IBHP: “Is there anything that you’ve done in that past at night that has helped you relax?”</td>
<td></td>
</tr>
<tr>
<td>Patient: “Well, when I was really stressed about work, sometimes taking a hot shower before bed would help me relax. Maybe that would help.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “That sounds like a great idea.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “So, we are just about out of time. We’ve covered quite a bit, and it sounds like you have a good next step. I’m wondering if you think it would be helpful to come back for a few more sessions to follow up.”</td>
<td></td>
</tr>
<tr>
<td>Patient: “Yes I think that’d be a good idea, this has been helpful.”</td>
<td></td>
</tr>
<tr>
<td>IBHP: “At this point in time, where do you stand with the medication?”</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
### Table 8.1 (Continued)

<table>
<thead>
<tr>
<th>Case Example</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient:</strong> John Smith, a patient of the referring PCP for the past 10 years</td>
<td></td>
</tr>
<tr>
<td><strong>Context:</strong> 25-minute initial visit, following a warm handoff</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Background Information and Dialogue</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient: “Now that I have a few ideas, I think I want to see how it goes shifting my planning schedule around before trying medication.”</td>
<td>The IBHP likely had the impulse to provide more information and suggestions about sleep behaviors during the interview.</td>
</tr>
<tr>
<td>IBHP: “Sounds good, I will let Dr. _____ know.”</td>
<td>Having a go-to handout that provides an overview of additional strategies can help the IBHP focus on other more salient aspects of the interaction (e.g., talking about what has worked for the patient in the past), knowing this information will be provided in a handout.</td>
</tr>
<tr>
<td>IBHP: “If you’re interested, I also have a few handouts on sleep tips and tricks and another on some strategies to relax and turn your mind off. This will also give you specific ideas about what to do if you find yourself lying in bed unable to sleep for more than 20–30 minutes. There is also a helpful handout on breathing techniques that slow down the part of your body that is activated when your mind is racing at night. [Patient nods.] Let’s schedule a follow-up in 2 weeks.”</td>
<td></td>
</tr>
</tbody>
</table>

**Closing the loop with the PCP:**
The IBHP checks in with the PCP at the end of the day to update her with details from meeting with the patient and the plan going forward.

“In talking with John earlier today, he said that he hasn’t been able to ‘turn his mind off’ and part of the reason might be that he has been planning and organizing details around putting his current home up for sale and trying to find a new home for him and his fiancé. He’s doing this later in the evening during the week. On top of that being inherently stressful, it’s also bringing up conflicted and painful emotions regarding his wife and the memories they had in that home. All of this causes him to ruminate on quite a few things as soon as he lies down to go to bed. He’s going to try to change the timing of when he does his planning to see if it has any effect. He voiced a preference to hold off on medications right now. He doesn’t meet full criteria for insomnia, and hasn’t tried any behavioral modification strategies yet, so I also agree with holding off on medication. I’ll see him in 2 weeks, and will keep you posted.”

This case example illustrates the ways in which clinical interviewing in integrated care utilizes the strengths of both the mental health and medical interviews. The IBHP is brought into a medical appointment to assist in the management of a patient’s insomnia. An introduction is initiated by the PCP. The PCP discusses the situation in front of the patient and explains the role of the IBHP in treatment. The IBHP then conducts a focused assessment of the presenting problem, rather than a comprehensive mental health assessment. A plan is developed that includes strategies the patient can use now to potentially improve his symptoms. Finally, the IBHP “closes the loop” by briefly reporting the plan (including decision on medication) back to the PCP.
CONCLUSION

The clinical interview in integrated primary care needs to meet the needs of mental health practice and the primary care setting. It needs to be efficient while maintaining flexibility and quickly building rapport with the patient. Although the demands are high on IBHPs, there are also advantages to practicing in integrated settings that aren’t typically available to most mental health providers. The IBHP has the support of the entire treatment team, which means more information and greater understanding about the patient, greater trust from the patient (conferred by the PCP), and better coordination and intervention support. When the IBHP knows how to adapt his or her clinical interviewing practice to a primary care setting, not only is the result acceptable, but in many cases, preferable for the patient and provider.

REFERENCES


Depressive disorders have been cited as one of the foremost causes of disability in adults and are among the most common psychological disorders seen in primary care (Qaseem, Barry, & Kansagara, 2016). In 2005, the estimated lifetime and 12-month prevalences of major depressive disorder (MDD) were reported to be 16.6% and 6.6%, respectively, and in 2010 the economic costs associated with depression were estimated at $210.5 billion, a 21.5% increase since 2005 (Greenberg, Fournier, Sisitsky, Pike, & Kessler, 2015; Kessler et al., 2005). In 2007, the World Health Organization reported that across 60 countries, depression produced the greatest reductions in health outcomes, both when reported as a single diagnosis and when diagnosed with comorbid medical conditions (Moussavi et al., 2007).

In primary care settings, the prevalence of major depression has been reported to range from 5% to 13% in adults and from 6% to 9% in geriatric populations, with 80% reported to have at least one medical comorbidity (Klinkman, 2003; O’Connor, Whitlock, Beil, & Gaynes, 2009). However, studies conducted in primary care settings indicate that only one-third to one-half of patients with major depression are accurately identified (Williams, Noël, Cordes, Ramirez, & Pignone, 2002), despite the fact that patients with depression diagnoses have significantly higher levels of both healthcare utilization and healthcare costs (Simon, VonKorff, & Barlow, 1995).

Less than optimal recognition of depressive disorders in primary care has fueled a demand for assessment instruments that can be administered easily in clinical settings to a wide variety of populations. In addition, newer requirements for evidence-based practices, quantifiable assessment data, and objective measures of treatment efficacy have created a need for valid and reliable self-report instruments, particularly in primary care where both mental health professionals and more comprehensive diagnostic measures may be unavailable. This chapter reviews three measures that have historically been used to assess depressive disorders in primary care populations: the Beck Depression Inventories (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), the Center for Epidemiological Studies Depression Scales (Radloff, 1977), and the Patient Health Questionnaires (Kroenke, Spitzer, & Williams, 2001).

THE BECK DEPRESSION INVENTORIES

Work on developing the Beck Depression Inventories began in the late 1950s by Aaron T. Beck, PhD, and his colleagues at the University of Pennsylvania (Beck et al., 1961). Since that time, the Beck measures have become among the most widely employed and systematically studied depression screening measures available, with several updated, revised, and derivative scales being published by the authors as new evidence regarding the nature and diagnostic standards of depressive syndromes has accumulated.
Beck Depression Inventory

The original version of the Beck Depression Inventory (BDI) was developed to describe symptoms and indicators of severity observed clinically that appeared to distinguish depressed versus nondepressed psychiatric patients. Developed initially to address low interclinician agreement in the diagnosis of depression, the measure was constructed as an atheoretical scale of depressive behaviors, and consisted of 21 item categories that assessed the severity of four to five graded behavioral and other observations of constructs such as “mood,” “guilty feeling,” “crying spells,” “social withdrawal,” “sleep inhibition,” “loss of appetite,” and “loss of libido” (Beck et al., 1961). Item category responses are rated on a numerical scale of 0–3 to indicate degree of symptom severity, for a range of scores between 0 and 63.

Beck Depression Inventory-Short Form

A shorter version of the BDI was modified for use in primary care in 1972. The Beck Depression Inventory-Short Form (BDI-SF) contained 13 of the original BDI items selected based on correlations to both overall BDI scores and clinician ratings of depression (Beck & Beck, 1972). Originally developed for use by primary care physicians as a screening tool to assist with decisions for the need for more comprehensive assessment of psychiatric symptoms, use of the instrument declined as criticisms regarding its use for diagnosis and treatment monitoring in both psychiatric and nonpsychiatric populations increased (Steer & Beck, 1985; Vredenberg, Krames, & Flett, 1985).

Beck Depression Inventory-Amended

An amended (revised) version of the original BDI, the Beck Depression Inventory-Amend (BDI-IA), was published by Beck and colleagues in 1979 in order to correct several technical issues, including making the number of item category responses consistent throughout, and simplifying the wording of many items (Beck, Rush, Shaw, & Emery, 1979). These changes involved reducing the number of responses for each item from either five or four to four only, eliminating redundancies in some item responses, and correcting unclear wording. Administration and scoring manuals for the BDI-IA were published in 1987 and 1993, and it was at this time that knowledge and use of the BDI-IA rather than the original BDI became more widespread; thus, many studies published using the Beck inventories employed the original BDI rather than the amended version up until the current BDI-II was published in 1996 (Brantley, Mehan, & Thomas, 2000).

Beck Depression Inventory for Primary Care

Subsequent to the publication of the BDI-II in 1996, Beck and associates revisited their previous objective of creating a short screener for use in primary care with the development of the Beck Depression Inventory for Primary Care (BDI-PC). Observational studies began to report high prevalence rates of clinical depression in medical inpatients and outpatients, and the lack of good screening instruments with high sensitivity and specificity led the authors to create a screening tool adapted from the BDI-II (Beck, Guth, Steer, & Ball, 1997). Available primary care screening instruments at the time tended to have lower diagnostic specificity, thus overestimating rates of depression due to overlap between the performance and somatic symptoms of depression and various aspects of medical illness, such as poor appetite or fatigue (Beck, Steer, Ball, Ciervo, & Kabat, 1997).
The PDI-PC consisted of seven nonsomatic items selected from the BDI-II based on whether they were consistent with the recently revised DSM-IV diagnostic criteria for MDD ("sadness" and "loss of pleasure"), whether they were indicative of clinical severity ("suicidal thoughts or wishes"), or whether they loaded highly on the cognitive factor of the BDI-II ("pessimism," "past failure," "self-dislike," "self-criticalness"). Similar to the other versions of the BDI, the seven item category responses were rated on a numerical scale of 0–3 to indicate degree of symptom severity, for a range of scores between 0 and 21. The goal of the BDI-PC was to assist medical providers in making decisions regarding the need for further evaluation by screening out medically ill patients who may demonstrate performance or somatic symptoms consistent with depression, but who do not display the affective or cognitive aspects of the disease (Beck, Steer et al., 1997). In 2000, an administration and scoring manual was published for this version of the BDI, and at that time the instrument was renamed the Beck Depression Inventory-FastScreen for Medical Patients (BDI-FS; Beck, Steer, & Brown, 2000).

### Beck Depression Inventory-Second Edition

With the adoption of new diagnostic criteria for the depressive disorders published in the third-revised and fourth editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R; DSM-IV), the need for a substantial revision to the BDI became evident as criticisms over the continued utility of the BDI-IA increased. These criticisms related most specifically to (1) lack of congruence between the item categories of the BDI-IA and the newer diagnostic criteria, (2) with the measure’s decreased efficacy in evaluating ongoing symptom severity in moderately and mildly depressed outpatients and short-term inpatients versus long-term hospitalized patients, and (3) finding that some items had been shown to reflect a gender bias (Moran & Lambert, 1983; Santor, Ramsay, & Zuroff, 1994; Vredenberg et al., 1985).

In 1996, Beck and colleagues published the second edition of the BDI, the Beck Depression Inventory-Second Edition (BDI-II), a substantially revised version of the scale that was developed specifically to assess symptoms and severity of depression in adolescents and adults based on the DSM-IV criteria for depressive disorders. Changes in the BDI-II included removing items related to "weight loss," "body image change," "somatic preoccupation," and "work difficulties," and replacing them with "agitation," "worthlessness," "concentration difficulties," and "loss of energy." Items relating to sleep and appetite were altered to allow for either increases or decreases in these symptoms, and 18 items assessing similar domains covered in the BDI-IA were retained, although several were renamed (e.g., "self-accusations" was renamed "self-criticalness"). Three items were left unaltered from the previous version: "punishment feelings," "suicidal thoughts or wishes," and "loss of interest in sex." Likewise, in order to be consistent with DSM-IV criteria for MDD, the respondent time frame was extended from one week to "past two weeks, including today" (Beck, Steer, Ball, & Ranieri, 1996). Scoring of the BDI-II is similar to past versions; 21 items are rated on a numerical scale of 0–3 to indicate degree of symptom severity, for a range of scores between 0 and 63. Finally, DSM-IV severity modifiers were incorporated into the scoring ranges, such that total BDI-II scores of 0–13 represent "minimal" depression; 14–19 represent "mild" depression; 20–28 represent "moderate" depression; and 29–63 represent "severe" depression (Beck, Steer, & Brown, 1996).

### Beck Depression Inventory-FastScreen for Medical Patients

As noted above, the BDI-FS was previously known as the BDI-PC, and consists of items 1–4 and 7–9 of the BDI-II. Similar to the other versions of the BDI, the seven item category responses are rated
on a numerical scale of 0–3 to indicate degree of symptom severity, for a range of scores between 0 and 21. Score ranges for the BDI-FS indicate that total scores of 0–3 represent “minimal” depression; 4–6 represent “mild” depression; 7–9 represent “moderate” depression; and 10–21 represent “severe” depression (Beck et al., 2000).

Psychometric Properties and Diagnostic Efficiency of the BDI-II and BDI-FS

BDI-II

Initial psychometric properties reported by the authors were derived from administering the BDI-II as part of a standard psychiatric intake battery to 500 outpatients and 120 normal controls from undergraduate psychology courses. The clinical sample consisted of 317 females and 183 males, 91% of whom were Caucasian, 4% African American, 4% Asian American, and 1% Latino, with a mean age of 37.20 (SD = 15.91). The control sample was predominately Caucasian, consisting of 67 females and 53 males with a mean age of 19.58 (SD = 1.84). Internal consistency reliability was .92 for the clinical samples and .91–.93 in the controls; 1-week test-retest reliability in the control sample was .93, and .92 in the clinical sample. Receiver operating curve (ROC) analysis was used to derive the aforementioned cut scores for minimal, mild, moderate, and severe depression, and convergent validity was demonstrated via correlations of .93 with the BDI-IA and .71 with the Hamilton Psychiatric Rating Scale for Depression (HAMD; Beck, Steer, & Brown, 1996; Hamilton, 1960). Nonsignificant relationships were also found between BDI-II scores and age, race, and ethnicity in a smaller sample of 140 psychiatric outpatients (Beck, Steer, Ball et al., 1996), and a consistent two-factor structure describing cognitive and somatic-affective symptoms of depression has been demonstrated in multiple psychiatric and nonpsychiatric samples (Beck, Steer, & Brown, 1996; Steer, Rissmiller, & Beck, 2000; Storch, Roberti, & Roth, 2004).

Norms were derived from administering the BDI-II to a sample of 7,500 community-dwelling adults in the Netherlands, with sex, age, and educational level hypothesized to predict BDI-II scores (Roelofs et al., 2012). Analyses indicated that while sex and educational level predicted changes in total scores, age was not a significant predictor. Table 9.1 lists raw BDI-II total scores and percentiles for sex and educational level. The authors note that any score at or above the 80th percentile can be

<table>
<thead>
<tr>
<th>Table 9.1 BDI-II Norms for Sex and Educational Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>Male:</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Female:</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total:</td>
</tr>
</tbody>
</table>

“Low” = no education, elementary school, or lower technical and vocational training
“Medium” = medium technical and vocational training, higher general secondary education, or pre-university education
“High” = bachelor’s degree and above

considered elevated depression levels, and scores at or above the 95th percentile can be considered high depression scores.

As comprehensively reviewed in 2013, the BDI-II has been translated and validated in 17 languages, and extensively studied in medical populations, including primary care, cardiology, neurology, brain injury, nephrology, chronic pain, chronic fatigue, oncology, and infectious disease (Wang & Gorenstein, 2013a, 2013b). In analyzing results with medically ill samples, the authors noted good sensitivity ranging from 72% to 100%, adequate specificity ranging from 60% to 100%, cut scores for major depression ranging widely between 7 and 22, and significant diagnostic accuracy of 82% or higher, as calculated via ROC analyses. They additionally note that differentiating between depressed and nondepressed groups depends not only on the sensitivity and specificity of cut scores, but also on the incidence of depression in the sample, and variation in sample characteristics such as inpatient or outpatient, comorbid medical disease, and which external gold standard criterion is used to diagnose depression. Thus, despite this variance in cut scores, the BDI-II is recommended as a screening tool to prevent excessive cases of false positives. However, caution is warranted when using the cut scores for criterion-referenced interpretation, or for using the BDI-II as a diagnostic instrument.

BDI-FS
The BDI-FS and its predecessor the BDI-PC have been less extensively studied than the BDI-II; however, psychometric properties published by Beck and colleagues for the BDI-FS have been reported for five samples of medical patients: 50 medically ill hospitalized patients subsequently referred to the psychiatric consultation and liaison (C&L) service; 94 family medicine clinic patients; 100 pediatric patients aged 12–17 receiving well-child visits; 120 university-based internal medicine clinic patients; and 75 geriatric medicine clinic patients (Beck et al., 2000; Scheinthal, Steer, Giffin, & Beck, 2001). Internal consistency reliabilities for the five groups were .86, .85, .88, .86, and .83 respectively; however, no test-retest reliability scores were reported. The correlation between the 7-item BDI-FS total score and the 21-item BDI-II total score was reported at .91 for samples of college students and medical outpatients, and the ROC analysis-derived cut score of four yielded sensitivities of 70%–100% and specificities of 83%–99% for detecting a diagnosis of MDD in adolescent, adult, and geriatric primary care outpatients. Convergent validity was demonstrated via adequate correlations with the depression subscale of the Hospital Anxiety and Depression Scale, $r = .62$ (Zigmond & Snaith, 1983), a clinically derived DSM-IV mood disorder diagnosis, $r = .69$, and the Geriatric Depression Scale-Short Form, $r = .81$ (Sheikh & Yesavage, 1986); however, divergent validity was not demonstrated with the Beck Anxiety Inventory for Primary Care, $r = .86$ (Mori et al., 2003).

More recently, Kliem, Mossle, Zenger, and Brahler (2014) reported reliability and validity data for the German language version of the BDI-FS based on a significantly larger sample of 2,467 who participated in a national general population survey of persons older than 14 years in Germany (Gierk et al., 2014). The final sample included 53% females with an average age of 49.4 years and a range of 14–91 years. Item and item-total correlations were found not to differ between men and women, and higher BDI-FS scores were associated with higher age, lower income, female sex, lower educational achievement, unemployment, and not having a romantic partner. Internal consistency reliability was reported at .84, and total BDI-FS scores demonstrated convergent validity with the PHQ-9, $r = .67$ (Kroenke et al., 2001); however adequate divergent validity was not demonstrated with the Generalized Anxiety Disorder Scale-2, $r = .60$ (Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007), or the Somatic Symptom Scale-8, $r = .57$ (Gierk et al., 2014). Finally, a confirmatory factor analysis revealed a very good fit for a one-factor solution (RMSEA = 0.58, 90% CI [.049, .074], CFI = 0.90, TLI = 0.986), and high factor loadings (.73–.90).
Consistency of the BDI-II and BDI-FS with DSM-5 and ICD-10 Diagnostic Criteria

As noted previously, the Beck measures were not developed to be diagnostic measures of clinical depressive disorders, and therefore do not incorporate all criteria necessary to make a diagnosis of depression by the current standards outlined in either the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association, 2013) or the International Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) (World Health Organization, 2016). As reflected in Table 9.2, BDI-II items do assess most of the individual “A” criteria and the 2-week time frame needed for a diagnosis of MDD; however, the other DSM-5 criteria (“B” through “E”) are not covered at all by either of the BDI measures. Criteria not assessed by the BDI-II or the BDI-FS include the following requirements: either depressed mood or anhedonia be present; the symptoms represent a change from previous functioning; the symptoms cause clinically significant distress or impairment in social, occupational, or other areas of functioning; other contributing or comorbid factors such as substance use, general medical conditions, or other psychiatric conditions be ruled out; or the symptoms are not better accounted for by bereavement or the physiological effects of a medication or other medical condition. Likewise, not all of the severity and course specifiers listed for MDD are assessed by the BDI measures, including discriminations between single, recurrent, or unspecified episodes; whether the depressive episode is in full or partial remission; or whether the current episode occurs with anxious distress, mixed features, melancholic features, atypical features, mood-congruent psychotic features, mood-incongruent psychotic features, catatonia, peripartum onset, or a seasonal pattern. In addition, by definition as a primary care screener, the BDI-FS does not include items that assess the DSM-5 somatic or performance symptoms of depression (see Table 9.2); and neither BDI measure discriminates clinically between MDD and other DSM-5 depressive disorders. Therefore, the use of supplemental measures, diagnostic interviews, and/or clinical judgment is needed in order to make a definitive diagnosis of clinical depression when using the BDI-II and BDI-FS.

In a similar manner to the incongruences noted earlier between the Beck measures and DSM-5 diagnostic criteria, inconsistencies are also present between the BDI-II, the BDI-FS, and the primary ICD-10-CM diagnostic criteria for a depressive episode (see Table 9.3). For example, while the BDI-II was constructed with DSM “A” criteria in mind, some of the diagnostic symptoms of depression noted in the ICD-10-CM are either not present or mentioned only in the specifiers of

<table>
<thead>
<tr>
<th>DSM-5 Criteria</th>
<th>BDI-II Items (item #)</th>
<th>BDI-FS Items (item #)</th>
</tr>
</thead>
<tbody>
<tr>
<td>depressed mood</td>
<td>sadness; crying (1; 10)</td>
<td>Sadness (1)</td>
</tr>
<tr>
<td>anhedonia</td>
<td>loss of pleasure; loss of interest (4; 12)</td>
<td>loss of pleasure (4)</td>
</tr>
<tr>
<td>weight or appetite loss/gain</td>
<td>appetite changes (18)</td>
<td>none</td>
</tr>
<tr>
<td>insomnia or hypersomnia</td>
<td>sleep changes (16)</td>
<td>none</td>
</tr>
<tr>
<td>psychomotor retardation or agitation fatigue</td>
<td>agitation (11)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>loss of energy; tiredness or fatigue (15; 20)</td>
<td>none</td>
</tr>
<tr>
<td>feelings of worthlessness or guilt concentration difficulties</td>
<td>worthlessness; guilty feelings (14; 5)</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>indecisiveness; concentration difficulty (13; 19)</td>
<td>none</td>
</tr>
<tr>
<td>suicidal ideation</td>
<td>suicidal thoughts or wishes (9)</td>
<td>suicidal thoughts or wishes (7)</td>
</tr>
</tbody>
</table>
Table 9.3 BDI-II/BDI-FS Consistency With ICD-10-CM Criteria for Depressive Episode

<table>
<thead>
<tr>
<th>ICD-10 Criteria</th>
<th>BDI-II Items (item #)</th>
<th>BDI-FS Items (item #)</th>
</tr>
</thead>
<tbody>
<tr>
<td>depressed mood</td>
<td>sadness; crying (1; 10)</td>
<td>sadness (1)</td>
</tr>
<tr>
<td>reduction in energy; decrease in activity</td>
<td>loss of energy; tiredness or fatigue (15; 20)</td>
<td>none</td>
</tr>
<tr>
<td>anhedonia</td>
<td>loss of pleasure; loss of interest (4; 12)</td>
<td>loss of pleasure (4)</td>
</tr>
<tr>
<td>reduced concentration</td>
<td>concentration difficulty (19)</td>
<td>none</td>
</tr>
<tr>
<td>tiredness</td>
<td>tiredness or fatigue (20)</td>
<td>none</td>
</tr>
<tr>
<td>sleep disturbance</td>
<td>sleep changes (16)</td>
<td>none</td>
</tr>
<tr>
<td>reduced appetite or weight loss</td>
<td>appetite changes (18)</td>
<td>none</td>
</tr>
<tr>
<td>reduced self-esteem</td>
<td>self-dislike (7)</td>
<td>self-dislike (7)</td>
</tr>
<tr>
<td>reduced self-confidence</td>
<td>self-criticalness (8)</td>
<td>self-criticalness (8)</td>
</tr>
<tr>
<td>guilt; worthlessness</td>
<td>guilty feelings; worthlessness (5; 14)</td>
<td>none</td>
</tr>
<tr>
<td>depression worse in morning</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>psychomotor retardation</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>agitation</td>
<td>agitation (11)</td>
<td>none</td>
</tr>
<tr>
<td>loss of libido</td>
<td>loss of interest in sex (21)</td>
<td>none</td>
</tr>
</tbody>
</table>

the DSM-5. Therefore, while a few of the ICD-10-CM depression criteria are analogously absent in the BDI measures, such as “depression worse in the morning” and “psychomotor retardation,” others are included that are not represented in the DSM-5 criteria, such as “reduced self-esteem” and “reduced self-confidence.” Like the inconsistencies noted earlier, the ICD-10-CM severity modifiers and other diagnostic categories are also not as well represented in the BDI measures, in addition to several older terms describing depressive symptom clusters, such as the presence of “psychogenic,” “reactive,” or “vital” depression. These include the ICD-10-CM specific definitions and diagnostic categories of mild, moderate, severe without psychotic features, severe with psychotic features, atypical, “masked,” in remission, dysthymia, depression not otherwise specified, and recurrent depression. As a result, the use of supplemental measures, diagnostic interviews, and/or clinical judgment is needed in order to make a diagnosis of clinical depression when using the BDI-II and BDI-FS within the ICD-10-CM classification scheme.

**CENTER FOR EPIDEMIOLOGIC STUDIES DEPRESSION SCALES**

The Center for Epidemiologic Studies Depression Scale (CES-D) was published in 1977 by Lenore Sawyer Radloff, a psychometrician and biostatistician at the National Institute of Mental Health (NIMH). Radloff developed the CES-D to serve as a psychometrically sound measure that could be used to assess depressive symptoms in large epidemiological surveys of the general population. The CES-D is also one of the most widely utilized depression screening measures available, and since its initial publication, other versions have been developed by multiple authors for use in various clinical and nonclinical populations.

**Center for Epidemiologic Studies Depression Scale**

The Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977) is a 20-item self-report measure developed by the National Institute of Mental Health (NIMH) for use in epidemiological studies to identify depressive symptomatology in nonpsychiatric community samples. The CES-D differs from previous scales in that it was not developed to be a clinical measure of disease diagnosis, course,
Using the scale below, indicate the number that best describes how often you felt or behaved this way during the past week:

<table>
<thead>
<tr>
<th></th>
<th>Rarely or none of the time</th>
<th>Some or little of the time</th>
<th>Occasionally or a moderate amount of time</th>
<th>Most or all of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(less than 1 day)</td>
<td>(1–2 days)</td>
<td>(3–4 days)</td>
<td>(5–7 days)</td>
</tr>
</tbody>
</table>

1. I was bothered by things that don’t usually bother me. 0 1 2 3
2. I did not feel like eating, my appetite was poor. 0 1 2 3
3. I felt that I could not shake the blues, even with help from my family and friends. 0 1 2 3
4. I felt that I was just as good as other people. 0 1 2 3
5. I had trouble keeping my mind on what I was doing. 0 1 2 3
6. I felt depressed. 0 1 2 3
7. I felt that everything I did was an effort. 0 1 2 3
8. I felt hopeful about the future. 0 1 2 3
9. I thought my life had been a failure. 0 1 2 3
10. I felt fearful. 0 1 2 3
11. My sleep was restless. 0 1 2 3
12. I was happy. 0 1 2 3
13. I talked less than usual. 0 1 2 3
14. I felt lonely. 0 1 2 3
15. People were unfriendly. 0 1 2 3
16. I enjoyed life. 0 1 2 3
17. I had crying spells. 0 1 2 3
18. I felt sad. 0 1 2 3
19. I felt that people disliked me. 0 1 2 3
20. I could not get “going.” 0 1 2 3

Adapted from the Center for Epidemiological Studies Depression Scale.

Table 9.4 The Center for Epidemiologic Studies Depression Scale (CES-D)

or symptom severity, but rather a cross-sectional scale designed to measure associations between the affective components of depressive symptoms and other variables of interest in the general population. Items were derived from a combination of those used in previously validated instruments (see Beck et al., 1961; Dahlstrom & Welsh, 1961; Raskin, Schulterbrandt, Reatig, & McKeon, 1969; Zung, 1965) and primary components identified from factor analytic studies and clinical judgment. The primary components from which the items were derived included “depressed mood,” “feelings of guilt and worthlessness,” “feelings of helplessness and hopelessness,” “psychomotor retardation,” “loss of appetite,” and “sleep disturbance.” Items are both positively and negatively worded, and respondents rate the frequency of occurrence of each symptom on a 4-point Likert scale ranging from “Rarely or none of the time (less than 1 day)” to “Most of the time or all of the time (5–7 days).” To score the CES-D, 20 items are rated on a numerical scale of 0–3 to indicate how often the respondent “felt or behaved this way in the past week,” for a range of scores between 0 and 60. The CES-D is available in the public domain, and can be administered, duplicated, and published free of charge. Items and responses for the CES-D are reflected in Table 9.4.

Center for Epidemiologic Studies Depression Scale-Revised

More recently, Eaton, Smith, Ybarra, Muntaner and Tien (2004) revised the original CES-D with the goal of more accurately reflecting current DSM diagnostic criteria for MDD. Several scale versions,
response categories, and scoring algorithms were tested, with original CES-D items compared to new
and revised items based on DSM-IV symptom clusters and previous work by Zimmerman and Coryell
(1994). Items assessing “anhedonia,” “psychomotor retardation or agitation,” and “suicidal ideation”
were added, several items assessing positive affect were revised or removed, and reverse-scored items
were inversely worded or removed. The scale is categorized via DSM diagnostic criteria such that
three items each assess “dysphoria” and “sleep disturbance,” and two items each assess “anhedonia,”
“appetite,” “thinking and concentration,” “guilt,” “fatigue,” “agitation,” and “suicidal ideation.”
Congruent with DSM criteria, the scale also adds a fifth response option to reflect a 2-week duration
of symptoms (Eaton et al., 2004). The final scale consists of 20 items with five response categories
each, with the rate of frequency of occurrence of each symptom rated on a 4-point Likert scale rang-
ing from “Not at all, or less than one day” (in the past week) to “Nearly every day for 2 weeks.” In
order to retain the same scoring range as the original version, the CESD-R uses an uppermost value
of “3” for both of the two most frequent response categories, resulting in a range of scores between
0 and 60.

To score the CESD-R, the 20 items are rated on a numerical scale of 0–3 to indicate how often the
respondent “felt or behaved in the past week or so.” The current version of the CESD-R is available
via the Internet, and information posted with the web-based version includes a scoring algorithm for
the CESD-R, such that possible depressive diagnostic categories are proposed as follows (http://cesd-r.
com/):

- **Meets criteria for major depressive episode**: Anhedonia or dysphoria nearly every day for the past two
  weeks, plus symptoms in an additional four DSM symptom groups noted as occurring nearly every
day for the past two weeks;
- **Probable major depressive episode**: Anhedonia or dysphoria nearly every day for the past two weeks,
  plus symptoms in an additional three DSM symptom groups reported as occurring either nearly
every day for the past two weeks, or 5–7 days in the past week;
- **Possible major depressive episode**: Anhedonia or dysphoria nearly every day for the past two weeks,
  plus symptoms in an additional two other DSM symptom groups reported as occurring either
nearly every day for the past two weeks, or 5–7 days in the past week;
- **Subthreshold depression symptoms**: People who have a CESD-style score of at least 16 but do not
  meet above criteria;
- **No clinical significance**: People who have a total CESD-style score less than 16 across all 20 ques-
tions.

Items and responses for the CESD-R are presented in Table 9.5.

**Table 9.5 The Center for Epidemiologic Studies Depression Scale-Revised (CESD-R)**

Below is a list of the ways you might have felt or behaved.

Please check the boxes to tell me how often you have felt this way in the past week or so.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not at all or Less than 1 day</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1–2 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3–4 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5–7 days</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Nearly every day for 2 weeks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. My appetite was poor.
2. I could not shake off the blues.
Table 9.5 (Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>I had trouble keeping my mind on what I was doing.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4.</td>
<td>I felt depressed.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5.</td>
<td>My sleep was restless.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>6.</td>
<td>I felt sad.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>7.</td>
<td>I could not get going.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8.</td>
<td>Nothing made me happy.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>9.</td>
<td>I felt like a bad person.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>10.</td>
<td>I lost interest in my usual activities.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>11.</td>
<td>I slept much more than usual.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>12.</td>
<td>I felt like I was moving too slowly.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>13.</td>
<td>I felt fidgety.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>14.</td>
<td>I wished I were dead.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>15.</td>
<td>I wanted to hurt myself.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>16.</td>
<td>I was tired all the time.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17.</td>
<td>I did not like myself.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18.</td>
<td>I lost a lot of weight without trying to.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19.</td>
<td>I had a lot of trouble getting to sleep.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20.</td>
<td>I could not focus on the important things.</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Adapted from the Center for Epidemiological Studies Depression Scale—Revised.

Psychometric Properties and Diagnostic Efficiency of the CES-D and CESD-R

**CES-D**

Initial psychometric properties were derived from administering the CES-D as part of a larger 300-item NIMH household interview of 3,934 homes in the Midwest selected by independent probability sampling, and in two clinical samples of 105 psychiatric patients. Internal consistency reliability was reported to be .84–.85 for the household samples and .90 for the clinical samples; split-half reliabilities were .76–.77 and .85; and Spearman-Brown predictions were .86–.87 and .92, respectively. Test-retest reliability analyses ranging from 2 to 8 weeks in the community sample and from 3 to 12 months in the psychiatric sample were performed, with values ranging from .51–.67 and .32–.54, respectively. Normative analyses reported in the initial samples indicated a consistent factor structure and good internal consistency reliability estimates of .80 and above in men, women, Caucasians, African Americans, varying age groups (under 25 years, 25–64 years, and 65 years and above), and varying educational ranks (below high school level, high school level, and above high school) (Radloff, 1977).

Principal components analyses in each initial sample and subgroup identified a consistent four-factor structure, with the factors described as, “Depressed affect (blues, depressed, lonely, cry, sad)”; “Positive affect (good, hopeful, happy, enjoy)”; “Somatic and retarded activity (bothered, appetite, effort, sleep, get going)”; and “Interpersonal (unfriendly, dislike),” although due to high internal consistency values, Radloff (1977) recommends only the use of a single total score when the scale is used for epidemiological studies. In addition, normative analyses were performed using a sample of 3,059 respondents aged 25 to 74 years from the first National Health and Nutrition Survey (NHANES-I; Centers for Disease Control and Prevention & National Center for Health Statistics, 1971), and mean scores of 7.1 (SD = 7.2) were reported for men and 10.0 (SD = 9.1) were reported
for women (Sayetta & Johnson, 1980). No statistical differences for various age and racial subgroups were reported.

Convergent and concurrent validity was demonstrated via significant correlations with other measures of depression including the CES-D’s ability to discriminate between the community and clinical samples at an arbitrary cut score of 16, a correlation of .56 with psychiatric nurse ratings of depression severity in the clinical samples, positive correlations between CES-D total scores and self-reported negative life events, and significant positive correlations with other clinician and self-report measures of depression, $r = .51$–.89, including the Hamilton Depression Scale (Hamilton, 1960), the Beck Depression Inventory (Beck et al., 1961), the Raskin Rating Scale (Raskin et al., 1969), and the Symptom Checklist-90 (Derogatis, Lipman, & Covi, 1973).

More recently, Carleton et al. (2013) have used confirmatory factor analyses and item response theory to propose a 14-item, three-factor model that they describe as more consistent with current diagnostic criteria for MDD. In order to test all previously reported factor structures for the CES-D, new samples of undergraduate students ($n = 948; 74\%$ women), community members ($n = 254; 71\%$ women), tertiary care rehabilitation patients ($n = 522; 53\%$ women), and clinically depressed outpatients ($n = 84; 77\%$ women) from across Canada were analyzed, as well as previously collected community-based samples derived from the NHANES-I data ($n = 2,814; 56\%$ women) (Centers for Disease Control and Prevention & National Center for Health Statistics, 1971). Their analyses suggest a more robust three-factor solution, with the factors interpreted as “negative affect,” “anhedonia,” and “somatic symptoms.” They also note that item 17, “I had crying spells,” performed differently in women versus men, which resulted in inflated total scores for women, particularly at higher levels of depression. Therefore, in addition to item 17, the items eliminated from their final model were item 9, “I thought my life had been a failure”; item 10, “I felt fearful”; item 13, “I talked less than usual”; item 15, “People were unfriendly”; and item 19, “I felt that people disliked me” (Carleton et al., 2013).

As noted earlier, Radloff (1977) originally chose an arbitrary cut score of 16, which in the initial samples discriminated both between the patient and community-dwelling samples, and between patient samples with varying degrees of depression. Several studies have suggested that a total score of 16 or above overestimates depression prevalence and sacrifices specificity for sensitivity. In their comprehensive review and meta-analysis of the use of CES-D to screen for depression in both general population and primary care samples, Vilagut, Forero, Barbaglia, and Alonso (2016) reported on 28 studies that assessed sensitivity, specificity, and the diagnostic odds ratio of the CES-D at cut-points of 16 ($n = 22$ studies), 20 ($n = 12$ studies) and 22 ($n = 7$ studies). The gold standard used to diagnose depression varied between studies, but most commonly included the Diagnostic Interview Schedule (Robins, Helzer, Croughan, & Ratcliff, 1981) ($n = 7$ studies); the Structured Clinical Interview for DSM (Spitzer, Gibbon, & First, 1990) ($n = 6$ studies); the Composite International Diagnostic Interview (Robins et al., 1988) ($n = 4$ studies); or the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998) ($n = 4$ studies). The median prevalence of major depression was calculated at 8.8% across studies. Settings in which the studies were conducted included 10 in primary care, eight in the general population, six in school settings, and four in geriatric settings. Ten studies used foreign language versions of the CES-D, 11 studies focused exclusively on geriatric populations, and six were conducted in adolescent populations.

For all studies analyzed, Vilagut et al. (2016) reported a median sensitivity of .85 (range .40–1.0) and median specificity of .72 (range .44–.90). Sensitivity, specificity, and diagnostic odds ratio for studies employing a cut score of 16 were .87, .70, and 16.24, respectively, whereas values for a cut score of 20 were .83, .78, and 16.64, and those employing a cut score of 22 were .79, .80, and 14.68. This suggests that a cut score of 20 provided the best balance of diagnostic efficiency. Similarly, while Vilagut and colleagues recommend a cut score of 20 be utilized when the CES-D is administered in
research and epidemiological settings, they state unequivocally that the CES-D should not be used as a stand-alone measure of depression diagnosis or severity. In addition, several previous studies have alternatively suggested that cut scores of 25 to 27 be used when screening for depression in psychiatric, chronic disease, and elderly populations, but that other, more extensive assessment measures must supplement its use for clinical diagnosis or treatment planning (Breslau, 1985; Haringsma, Engels, Beekman, & Spinhoven, 2004; Santor, Zuroff, Ramsay, Cervantes, & Palacios, 1995; Schulberg et al., 1985).

**CESD-R**

Initial psychometric properties on the final scale version were reported by Eaton and colleagues (2004), and were derived via administering the CES-D-R and the original CES-D to a sample of 868 female nursing assistants employed in 50 nursing homes throughout West Virginia and Ohio. Internal consistency reliability for the CESD-R was reported at .93, and item-total correlations ranged from .32 to .75. Correlation between total CES-D and total CESD-R scores was .88, however, a lower proportion of the sample was reported to meet a cut score of 16 on the revised version versus the original (CES-D ≥ 16 = 73.5%; CESD-R ≥ 16 = 52.4%).

More recently, Van Dam and Earleywine (2011) reported the results of a much larger validation study of the CESD-R conducted in two samples: respondents to a web-based listserv, and undergraduate students. Participants in the web-based listserv responded to an email request issued to the National Organization for the Reformation of Marijuana Laws listserv (NORML; Washington, DC). This sample was composed of 6,971 respondents who completed the web-based version of the CESD-R, the majority of whom were young white males (male = 80.7%; white = 89.4%; mean age = 30.6 years). The undergraduate sample was composed of 243 undergraduate psychology students who completed the paper-and-pencil version of the scale, the majority of whom were young white females (female = 62.0%; white = 72.8%; mean age = 19.6 years). In the web-based sample, mean CESD-R scores were 10.3 with a standard deviation of 11.7; in the undergraduate sample, mean CESD-R scores were 16.4 with a standard deviation of 13.5. Internal consistency reliabilities were reported at .92 and .93, respectively. Exploratory factor analysis conducted on a random split-half of the web-based sample (n = 3,528) supported a two-factor solution, with the factors interpreted as “negative mood” and “functional impairment.” However, two confirmatory factor analyses performed in the second split half of the web-based sample (n = 3,443) and in the undergraduate sample (n = 207) found a one-factor solution to be most parsimonious, due to high interfactor correlational coefficients (.94 and .98, respectively).

Convergent validity was demonstrated in the undergraduate sample by a significant positive correlation (r = .58, p ≤ .001) with the Negative Affect Scale of the Positive and Negative Affect Scale (Watson, Clark, & Tellegen, 1988), and divergent validity demonstrated by a significant negative correlation (r = −.26, p ≤ .001) with the Positive Affect Scale in the same sample. However, adequate divergent validity was not demonstrated between the CESD-R and total scores on the State-Trait Inventory for Cognitive and Somatic Anxiety (Grös, Antony, Simms, & McCabe, 2007) due to significant positive correlations in the both the web-based sample (r = .73, p ≤ .001) and the undergraduate sample (r = .65, p ≤ .001). Finally, based on the score ranges and diagnostic categories described earlier, the authors compared the traditional CES-D cut score of 16 with the algorithmic classification described by Eaton et al. (2004) for “probable depression episode,” which calculates a cut score of 24 with the inclusion of items endorsing either dysphoria or anhedonia. They report that in the web-based sample, the proportion of respondents meeting the algorithmic classification of probable depression was 4.6% (95% CI = 4.98–5.57), a figure similar to the 5.28% 12-month prevalence of MDD reported by Hasin and colleagues (2005) based on the first wave of the National Epidemiologic Survey on
Alcoholism and Related Conditions (National Institute on Alcohol Abuse and Alcoholism, 2001). In contrast, using the 16 cut score in the web-based sample resulted in 16.5% of the sample being classified as depressed (Van Dam & Earleywine, 2011).

**Short Forms of the CES-D and CESD-R**

An 8-item version of the CES-D assessing “depressed mood,” “sleep disturbance,” “unhappiness,” “loneliness,” “anhedonia,” “sadness,” and “amotivation” has been tested extensively in European populations as part of the European Social Survey (ESS Round 3: European Social Survey, 2014) and the Survey of Health, Ageing, and Retirement (Börsch-Supan et al., 2008). These large epidemiological studies have found good internal consistency reliability ($\alpha = .81–.84$), a single-factor structure, and measurement invariance between men and women of 25 European countries (Van de Velde, Bracke, Levecque, & Meuleman, 2010), and between middle-aged (50–64 years) and older respondents (65–79 years) of 11 European countries (Missinne, Vandeviver, Van de Velde, & Bracke, 2014). Consistent with other large-scale epidemiological studies, these analyses also found significantly higher depression scores in women versus men in all countries except Ireland and Finland, and in older versus middle-aged respondents in all countries except Switzerland.

Likewise, a previously developed 10-item version of the CES-D (Andresen, Malmgren, Carter, & Patrick, 1994) assessing “irritability,” “concentration problems,” “depressed mood,” “anhedonia,” “hopelessness,” “anxiety,” “sleep disturbance,” “unhappiness,” “loneliness,” and “amotivation” has been more recently tested in large samples of Alzheimer’s and elder caregivers (Andresen, Byers, Friary, Kosloski, & Montgomery, 2013), psychiatric partial-day program patients (Bjorgvinsson, Kertz, Bigda-Peyton, McCoy, & Aderka, 2013), and an epidemiological study of Spanish-speaking respondents of the Hispanic Community Health Study/Study of Latinos (Gonzalez et al., 2016). Good internal consistency reliability ($\alpha = .80–.89$) and a single-factor structure was demonstrated in all three studies; measurement invariance was also demonstrated between three separate samples in the caregiver study, and between respondents of six different Hispanic/Latino background groups. The Hispanic/Latino study also reported good 1–3 week and 3–6 month test-retest reliabilities (.70 and .53, respectively), strong positive convergent validity with the Patient Health Questionnaire-9, $r = .80$ (Kroenke et al., 2001), but a lack of divergent validity with the Trait subscale of the State-Trait Anxiety Inventory, $r = .72$ (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983).

In the psychiatric sample, Bjorgvinsson and colleagues (2013) reported strong positive convergent validity between the CESD-10 and the Depression and Functioning scale of the Behavior and Symptom Identification Scale-24 (BASIS-24), $r = .86$ (Eisen, 2011); acceptably moderate correlation with the Penn State Worry Questionnaire-Abbreviated, $r = .46$ (Meyer, Miller, Metzger, & Borkovec, 1990); and low correlation with the Psychotic Symptom and Substance Abuse subscales of the BASIS-24 ($r = .18$ and .09, respectively). Divergent validity was demonstrated in the psychiatric sample by negative correlations with the Schwartz Outcome Scale, $r = -.54$ (Blais et al., 1999) and the Emotion Regulation Questionnaire, $r = -.23$ (Gross & John, 2003); and concurrent validity was reported in the portion of the sample diagnosed as having a current major depressive episode versus those not ($U = 46835.50$, $Z = -7.96$, $p \leq .001$, Mdn = 20.00 and Mdn = 14.00). In addition, an ROC analysis compared sensitivity (Sn), specificity (Sp), and positive and negative predictive values (PPV and NPV, respectively) for various cut scores, concluding that in this psychiatric sample, a cut score of 15 demonstrated the best diagnostic efficiency (Sn = .76; Sp = .75; PPV = .89; NPV = .53) as opposed to the cut score of 10 (Sn = .89; Sp = .47; PPV = .82; NPV = .62) previously suggested in studies with nonpsychiatric samples (Andresen et al., 1994).
In addition, Haroz, Ybarra, and Eaton (2014) have recently reported on a shorter, 10-item version of the CESD-R that has been revised to assess depression in adolescents. Referred to as the Center for Epidemiologic Studies Depression Scale-Revised 10-item Version for Adolescents (CESDR-10), this scale was derived by using the nine items from the CESD-R the authors report to be most predictive of diagnostic depression in adolescents, and adding an item assessing “irritability,” a feature more frequently seen in adolescent versus adult depression. In two separate samples of adolescents between the ages of 13 to 18 (older sample: $N = 3,777$; age = 15.7 years; female = 56.6%; white = 75.2%, in high school = 72.0%) and 10 to 15 (younger sample: $N = 1150$; age = 14.5 years; male = 50.6%; white = 74.3%, in high school = 62.7%), excellent internal consistency reliability ($\alpha = .91, .90$), a single-factor structure, and measurement invariance between males and females was demonstrated. The authors also report that in the older sample, total CESDR-10 scores showed divergent validity with a measure of self-esteem, $r = -.56$ (Rosenberg, 1965) and convergent validity with negative parent–child relationships, $r = .35$ (Finkelhor, Mitchell, & Wolak, 2000). In the younger sample, convergent validity was reported with measures of substance use, $r = .19$ (Brener, Collins, Kann, Warren, & Williams, 1995), aggressive behavior, $r = .44$ (Udry, 1995), and negative parent–child relationships, $r = .42$, although it should be noted that scale stability and significance levels of these relationships were not reported by the authors. Similar to what has been reported in the adult literature, higher depression scores for females versus males were reported for all age groups in both adolescent samples.

**Consistency of the CES-D and CESD-R With DSM-5 and ICD-10-CM Diagnostic Criteria**

Like the Beck measures, the original version of the CES-D was not developed for use as a clinical diagnostic measure, and in fact, was not developed for clinical use at all. As noted earlier, the CES-D was developed to serve as an epidemiological measure of depressive symptoms in nonclinical populations, and as such suffers from more diagnostic incongruencies with a DSM-5 diagnosis of MDD than noted previously for the BDI measures.

As outlined in Table 9.6, while many of the CES-D items appear to cover most of the DSM-5 “A” criteria, there is a preponderance of items assessing depressed mood, only one item each assessing anhedonia and suicidal ideation, and no items assessing psychomotor retardation or agitation. Whereas the 1-week time frame assessed by the CES-D precludes a DSM-5 diagnosis of MDD, the CESD-R was developed to specifically incorporate all of the DSM “A” criteria, including the requirement that either dysphoria or anhedonia be present, and the 2-week time frame of symptoms. However, like the BDI measures, neither the CES-D nor the CESD-R assesses DSM criteria “B” through “E,” including that the symptoms represent a change from previous functioning; that the symptoms cause clinically significant distress or impairment in social, occupational, or other areas of functioning; that other contributing or comorbid factors such as substance use, general medical conditions, or other psychiatric conditions be ruled out; or that the symptoms are not better accounted for by bereavement or the physiological effects of a medication or other medical condition. None of the severity and course specifiers listed for MDD are assessed by either version, including discriminations between mild, moderate, severe, with psychotic features, single, recurrent or unspecified episodes; whether the depressive episode is in full or partial remission; or whether the current episode occurs with anxious distress, mixed features, melancholic features, atypical features, mood-congruent psychotic features, mood-incongruent psychotic features, catatonia, peripartum onset, or a seasonal pattern.
Table 9.6  CES-D/CESD-R Consistency With DSM-5 Criteria for Major Depressive Disorder

<table>
<thead>
<tr>
<th>DSM-5 Criteria</th>
<th>CES-D Items (item #)</th>
<th>CESD-R Items (item #)</th>
</tr>
</thead>
<tbody>
<tr>
<td>depressed mood</td>
<td>I felt depressed; I could not shake the blues; I was happy; I had crying spells; I felt sad (6; 3; 12; 17; 18)</td>
<td>I could not shake off the blues; I felt depressed; I felt sad (2; 4; 6)</td>
</tr>
<tr>
<td>anhedonia</td>
<td>I enjoyed life (16)</td>
<td>Nothing made me happy; I lost interest in my usual activities (8; 10)</td>
</tr>
<tr>
<td>weight or appetite loss/gain</td>
<td>I did not feel like eating, my appetite was poor (2)</td>
<td>My appetite was poor; I lost weight (1; 18)</td>
</tr>
<tr>
<td>insomnia or hypersomnia</td>
<td>My sleep was restless (11)</td>
<td>My sleep was restless; I slept much more than usual; I had a lot of trouble getting to sleep (5; 11; 19)</td>
</tr>
<tr>
<td>psychomotor retardation or agitation</td>
<td>none</td>
<td>I felt like I was moving too slowly; I felt fidgety (12; 13)</td>
</tr>
<tr>
<td>fatigue</td>
<td>I felt that everything I did was an effort; I could not get “going”; (7; 20)</td>
<td>I could not get going; I was tired all the time (7; 16)</td>
</tr>
<tr>
<td>feelings of worthlessness or guilt</td>
<td>I felt that I was as good as others; I thought my life had been a failure (4; 9)</td>
<td>I felt like a bad person; I did not like myself (9; 17)</td>
</tr>
<tr>
<td>concentration difficulties</td>
<td>I had trouble keeping my mind on what I was doing (5)</td>
<td>I had trouble keeping my mind on what I was doing; I could not focus on the important things (3; 20)</td>
</tr>
<tr>
<td>suicidal ideation</td>
<td>I felt hopeful about the future (8)</td>
<td>I wished I were dead; I wanted to hurt myself (14; 15)</td>
</tr>
</tbody>
</table>

*Italicics = reverse scored item*

Finally, neither the CES-D nor the CESD-R is able to discriminate clinically between the various DSM-5 depressive disorders and MDD. Therefore, like the Beck scales, the use of supplemental measures, diagnostic interviews, and/or clinical judgment is needed in order to make a definitive diagnosis of clinical depression when using the CES-D or the CESD-R.

Likewise, incongruences are also present between the CES-D, the CESD-R, and the primary ICD-10-CM diagnostic criteria for a depressive episode (see Table 9.7). For example, some of the diagnostic symptoms of depression noted in the ICD-10-CM are either not present or are analogously absent in the CES-D and CESD-R, such as “reduced self-confidence,” “depression worse in the morning,” and “loss of libido” for both measures and “psychomotor retardation” for the CES-D. As also noted earlier, the ICD-10-CM severity modifiers and other diagnostic categories are also not well represented in the CES-D or the CESD-R, as well as several older terms describing depressive symptom clusters, such as the presence of “psychogenic,” “reactive,” or “vital” depression. Excluded from both measures are the ICD-10-CM specific definitions and diagnostic categories of mild, moderate, severe without psychotic features, severe with psychotic features, atypical, “masked,” in remission, dysthymia, depression not otherwise specified, and recurrent depression. Hence, the use of supplemental measures, diagnostic interviews, and/or clinical judgment is needed in order to make a diagnosis of clinical depression when using either the CES-D or the CESD-R within the ICD-10-CM classification scheme.
The Patient Health Questionnaires (PHQ) have been developed more recently as self-report scales based on the various modules of the Primary Care Evaluation of Mental Disorders (PRIME-MD; Spitzer et al., 1994) by Kurt Kroenke, MD, and his colleagues with the Patient Health Questionnaire Primary Care Study Group (Kroenke et al., 2001). Based on their widespread availability, defined correspondence to DSM diagnostic criteria, and adoption as recommended scales in recent clinical practice guidelines, two PHQ depression scales have become increasingly popular for both clinical and epidemiological applications in recent years.

### The Patient Health Questionnaire-9

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item patient self-report measure adapted from the mood module of the Primary Care Evaluation of Mental Disorders (PRIME-MD; Spitzer et al., 1994),
Table 9.8 The Patient Health Questionnaire-9 (PHQ-9)

Over the last 2 weeks, how often have you been bothered by any of the following problems?

<table>
<thead>
<tr>
<th>Item</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Not at all</td>
</tr>
<tr>
<td>1</td>
<td>Several days</td>
</tr>
<tr>
<td>2</td>
<td>More than half the days</td>
</tr>
<tr>
<td>3</td>
<td>Nearly every day</td>
</tr>
</tbody>
</table>

1. Little interest or pleasure in doing things
2. Feeling down, depressed, or hopeless
3. Trouble falling or staying asleep, or sleeping too much
4. Feeling tired or having little energy
5. Poor appetite or overeating
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down

(Continued)
Table 9.8 (Continued)

<table>
<thead>
<tr>
<th>No.</th>
<th>Item</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td>Trouble concentrating on things, such as reading the newspaper or watching television</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Moving or speaking so slowly that other people could have noticed? Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Thoughts that you would be better off dead or of hurting yourself in some way</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

<table>
<thead>
<tr>
<th>Not difficult at all</th>
<th>Somewhat difficult</th>
<th>Very difficult</th>
<th>Extremely difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

**Patient Health Questionnaire-2**

The PHQ-2, described as an “ultra-brief” screener, was published in 2003 by Kroenke, Spitzer, and Williams, and consists of the first two items of the PHQ-9, which correlate to the DSM “required” diagnostic symptoms of depression and anhedonia. This measure was published with the intent of serving as a depression screening instrument to assist healthcare providers in identifying patients for whom further assessment may be warranted. Similar to the other versions of the PHQ, the two item category responses are rated on a numerical scale of 0–3 to indicate frequency of endorsement, for a range of scores between 0 and 6. The authors propose a cut score of 3, which indicates a frequency of “Nearly every day” for either of the items, and recommend patients scoring above this threshold be administered the full PHQ-9 (Kroenke et al., 2003).

**Psychometric Properties and Diagnostic Efficiency of the PHQ-9 and PHQ-2**

**PHQ-9**

Initial psychometric properties reported by the authors (Kroenke et al., 2001) were derived by administering the more extensive version of the PHQ scale to two large outpatient samples of primary care patients (N = 1,422 from five general internal medicine clinics; N = 1,578 from three family medicine clinics) and a sample of obstetrics/gynecological patients (N = 3,000 from seven clinics). Characteristics of the primary care sample indicated that respondents were middle-aged (M = 46 ± 17), predominately Caucasian (79%) females (66%), who had at least some college education (73%), and 25% of whom had some comorbid medical condition. Characteristics of the OB/GYN sample indicated that respondents were all female, younger (M = 46 ± 17), included equal numbers of Caucasians and Hispanics (39%), and had at least some college education (84%). Internal consistency reliabilities were reported to be .89 in the primary care sample and .86 in the OB/GYN sample, and 48-hour test-retest reliability was .84 for the combined sample. Pragmatically selected cut scores of 5, 10, 15, and 20 were used to classify “mild,” “moderate,” “moderately severe,” and “severe” depression, and score distributions within those classifications were 19.6%, 4.9%, 1.7%, and 0.4%, respectively. ROC analysis indicated that sensitivity, specificity, and positive likelihood ratios on a subsample of 580 respondents who were determined via a telephone-administered version of the PRIME-MD and additional SCID items 48 hours later to have a
diagnosis of MDD were $Sn = .95; Sp = .84; LR + 6.0$ for a cut score of 9, and $Sn = .88; Sp = .88; LR + 7.1$ for a cut score of 10.

Significant positive correlations were demonstrated between PHQ-9 total scores and the Mental Health, $r = .73$, Medical Health Perceptions, $r = .55$, and Social Functioning, $r = .52$, subscales of the Medical Outcomes Study Short-Form General Health Survey (SF-20; Stewart, Hays, & Ware, 1988), as well as between PHQ-9 total scores and self-assessed symptom-related difficulty in activities and relationships, $r = .55$; self-reported estimates of disability days in the past three months, $r = .39$; and self-reported estimates of physician visits in the past three months, $r = .24$. No relationships for various age and racial subgroups were reported (Kroenke et al., 2001).

More recently, Beard, Hsu, Rifkin, Busch, and Bjorgvinsson (2016) validated the PHQ-9 in a psychiatric sample of 1,023 patients attending a partial-day hospital treatment program. In this sample, the authors reported strong positive convergent validity between the PHQ-9 and the CESD-10, $r = .80$, and higher scores were reported in the portion of the sample diagnosed as having a current major depressive episode versus those not, $t(847) = 21.93, p \leq .001$. Divergent validity was demonstrated by negative correlations with the Schwartz Outcome Scale, $r = -.65$ (Blais et al., 1999); however, divergent validity was not demonstrated between the PHQ-9 and the GAD-7, $r = .61$ (Spitzer et al., 2010). In addition, an ROC analysis compared sensitivity, specificity, and positive and negative predictive values for various cut scores, concluding that in this psychiatric sample, a cut score of 13 demonstrated the best diagnostic efficiency ($Sn = .83; Sp = .72; PPV = .82; NPV = .74$) as opposed to the recommended cut score of 10 ($Sn = .93; Sp = .52; PPV = .74; NPV = .84$).

Norms have been derived from administering the PHQ-9 to two samples of community-dwelling adults in Germany ($N = 2,500$ and $2,518$), with sex, age, and educational level hypothesized to predict higher scores (Kocalevent, Hinz, & Braher, 2013). In this study, good internal consistency reliability ($\alpha = .87$) was reported, and both exploratory and confirmatory factor analyses indicated that a unidimensional factor structure best described the data, with all nine items having factor loadings between .65 and .77. Unlike the findings reported in the initial validation studies, the PHQ-9 demonstrated significant negative correlations with the Mental Health, $r = -.61$, and Social Functioning, $r = -.60$ subscales, as well as the Mental Component Summary measure, $r = -.68$, of the Medical Outcomes Study Short-Form-12 (SF-12) (Ware, Kosinski, & Keller, 1996). The overall prevalence rate for moderate to high depression severity scores was 5.6%, and as predicted, higher PHQ-9 scores were found in female, elderly, and unemployed respondents, as well those living alone and those with lower income and educational levels. For the total population of both sexes and all age groups, PHQ-9 scores of 5 and 10 fell at percentile 79.3 and 95.1, respectively. Table 9.9 lists raw total PHQ-9 scores that fall at approximately the 80th and 95th percentiles for sex and age bands between 14 and $\geq 75$.

As noted earlier, Kroenke and colleagues (2001) originally reported pragmatically selected cut scores of 5, 10, 15, and 20 for mild, moderate, moderately severe, and severe depression. In their

<table>
<thead>
<tr>
<th>Table 9.9 PHQ-9 Norms for Sex and Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

recent meta-analytic review of the diagnostic utility of using this algorithmic scoring method, Manea, Gilbody, and MacMillan (2015) found that using these cut-points produced unacceptably low sensitivity in diagnosing MDD. While they report high heterogeneity between the studies reviewed, diagnostic efficiency for 27 studies using the recommended scoring algorithm conducted in either primary care, hospital-based, or medical subspecialty clinics indicated that pooled sensitivity was .58 (CI .50–.66), pooled specificity was .94 (CI .92–.96), and the diagnostic odds ratio was 24.92 (CI 16.73–37.12). However, in 13 studies that compared both methods, sensitivity was increased by using the total summed-items scoring method and a cut-point of 10 or greater to make a diagnosis with pooled sensitivity of .77 (CI .66–.85), pooled specificity of .85 (CI .79–.90), and diagnostic odds ratio of 21.53 (15.68–29.58). One explanation posited by the authors for the low sensitivity of the algorithmic scoring method could be that some respondents find the middle score options provided by the PHQ-9 to be confusing, and that responses coded as “1 = Several days” and “2 = More than half the days” are not always discriminated well from one another.

A contemporaneous meta-analysis by the same group (Moriarty, Gilbody, McMillan, & Manea, 2015) reviewed studies utilizing a range of total summed-item cut scores between 7 and 15 in studies conducted in primary care, hospital-based settings, and community clinics, and identified that the best balance of sensitivity and specificity for the studies reviewed was a cut score of 11 or greater, pooled sensitivity of .83 (CI .71–.90), pooled specificity of .89 (CI .83–.93), and diagnostic odds ratio 43.11 (CI 21.63–85.91). The authors also report that sensitivity values across the range of cut scores reviewed showed an unusual trend, in that cut scores of 8, 9, and 11 demonstrated comparable sensitivity values, whereas a cut score of 10 produced a drop in sensitivity value. Likewise, cut scores of 12 and 13 demonstrated comparable sensitivity values, with two more drops occurring at cut scores of both 14 and 15. Therefore, while the authors again note significant heterogeneity between the studies reviewed, they report that their data indicates that (1) the sensitivity of the PHQ-9 is lower than that reported in the original validation studies, (2) it is currently unclear how the PHQ-9 performs at different cut-points, and (3) it may perform differently in different populations. Similarly, a review of the norms reported by Kocalevent and colleagues (2013; see Table 9.9) echo this interpretation, and suggest that cut scores of 9 may be more appropriate in younger populations, and cut scores of 10 or 11 may be more appropriate in those aged 45 and older.

**PHQ-2**

Initial psychometric properties for the PHQ-2 were derived using the same validation samples as those used for the PHQ-9 (Kroenke et al., 2003). ROC analysis indicated that a cut score of 3 provided the best balance of sensitivity and specificity, with values in the subsample of 580 respondents who were determined 48 hours later to have a diagnosis of MDD reported as Sn = .83; Sp = .90; LR + 2.9. Positive correlations were reported separately for the primary care and OB/GYN samples between PHQ-2 total scores and the Mental Health (r = .70 and .63, respectively), Medical Health Perceptions (r = .47 and .46, respectively), and Social Functioning (r = .46 and .36, respectively) subscales of the Medical Outcomes Study Short-Form General Health Survey (Stewart et al., 1988).

A cut score of 3 was also reported in a follow-up validation study (Lowe, Kroenke, & Gräfe, 2005) of 1,619 medical outpatients and a subsample of 520 SCID diagnosed patients who were evaluated one year later (Sn = .87; Sp = .78; LR + 1.3). This study also reported significant convergent validity with the Hospital Anxiety and Depression Scale, r = .67 (Zigmond & Snaith, 1983). However, unlike the earlier study, the PHQ-2 demonstrated significant negative correlations with the Mental Health, r = −.71, and Physical Health, r = −.23, Component Summary measures of the SF-12 (Ware et al., 1996), as well as with the World Health Organization Five-Item Well-Being Index, r = −.69
(World Health Organization, 1988). In addition, this study reported sensitivity to change by comparing improvements in SCID diagnostic status to change in total PHQ-2 scores, and reported that mean PHQ-2 scores changed by +2.3 points, +0.4 points and −1.3 points in the “improved,” “unchanged,” and “deteriorated” subgroups, respectively (Lowe et al., 2005).

In their recent meta-analysis, Manea and colleagues (2016) reviewed 36 studies that compared the performance of the PHQ-2 to a gold standard criterion reference for diagnosis of MDD. The studies reviewed were conducted in a variety of primary care, outpatient, and hospital-based specialty care clinics, with a few also stemming from community-based health services or clinics. Nineteen studies reported the standard cut score of 3 or greater, and for those studies the pooled sensitivity was .76 (CI .68–.82), the pooled specificity was .87 (CI .82–.90), and the pooled diagnostic odds ratio was 22.20 (CI 14.00–35.19). For 17 studies that reported a cut score of 2 or greater, pooled sensitivity was .91 (CI .85–.94) and pooled specificity was .70 (CI .64–.76). The authors of this study conclude that in populations with low base rates of depression, the higher sensitivity afforded by a cut score of 2 may be preferable, in order to decrease the risk of cases being missed.

Consistency of the PHQ-9 and PHQ-2 With DSM-5 and ICD-10 Diagnostic Criteria

While the Beck scales, the CES-D measures, and the PHQ-2 were not developed to be diagnostic measures of clinical depressive disorders, and therefore do not incorporate all criteria necessary to make a diagnosis of depression according to either the DSM-5 or ICD-10-CM criteria, the PHQ-9 was specifically developed as a diagnostic measure of DSM criteria for MDD (Hahn et al., 2004). However, it is interesting to note that the scoring algorithms provided in the Instructional Manual for the PHQ scales concede that the PHQ-9 cannot be used as a stand-alone measure to diagnose MDD. Rather these algorithms suggest diagnoses of “Major Depressive Syndrome” and “Other Depressive Syndrome,” diagnostic categories that do not exist in either the DSM or the ICD classification schemes. Rather, it is stated that

the diagnoses of Major Depressive Disorder and Other Depressive Disorder requires ruling out normal bereavement (mild symptoms, duration less than 2 months), a history of a manic episode (Bipolar Disorder) and a physical disorder, medication or other drug as the biological cause of the depressive symptoms.

(Spitzer et al., 2010, p. 4)

As reflected in Table 9.10, the PHQ-9 items accurately assess all of the individual “A” criteria, the “B” criteria, and the 2-week time frame needed for a diagnosis of MDD; however, many of the other DSM-5 criteria (“C” through “E”) are not covered by the PHQ-9. While the PHQ-9 is the only measure reviewed to assess criteria “B,” that the symptoms cause “clinically significant distress or impairment in social, occupational, or other important areas of functioning”; criteria not assessed by the scale include the requirements that the symptoms represent a change from previous functioning; that other contributing or comorbid factors, such as substance use, general medical conditions or other psychiatric conditions, be ruled out; or that the symptoms are not better accounted for by bereavement or the physiological effects of a medication or other medical condition. Likewise, similar to the Beck scales and CES-D measures previously reviewed, the majority of the severity and course specifiers listed for MDD are also not assessed by the PHQ-9, including discriminations between single, recurrent, or unspecified episodes; whether the depressive episode is in full or partial remission; or
whether the current episode occurs with anxious distress, mixed features, melancholic features, atypical features, mood-congruent psychotic features, mood-incongruent psychotic features, catatonia, peripartum onset, or a seasonal pattern. Finally, by definition as a screening device, the PHQ-2 does not include items that assess any but the two cardinal “A” DSM-5 criteria (see Table 9.10), and neither PHQ measure discriminates clinically between MDD and other DSM-5 depressive disorders. Therefore, the use of supplemental measures, diagnostic interviews, and/or clinical judgment is needed in order to make a definitive diagnosis of clinical depression when using the PHQ-9 and PHQ-2.

Likewise, in a similar manner to the incongruences previously reviewed between the Beck and CES-D scales and the primary ICD-10-CM diagnostic criteria for a depressive episode, similar discrepancies exist for the PHQ measures (see Table 9.11). As noted previously, some of the diagnostic symptoms of depression noted in the ICD-10-CM are either not present, or mentioned

<table>
<thead>
<tr>
<th>ICD-10 Criteria</th>
<th>PHQ-9 Items (item #)</th>
<th>PHQ-2 Items (item #)</th>
<th>PHQ-2 Items (item #)</th>
</tr>
</thead>
<tbody>
<tr>
<td>depressed mood</td>
<td>Feeling down, depressed, or hopeless (2)</td>
<td>Feeling down, depressed, or hopeless (2)</td>
<td>Feeling down, depressed, or hopeless (2)</td>
</tr>
<tr>
<td>anhedonia</td>
<td>Little interest or pleasure in doing things (1)</td>
<td>Little interest or pleasure in doing things (1)</td>
<td>Little interest or pleasure in doing things (1)</td>
</tr>
<tr>
<td>weight or appetite loss/gain</td>
<td>Poor appetite or overeating (5)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>insomnia or hypersomnia</td>
<td>Trouble falling or staying asleep, or sleeping too much (3)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>psychomotor retardation or</td>
<td>Moving or speaking slowly, or being fidgety or restless (8)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>agitation</td>
<td>Feeling tired or having little energy (4)</td>
<td>Feeling bad about yourself (6)</td>
<td>Feeling bad about yourself (6)</td>
</tr>
<tr>
<td>fatigue</td>
<td>Trouble concentrating on things (7)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>feelings of worthlessness or</td>
<td>Thoughts that you would be better off dead or of hurting yourself (9)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>guilt; concentration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>difficulties</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>suicidal ideation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Table 9.11 PHQ-9/PHQ-2 Consistency With ICD-10-CM Criteria for Depressive Episode |
|------------------------------|----------------------|----------------------|
| depressed mood               | Feeling down, depressed, or hopeless (2) | Feeling down, depressed, or hopeless (2) |
| reduction in energy; decrease in activity anhedonia | Feeling tired or having little energy (4) | Feeling tired or having little energy (4) |
| anhedonia                    | Little interest or pleasure in doing things (1) | Little interest or pleasure in doing things (1) | Little interest or pleasure in doing things (1) |
| reduced concentration        | Trouble concentrating on things (7) | None | None |
| tiredness                    | Feeling tired or having little energy (4) | None | None |
| sleep disturbance            | Trouble falling or staying asleep, or sleeping too much (3) | None | None |
| reduced appetite or weight loss | Poor appetite or overeating (5) | None | None |
| reduced self-esteem          | Feeling bad about yourself (6) | None | None |
| reduced self-confidence      | Feeling that you are a failure (6) | None | None |
| guilt; worthlessness          | Feeling bad about yourself (6) | None | None |
| depression worse in morning  | None | None | None |
| psychomotor retardation      | Moving or speaking slowly (8) | None | None |
| agitation                    | Being fidgety or restless (8) | None | None |
| loss of libido                | None | None | None |
only in the specifiers of the DSM-5. Therefore, while a few of ICD-10-CM depression criteria are analogously absent in the PHQ measures, such as “depression worse in the morning” and “loss of libido,” others are included that are not represented in the DSM-5 criteria, such as “reduced self-esteem” and “reduced self-confidence,” although these constructs can be extrapolated from item 6 of the PHQ-9.

The ICD-10-CM severity modifiers and other diagnostic categories are also not as well represented in the PHQ measures, in addition to several older terms describing depressive symptom clusters, such as the presence of “psychogenic,” “reactive,” or “vital” depression. These include the ICD-10-CM specific definitions and diagnostic categories of mild, moderate, severe without psychotic features, severe with psychotic features, atypical, “masked,” in remission, dysthymia, depression not otherwise specified, and recurrent depression. As a result, the use of supplemental measures, diagnostic interviews, and/or clinical judgment is needed in order to make a diagnosis of clinical depression when using the PHQ-9 and PHQ-2 within the ICD-10-CM classification scheme.

IMPLEMENTATION ISSUES, SCREENING, AND TREATMENT DECISIONS BASED ON THE BDI, CES-D, AND PHQ SCALES

As noted previously, the BDI measures were originally developed to describe the presence and severity of depressive symptoms in clinical psychiatric populations, and the original CES-D was developed as an epidemiological measure for use in large studies of normal populations. The PHQ-9 was developed to specifically diagnose MDD using DSM criteria, and the PHQ-2 was published as a short screener for the same diagnosis. Whereas the Beck, CES-D, and PHQ-2 measures are intended to characterize clusters of symptoms consistent with depression, but not to diagnose or differentiate these symptoms from clinical depressive disorders or other possible etiologies of those symptoms, the stated intent of the PHQ-9 is diagnosis. However, as reviewed earlier, high scores on all of these measures may be present in patients who do not have a clinical depressive disorder, who have other psychiatric or medical disorders, as the result of medications or substances, or as the result of transient stressors. Accordingly, clinicians must be aware that the literature does not provide evidence that high scores on these scales suggest any particular etiology of a patient’s symptoms, and also no particular differential diagnosis, prognosis, suggested treatment, or predicted response to treatment.

Several authors have found that elevated scores on the CES-D are just as likely to be associated with anxiety disorders as depressive disorders, leading to the suggestion that the CES-D may more accurately be assessing general psychological distress rather than clinical depression (Breslau, 1985; Fechner-Bates, Coyne, & Schwenk, 1994; Kirmayer, Robbins, Dworkind, & Yaffe, 1993; Roberts & Vernon, 1983). Similarly, many of the current measures reviewed earlier also failed to report adequate discriminant validity with measures of anxiety. While some overlap in depression and anxiety symptoms is to be expected, and is reflected in the new DSM-5 modifier of MDD “with Anxious Distress,” the very high positive correlations observed in the BDI-FS (r = .86 and .60), the CESD-R (r = .73 and .65), the CESD-10 (r = .72), and the PHQ-9 (r = .61) scales suggest that these depression scales and the referent anxiety measures may be assessing the same underlying construct. Therefore, these measures cannot be expected to differentiate between the various diagnostic entities of depression and anxiety, and the use of these scales in primary care must be supplemented with additional relevant information based on each patient’s unique presentation and history.

While the BDI, CES-D, and PHQ scales should not serve as stand-alone diagnostic scales or treatment outcome measures in either normal or clinical populations, their brevity, ease of use, and generally good psychometric properties have nonetheless led to their widespread utilization for all of
these purposes in both primary and specialty care. The high prevalence of depressive symptoms in primary care coupled with the poorer health outcomes and higher healthcare utilization in patients whose depression goes unrecognized have fueled a demand for instruments that can be administered easily in clinical settings to a wide variety of populations. In addition, newer requirements for evidence-based practices, quantifiable assessment data, and objective measures of treatment efficacy have created a need for valid and reliable self-report instruments, particularly in primary care where both mental health professionals and more comprehensive diagnostic measures may be unavailable.

All versions of the BDI, CES-D, and PHQ scales described in this chapter can be quickly and easily administered by nursing or ancillary healthcare staff in 15 minutes or less, have reading levels of sixth grade or less, and can be scored and interpreted by physicians or other treating healthcare providers with proper training in the assessment of suicide risk, the availability of treatment options for depression, and knowledge of appropriate referral mechanisms for depressive disorders.

It should be noted that in their administration, particular attention must be paid to Item #7 on the BDI-FS, and Items #9 on both the BDI-II and PHQ-9: “Suicidal Thoughts or Wishes” and “Thoughts that you would be better off dead or of hurting yourself in some way.” If any of the scored item choices is endorsed (i.e., “1 = I have thoughts of killing myself, but I would not carry them out”; “2 = I would like to kill myself”; “3 = I would kill myself if I had the chance”; “1 = Several days”; “2 = More than half the days”; “3 = Nearly every day”), a comprehensive suicide risk assessment should be performed, and referral to a mental health professional or psychiatric emergency room should be strongly considered. Assessment of suicide risk is addressed in Chapter 11 of this book.

In addition, in settings where mental healthcare providers and more complex diagnostic tools are available, it is also recommended that primary care providers refer patients scoring above an applicable cut score for a more complete multimethod assessment of depressive symptomatology, which would include readministration of the screening measure(s) along with a validated clinical diagnostic interview and comprehensive assessments of medical and psychiatric history and symptoms (Kendall, Hollon, Beck, Hammen, & Ingram, 1987).

With regard to selection of specific cut scores, clinicians need to consider both their specific patient population and the purpose of screening. Higher cut scores may be employed if the patient population being assessed has a higher baseline rate of depression, or the clinician is interested in identifying individuals more likely to meet full diagnostic criteria for a depressive disorder. Lower cut scores may be employed if the goal is to identify all possible cases of depression. Therefore, cut scores on the BDI, CES-D, and PHQ measures may be altered to adjust the balance between sensitivity (i.e., detection of true positives) and specificity (i.e., detection of true negatives). Based on the ROC analyses, norms, and cut scores reviewed earlier, recommended cut scores for determining further evaluation for normal, primary care, or lower base rate (↓ rate score) populations, and psychiatric, chronic disease, elderly, or higher base rate (↑ rate score) populations for each scale are provided in Table 9.12.

Recommended scale versions of each measure previously reviewed are also indicated in bold in Table 9.12. Scale versions recommended for use as screening instruments in primary care are the BDI-II, the original CES-D, the 10-item CESD-10, and the PHQ-9. These versions were selected based on methodological strength of studies reviewed, the availability of updated psychometric characteristics, and a large recent literature with evidence of stability in a wide variety of populations; however clinicians utilizing these scales and cut scores may wish to consult referent psychometric analyses or norm tables based on the profile of their particular patient population.

While the CESD-R and CESDR-10 may have promise as alternative measures, there is a dearth of published studies utilizing a single, standard version of either scale, with both consistent items and a consistent response set, thus making it difficult to adequately evaluate these measures’ psychometric properties. Similarly, there are only two published studies on the CESD-R, and only one published
study on the CESDR-10, all of which utilized limited demographic samples, making it difficult to assess either measures’ suitability for primary care, or generalizability to diverse populations. Likewise, the PHQ-2 is not recommended because the PHQ-9 is just as accessible, takes negligible additional time to administer and interpret, and because it adds no new information to the diagnostic picture that is not already provided by the PHQ-9. The authors of the PHQ-2 recommend administering the PHQ-2 first, and then following up with a PHQ-9 if the applicable cut score is reached. However, at least one study recently indicated that primary care physicians rarely follow up with a full PHQ-9 (Fuchs et al., 2015). Given the time constraints of busy primary care providers, simply administering the PHQ-9 alone would be more efficient.

With regard to treatment monitoring, the BDI-II has been reported to be sensitive to change in prior studies of various treatment modalities, including pharmacotherapy, psychotherapy, and electroconvulsive therapy (Katz, Katz, & Shaw, 1999), and several studies have demonstrated that the BDI-II is sensitive to measures of clinically significant change in both symptomatic and asymptomatic samples when evaluated by reliable change indices, item response theory methods, general linear modeling, and traditional ROC analyses (Brouwer, Meijer, & Zevalkink, 2013; Button et al., 2015; Seggar, Lambert, & Hansen, 2002). In these studies, BDI-II score changes of 3 or 4 points were clinically significant in low or asymptomatic individuals, whereas score changes of 14 to 17 points were needed to demonstrate clinically significant change in highly symptomatic individuals.

Based on the initial validation studies of the PHQ-9, the authors propose that a posttreatment score of 9 or greater, along with the commonly used criterion of a 50% reduction from the baseline total score, can be used to define clinically significant improvement with this scale (Kroenke et al., 2001). However, concerns raised above about the PHQ-9’s diagnostic sensitivity, recommended algorithmic scoring classification scheme, performance characteristics in differing populations, and stability in discriminating between various levels of depression severity suggest that caution should be applied when utilizing the PHQ-9 as a stand-alone treatment monitoring measure. Rather, clinicians may wish to employ supplemental measures or definitions of improvement, such as a reliable change index score, or minimal clinically important difference criteria, which suggest a posttreatment change of 9 or greater and a reduction in PHQ-9 total scores of 5 or greater (Löwe, Kroenke, Herzog, & Gräfe, 2004; McMillan, Gilbody, & Richards, 2010).

| Table 9.12 Recommended Versions and Cut Scores for the BDI, CES-D, and PHQ Scales |
|-------------------------------------|-----|-----|
| Scale Version:                      | ↓ Rate Score | ↑ Rate Score |
| Original Beck Depression Inventory—BDI | 10  | 19  |
| BDI-Short Form                       | ND  | ND  |
| BDI-IA                              | 10  | 19  |
| **BDI-II**                           | **14** | **20** |
| BDI-FS (BDI-PC)                     | 4   | 6   |
| **Center for Epidemiological Studies Depression Scale—CES-D** | **20** | **25** |
| CESD-8                              | ND  | ND  |
| **CESD-10**                          | **10** | **15** |
| Center for Epidemiological Studies Depression Scale-Revised—CESD-R | 18  | 24  |
| CESDR-10                            | ND  | ND  |
| **Patient Health Questionnaire-9—PHQ-9** | **9**  | **11** |
| Patient Health Questionnaire-2—PHQ-2 | 2   | 3   |

ND = no data available  Bold = recommended scale versions
↓ rate score = cut score for normal, primary care, or lower base rate populations
↑ rate score = cut score for psychiatric, chronic disease, elderly, or higher base rate populations
While Radloff (1977) reported the CES-D to be sensitive to change in one sample of psychiatric inpatients, as demonstrated by 10-point reductions in total score after both 1 week and 4 weeks, there are no recent comprehensive studies investigating the use of any of the CES-D scales for treatment monitoring. Therefore, it is not currently recommended that the CES-D or the CESD-R be used to evaluate treatment outcomes or efficacy.

**CASE STUDY**

Ms. D is a 25-year-old single female with no children who presented to an outpatient clinic for treatment of depression. A 3-year relationship had recently ended, she was having difficulty on her job as an insurance salesperson, and she was experiencing financial trouble. Ms. D presented with depressed mood, sleep difficulties, and suicidal ideation.

Ms. D had no prior history of depression and reported that her mother had experienced episodes of depression throughout her adult life. She described her father as “hard working” and “strict” and reported that he was somewhat removed from the family. Ms. D had done well in school, had graduated high school, and completed a bachelor’s degree in marketing at the age of 23. Since that time she had been working as an insurance agent, which she reported was stressful due to the commission-based salary, long hours, and difficulties with her current supervisor. Ms. D had been in a serious relationship for the past three years, which ended unexpectedly 2 months ago when her significant other abruptly terminated the relationship. Ms. D reported that she was very distraught by this breakup, as she had anticipated that this relationship would eventually result in marriage. Ms. D reported onset of depressive symptoms at the time of the breakup and noted that symptoms have increased in intensity due to compounded work and financial stressors.

Ms. D’s BDI-II score upon initial interview was 33, indicating that she was severely depressed. She endorsed statements rated as 2 or 3 for the following symptoms: sadness, loss of pleasure, self-dislike, suicidal thoughts, crying, worthlessness, and changes in sleeping pattern.

The clinician took note of Ms. D’s response to the item regarding suicidal thoughts and conducted a complete suicide assessment during the first interview. Ms. D reported that she had been experiencing vague thoughts of “not wanting to be here anymore,” without a current plan or the intent for suicide. The clinician discussed options for Ms. D should her suicidal ideation become more intense, such as calling a local suicide hotline, phoning the psychological center, or presenting to the emergency room if necessary. Ms. D contracted that she would not harm herself and would take any or all of these actions if necessary to prevent harming herself.

Item analysis of the BDI also alerted the clinician to Ms. D’s experience of anhedonia, which was not a presenting complaint, as well as her negative self-evaluation. The clinician was able to address anhedonia with Ms. D and proceed with behavioral treatment of this particular depressive symptom. In addition, examination of her responses on the BDI to the self-dislike and worthlessness questions assisted the clinician in introducing pertinent examples of cognitive restructuring. Ms. D’s BDI score varied between 25 and 38 during the first five weeks of therapy. By the tenth week of therapy her score was 17, and between weeks 12 and 14 her score was stable at 9. By this time, Ms. D and her therapist agreed that treatment could be terminated, allowing for follow-up sessions as needed.

This case demonstrates how the BDI-II can be used in treatment planning by alerting the clinician to severe symptomatology (e.g., suicidal ideation) and to salient cognitive and affective depressive symptoms. The BDI-II can be an excellent point for initiating cognitive therapy, as patients may be more apt to endorse depressive self-statements rather than initiate them on their own. The BDI-II was also useful in this case with regard to treatment monitoring and assisting in the determination of treatment termination.
SUMMARY

The BDI, CES-D, and PHQ-2 scales, while inappropriate for use as measures for the primary diagnosis and treatment planning of clinical depressive disorders, can be utilized as screening tools to identify patients who may need further evaluation of their depressive symptoms. This can be particularly important in primary care settings, where high prevalence rates of depression have been observed, and nonrecognition of symptoms is associated with both poorer health outcomes and higher healthcare utilization. The PHQ-9, while covering many of the DSM-5 inclusion criteria for MDD, also cannot serve as a stand-alone instrument for either diagnosis or treatment planning.

Primary strengths of these measures include their brevity, ease of use, low cost and time burden of administration, applicability and acceptability to wide patient populations, and generally good psychometric characteristics. Primary limitations of these instruments include their inability to provide a definitive diagnosis of clinical depression under either the DSM-5 or the ICD-10-CM classification schemes, their inability to differentiate between various medical or clinical depressive and anxiety disorders, and in the case of the CES-D scales, their lack of available evidence supporting use for monitoring treatment efficacy or outcomes.

Based on the recent literature evaluating psychometric properties and diagnostic efficiency, scales recommended for use as screening instruments in primary care settings include the BDI-II, CES-D, CESD-10, and PHQ-9. Recommended cut scores for further evaluation for respective low and high prevalence populations for each of these measures are 14 and 20 for the BDI-II, 20 and 25 for the CES-D, 10 and 15 for the CESD-10, and 9 and 11 for the PHQ-9.

Newer requirements for evidence-based practices, quantifiable assessment data, and objective measures of treatment efficacy have created a need for valid and reliable self-report instruments, and for these purposes, the scales reviewed therein can provide primary care clinicians with a tool to rapidly identify patients who warrant further evaluation of their depression symptoms. Finally, the BDI, CES-D, and PHQ scales all provide important information on frequency or severity of depressive symptoms, which may help primary care practitioners make decisions about the need for referral to a mental health provider, or about whether to initiate treatment.

AUTHOR’S NOTE


REFERENCES


The Geriatric Depression Scale (GDS) is a widely used instrument developed in the early 1980s as a measure of geriatric depressive symptomatology. The GDS has several features that make it a good option for screening of and treatment planning for depressive symptoms in older adults. These features include brevity, strong psychometric qualities, and applicability for a wide range of older adults. An extensive literature on the GDS exists and much of it will be reviewed in this chapter. From this literature suggestions for its use in treatment planning, treatment monitoring, and treatment outcomes assessment will be developed.

OVERVIEW OF THE INSTRUMENT

While there are many instruments available to measure depression, the GDS is a valid and reliable instrument that was developed specifically for use with older adults by recognizing that depression symptoms are expressed differently across the life span. The use of the GDS or abbreviated versions of the GDS are well suited for the primary care setting, but providers should be mindful about the influence of cultural factors and cognitive function on the utility of the instrument. This section provides additional information about using the GDS as a depression screener, as well as ways to provide feedback regarding findings and using these findings to guide treatment approaches.

Development

The development of the GDS was motivated by the perception that existing self-report measures of late-life depression were lacking (e.g., Jarvik, 1976; Kane & Kane, 1981; Salzman & Shader, 1978). The most frequently expressed concerns were with the inclusion of somatic symptoms of depression, including decreased energy, sleep deficits, and gastrointestinal difficulties, as these tended to be unreliable indicants of depression when evaluated in older adults. The use of somatic items can be problematic, as physical concerns that may be a normal part of aging may be misinterpreted as depressive symptoms (Peach, Koob, & Kraus, 2001).

Other concerns included the confusion often engendered by the multiple-response format of extant measures, especially with older adults experiencing mild to moderate cognitive impairment. Additionally, extant measures lacked norms for older adults. These concerns coincided with the maturation of the field of geriatric mental health and motivated investigators to develop a self-report depression instrument specifically geared for older adults.

The results of initial development efforts of the GDS were detailed in articles published in the early 1980s. A group of investigators at Stanford University and the Palo Alto Veteran’s Administration Medical Center, led by T.L. Brink and Jerome Yesavage, published these initial studies. The first (Brink, Yesavage, Lum, Heersema, Adey, & Rose, 1982) reported in rather truncated fashion the
development of the scale. The second (Yesavage et al., 1983) is a much more detailed version of the same initial development and validation and will be reviewed for the purposes of this chapter.

The development of the GDS began with a team of geriatric mental health specialists selecting 100 items believed to be useful for distinguishing between depressed and nondepressed older adults. An effort was made to include items that covered the range of depressive symptoms, including memory loss, cognitive complaints, somatic complaints, and self-image. A yes/no response format was chosen based on the experiences of the authors that multiple-response formats often confused older adults. The 100 items were administered to 47 male and female older adults (age >55) who were either community-dwelling with no complaints of depression or were hospitalized for depression. The authors used a bootstrapping strategy to select items whereby those items evidencing the best correlation to the total score were retained for further validation. The rationale provided for this strategy was that the 100 items generated would provide the best measure of the geriatric depression construct. A decision to select 30 items was made, presumably, to minimize fatigue effects. The item-total correlations ranged from .47 to .83. Interestingly, the 12 items related to somatic concerns were not selected based on this procedure; that is, the item-total correlations for these items were not in the top 30.

The next step in the development of the GDS involved cross-validation with a new set of participants. The sample was comprised of 40 community-dwelling older adults without depression and 60 older adults in treatment for depression. The depressed sample was divided into mild (n = 26) and severe (n = 34) cases based on the number of endorsed depressive symptoms. Participants were administered the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967), the Zung Self-Rating Scale for Depression (SDS; Zung, 1965), and the GDS in random order.

Validity and Reliability of the GDS

Consistency and reliability estimates for the GDS were impressive. The alpha coefficient was .94, split-half reliability was also .94, and test-retest reliability over a 1-week interval was .85. Validity was examined by using the classifications of nondepressed, mildly depressed, and severely depressed as between-subjects variables, and then comparisons of the scores obtained on the GDS, HRSD, and SDS were made. Scores on each of the measures reliably distinguished the three grades of severity, suggesting discriminate validity. Concurrent validity for the GDS was confirmed by strong correlations with the other two instruments (.84 with the SDS and .83 with the HRSD).

Yesavage and colleagues also suggested a cutoff score of 11 for identifying depression (i.e., 0–10 nondepressed, >11 depressed). This cut score yielded sensitivity (the ability to correctly classify depressed patients) and specificity (the ability to correctly classify “normal” patients) rates of 84% and 95%, respectively.

The results of this initial validation study were encouraging. Although the GDS demonstrated adequate reliability and validity, the methods of this study did not permit the demonstration of superiority to other measures of depression. This article serves as the cornerstone of GDS literature, and follow-up validity studies have mostly supported these original findings. For instance, a more recent systematic review (Wancata, Alexandrowicz, Marquart, Weiss, & Friedrich, 2006) found similar criterion validity rates (sensitivity = .75; specificity = .77) to that reported in the original study. The GDS has also been found to have good known-groups validity in discriminating among the individuals classified as normal (scoring between 0 and 9 on the instrument), mildly depressed (10–19), and severely depressed (20–30; Peach et al., 2001).
Development of Abbreviated Versions of the GDS

To assist with the time demands of clinical settings, researchers have developed many different abbreviated versions of the GDS (e.g., 4-, 5-, 10- and 15-item versions; see Table 10.1). When developing the GDS-Short Form (GDS-15), Sheikh and Yesavage (1986) selected the 15 items on the GDS that best correlated with the total score using data from the original validation study. Cross-validation was conducted on these items with 35 participants (18 nondepressed individuals and 17 individuals being treated for dysthymia or depression). Participants completed both the full GDS and the short form version of the measure, yielding a strong correlation of .84. A more recent meta-analysis of the GDS-15 found a pooled sensitivity of .89 and specificity of .77 at the recommended cutoff score of 5 (Pocklington, Gilbody, Manea, & McMillan, 2016).

Although the GDS-15 has been found to be unidimensional, there has been evidence of item redundancy indicating that an even shorter version would be adequate at assessing late-life depression (Tang, Wong, Chiu, Lum, & Ungvari, 2005). As such, even shorter versions of the GDS have been developed, using 10 items (GDS-10), five items (GDS-5), and four items (GDS-4) based on the most discriminating items of the measure. The recommended cutoff scores for these abbreviated GDS are: 4/5 for the GDS-10, 2/3 for the GDS-5, and 2/3 for the GDS-4 (Almeida & Almeida, 1999; Hoyl et al., 1999). D’Ath, Katona, Mullan, Evans, and Katona (1994) found the GDS-10 and GDS-4 correlations with the GDS-15 to be .97 and .89, respectively. However, D’Ath et al. concluded that the GDS-4 should be considered primarily as a minimal screening procedure, as they found that sensitivity was 93% with a cut score of 1. A separate study (Rinaldi et al., 2003) assessed the validity of the GDS-5 for depression screening in older adults evaluated in three different settings (hospital, outpatient clinic, and nursing home). The GDS-5 had a sensitivity of .94 (95% Confidence Intervals [CI] = .91-.98), specificity of .81 (95% CI = .75-.87), interrater reliability of .88, and test-retest reliability of .84; similar values were obtained in each setting. This study concluded that the GDS-5 is as effective as the GDS-15 for screening for depression in cognitively intact older adults. Thus, there are viable options to the full-scale GDS when brevity is of particular concern.

Table 10.1 Geriatric Depression Scale (Short Form)

Choose the best answer for how you felt over the past week

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Are you basically satisfied with your life? yes/no</td>
</tr>
<tr>
<td>2.</td>
<td>Have you dropped many of your activities and interests? yes/no</td>
</tr>
<tr>
<td>3.</td>
<td>Do you feel that your life is empty? yes/no</td>
</tr>
<tr>
<td>4.</td>
<td>Do you often get bored? yes/no</td>
</tr>
<tr>
<td>5.</td>
<td>Are you in good spirits most of the time? yes/no</td>
</tr>
<tr>
<td>6.</td>
<td>Are you afraid that something bad is going to happen to you? yes/no</td>
</tr>
<tr>
<td>7.</td>
<td>Do you feel happy most of the time? yes/no</td>
</tr>
<tr>
<td>8.</td>
<td>Do you often feel helpless? yes/no</td>
</tr>
<tr>
<td>9.</td>
<td>Do you prefer to stay at home, rather than going out and doing new things? yes/no</td>
</tr>
<tr>
<td>10.</td>
<td>Do you feel you have more problems with memory than most? yes/no</td>
</tr>
<tr>
<td>11.</td>
<td>Do you think it is wonderful to be alive now? yes/no</td>
</tr>
<tr>
<td>12.</td>
<td>Do you feel pretty worthless the way you are now? yes/no</td>
</tr>
<tr>
<td>13.</td>
<td>Do you feel full of energy? yes/no</td>
</tr>
<tr>
<td>14.</td>
<td>Do you feel that your situation is hopeless? yes/no</td>
</tr>
<tr>
<td>15.</td>
<td>Do you think that most people are better off than you are? yes/no</td>
</tr>
</tbody>
</table>

Utility of the GDS

The GDS has been found to have good discriminant validity with older adults who are depressed, physically ill, and treated with cognitive behavioral therapy (Abraham, 1991; Friedman, Heisel, & Delavan, 2005; Kafonek, Ettinger, Roca, Kittner, Taylor, & German, 1989; Lesher, 1986; Trenteseau, Hyer, Verenes, & Warsaw, 1989; Yesavage et al., 1983). Through a number of different validity and reliability studies, the GDS has also been found to be a valid and reliable measure for screening depression in older adults in residential care settings (Abraham, Wofford, Lichtenberg, & Holroyd, 1994; Heiser, 2004; Parmelee, Lawton, & Katz, 1989), older medical or primary care patients (Bijl, van Marwijk, Ader, Beekman, & de Haan, 2006; Rapp, Parisi, Walsh, & Wallace, 1988), older home care patients (Marc, Raue, & Bruce, 2008), and hospitalized elders (Lyons, Strain, Hammer, Ackerman, & Fulop. 1989). While the GDS and GDS-15 have been found to be reliable and valid tools for depression screening in long-term care residents (Aikman & Oehlert, 2001; Jongenelis et al., 2005), researchers have also specifically constructed other GDS versions (GDS-8; GDS-12R; CS-GDS) to be used with this population (Jongenelis et al., 2007; Li et al., 2015; Sutcliffe et al., 2000).

Some studies have not supported the utility of the GDS, particularly concerning dementia patients (Burke, Houston, Boust, & Roccaforte, 1989; Kafonek et al., 1989). Compared with cognitively intact older adults, those with cognitive impairments are more likely to answer fewer GDS questions and display more positive response bias (Bedard et al., 2003). The GDS has been shown to be reliable in older adults with mild to moderate cognitive impairment (Mini-Mental State Examination [MMSE] ≥15; Jongenelis et al., 2005), but the sensitivity and specificity estimates considerably attenuate in older adults with more severe cognitive impairments (MMSE ≤14; McGivney, Mulvihill, & Taylor, 1994). In older adults with mild to moderate cognitive impairments, it has been suggested that the GDS-4 be used for screening of depression and GDS-15 for severity assessment (Isella, Villa, & Appollonio, 2002). Similarly, the GDS-15 has been shown to be a useful depression screener among the very old (age ≥85) with an MMSE score of 10 or more (Conradsson et al., 2013). Future efforts to strengthen the GDS for cognitively impaired older adults should consider validating the previously identified factor structure of dysphoria, meaninglessness, apathy, and cognitive impairment (Hall & Davis, 2009).

The utility of the GDS in identifying depression may also be compromised when cultural factors limit the patient’s understanding of the questions. For example, Flacker and Spiro (2003) found that four questions on the GDS-15 are “not understood” by at least 10% of African American respondents. On the other hand, there is a great deal of research supporting the use of the GDS across various cultures. The GDS has been adapted into numerous languages including Brazilian Portuguese (Zilenovski, 1991), Italian (Ferrario, Cappa, Bertone, Poli, & Fabris, 1990), and Spanish (Izal & Montorio, 1993). It has also been shown to be reliable and valid in many international countries including Iran (Malakouti, Fatollahi, Mirabzadeh, Salavati, & Zandi, 2006), China (Lai, Tong, Zeng, & Xu, 2010), and Portugal (Pocinho, Farate, Dias, Lee, & Yesavage, 2009). Furthermore, the GDS has been shown to be a reliable depression screener with older Chinese immigrants living in Western countries and Japanese American elders; however, there is limited validity data (Lin et al., 2016; Mui & Shibusawa, 2003). Yet, other studies have demonstrated the need to adapt the GDS cutoff points for different cultures (e.g., Vargas, Matsuo, & Blay, 2007).

USE OF THE INSTRUMENT IN PRIMARY CARE SETTINGS

Depression is one of the most commonly occurring mental health disorders in late life. In fact, depressive symptoms are present in 6%-15% of community-dwelling older adults (Akincigil et al., 2011;...
Geriatric Depression Scale

Unfortunately, symptoms of older individuals’ underlying mental health problems are often either ignored or misdiagnosed, or are simply attributed to the inevitability of the “aging process” and then left untreated (Butler, Lewis, & Sunderland, 1991; Simon & VonKorff, 1995). This is particularly concerning because untreated late-life depression can have devastating effects on a variety of domains, including cognitive performance (McBride & Abeles, 2000), life satisfaction (Lue, Chen, & Wu, 2010), and quality of life (Chachamovich, Fleck, Laidlaw, & Power, 2008), as well as social and physical functioning (Lenze et al., 2001; Liu, Leung, & Chi, 2011). In fact, depressive symptoms based on GDS screening have been shown to predict 5- and 8-year mortality (Rozzini & Trabucchi, 2012; Wada et al., 2011). Therefore, the use of depression screeners such as the GDS for older adults is desirable for a number of reasons.

Older adults are more likely to seek mental health treatment in primary care than in specialty mental health settings (Gum, Iser, & Petkus, 2010). Because of the role that the primary care physician and behavioral health clinician play in detection and treatment, accurate screening for depression is critical. Such screening can be accomplished quite effectively using the GDS. The GDS is, in these authors’ opinion, the screening instrument of choice for geriatric depression. It has all the desirable qualities: it is brief, well-tolerated by patients, cost-efficient, and exhibits sensitivity and reliability across a variety of populations. In addition, the GDS is in the public domain. That is, the instrument may be obtained for widespread use without payment or copyright requirements. For these reasons in particular, it is well suited for use in busy primary care settings. Unfortunately, Hammond (2004) explored doctors’ and nurses’ attitudes toward the GDS and found that only 10% of respondents would consider using it for routine screening. Specifically, doctors and nurses felt the GDS was “too depressing” for routine use and a barrier to rapport with the patient. To improve detection and treatment of late-life depression, additional implementation efforts should be made to increase the widespread use of the GDS.

Protocol for Use of the GDS as a Screener

To use the GDS as a screening measure for depression in older adults the following protocol could be followed. Either the GDS or the GDS-15 could be administered on a yearly basis to patients 60 years of age or older for the purpose of screening for depression. Additionally, if there is reason to suspect that a patient may currently be suffering from depression, it may be administered to that individual on a more frequent basis. When administering the GDS, it is also important to consider the cognitive status of the patient. If a patient is known to be or suspected to be moderately to severely cognitively impaired, a collateral source should complete the GDS for the patient or in addition to the patient’s GDS.

The validity of a collateral source completing the GDS for a patient has been established (Nitcher, Burke, Roccaforte, & Wengel, 1993). To compare patient and collateral source reports of depressive symptomatology, 170 patients completed the GDS, while their collateral sources completed a collateral version of the GDS (CS-GDS). Patients were identified as cognitively impaired or as cognitively intact, with results compared with the blind diagnoses made by geriatric psychiatrists. Results showed that good agreement was found among the clinical diagnoses that were made, as well as among responses on both the GDS and CS-GDS. It was concluded that the CS-GDS is an accurate depression screening instrument for both cognitively intact and impaired individuals. It is to be noted, however, that collateral sources reported 28 of the 30 depressive symptoms on the GDS significantly more often than did the patients. Thus, it is suggested that a higher cutoff (i.e., score of 18 or greater) score for depression be used on the CS-GDS than on the GDS (Chang, Edwards, & Lach, 2011).

In a primary care setting, research has shown that a cutoff score of 5 on the GDS-15 gives a sensitivity of .79 and specificity of .67 (Bijl et al., 2006). However, we recommend that providers establish
a low cutoff score for the GDS (7 for the GDS, 3 for the GDS-15). The purpose of setting a low cutoff score is to avoid false negatives (failures to identify a patient’s depression). When a collateral source is used, a higher cutoff score of 17 should be employed, as these sources tend to over-report depressive symptoms.

**Use of the Instrument’s Findings With Other Evaluation Data**

If a patient is found to score above the designated GDS cutoff score, providers need to then determine whether the patient’s primary treatment concern is depression or a condition with certain overlapping symptoms. Yet, providers are unable to determine whether the patient meets DSM-5 criteria for depression with the use of the GDS because some of the diagnostic criteria are not included in the measure, such as insomnia or hypersomnia, significant weight gain or loss, and suicidal ideation. Therefore, it is necessary to inquire further about the patient’s symptoms through semi-structured interviews. For example, a semi-structured clinical interview focusing on depressive symptomatology should be administered to both detect the presence of and to discriminate between various mood disorders, including major depression, dysthymic disorder, depressive disorder not otherwise specified, minor depression, adjustment disorder with depressed features, and mixed anxiety and depression. The diagnostic criteria for each of the aforementioned disorders are described in the *Diagnostic and Statistical Manual of Mental Disorders, 5th edition* (DSM-5; American Psychiatric Association, 2013).

Potentially useful semi-structured clinical interviews include the Diagnostic Interview Schedule (DIS-IV; Robins et al., 2000), as well as the Structured Clinical Interview for DSM-5 disorders (SCID-5; First, Williams, Karg, & Spitzer, 2014). These diagnostic interviews provide information regarding patients’ past and current disorders. In addition, they provide subtyping for depressive disorders that can, in turn, inform treatment planning when determining which treatment modality would be most effective. However, these interviews are time-consuming, require extensive knowledge of their proper administration, and may assess for more disorders than are of interest.

Other semi-structured clinical interviews may be more appropriate to use in a primary care setting if a quantitative estimate of depression severity is desired. To this end, the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1967) or the Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979) may be used. Scores of 17 or higher on the HRSD almost always results in a patient meeting DSM criteria for a depressive disorder (Scogin & McElreath, 1994).

**Determining a Treatment Approach**

Once a patient has been found to meet the diagnostic criteria for a depressive disorder through the use of the GDS and a follow-up clinical interview, an appropriate treatment plan must be established. First, however, the appropriate level of care required for the patient must be determined. Those individuals presenting severe depressive symptomatology (e.g., a patient with a score of 28 on the GDS who is unable to take care of himself or herself and indicates suicidal ideation) may require more intensive inpatient treatment.

Appropriate treatment modalities for the depressed individual may include antidepressant medication, psychotherapy, or a combination treatment including both treatment modalities. Pharmacotherapies (i.e., tricyclic antidepressants [TCAs]), selective serotonin reuptake inhibitors [SSRIs], and monoamine oxidase inhibitors [MAOIs]) have been identified as evidence-based interventions for late-life depression (Shanmugham, Karp, Drayer, Reynolds, & Alexopoulos, 2005; Wilson, Mottram, Sivananthan, & Nightingale, 2001). In comparing efficacy of TCAs and SSRIs, Mottram, Wilson, and
Strobl (2006) found both are efficacious; however, the total withdrawal rate due to side effects (i.e., participants that withdrew from the study) was significantly more for TCAs. SSRIs are thought to produce side effects that are less threatening to older adults than those associated with TCAs (Shanmugham et al., 2005). However, SSRIs may cause side effects such as nausea, diarrhea, insomnia, headache, agitation, and anxiety (Schneider, 1996). In addition, there are certain adverse side effects of SSRIs that are more common in older adults, including bradycardia, extrapyramidal symptoms, and abnormal secretion of antidiuretic hormones (Salzman, Schneider, & Alexopoulos, 1995). Thus, while antidepressant medications appear to be effective for treating geriatric depression, they are not without certain potential harmful side effects. Therefore, the use of antidepressant medication is more likely to be recommended by medical professionals as a treatment for older adults when depression is severe (Shanmugham et al., 2005). Older adults with scores of 20 or above on the 30-item GDS and for whom more rapid relief of symptoms is indicated, pharmacotherapy should probably be considered.

Although the effectiveness of antidepressant treatment has been well researched and documented, a significant percentage of those treated respond only partially to treatment or fail to respond altogether. It has been reported that a third of patients treated with an initial trial of antidepressants will not respond (Thase, 1997). Older patients may be even less treatment responsive than younger people (Thase & Rush, 1995). By the time patients have failed to respond to several different medications, the response rate per trial may be as low as 15%–25% (Thase & Rush, 1995). Fava and Davidson (1996) reported similar figures in a meta-analysis of 36 antidepressant therapy trials; 12%–15% of the patients exhibited only a partial response to therapy, and 19%–34% exhibited nonresponse. Studies focusing on older persons report similar figures; initial trial nonresponse rates as high as 40% have been reported for older patients (Flint & Rifat, 1996). For the present purposes, one should consider a score above 5 on the full GDS to be evidence of residual symptoms in a person currently being treated with pharmacotherapy or psychotherapy. Ideally, treatment would continue until symptoms are virtually nonexistent to protect against relapse and recurrence, but this is not always fiscally or time efficient.

Psychotherapy is another treatment option for geriatric depression that does not involve the adverse side effects of antidepressants. For persons with mild and moderate depression, as indicated by scores of around 5–20 on the full GDS, the authors suggest psychotherapy as the treatment of choice if it meets with client preference and there is a qualified therapist available. The efficacy of psychotherapy for depressed older adults has been established in a number of studies (e.g., Engels & Verme, 1997; Pinquart & Sörensen, 2001). Cuijpers, van Straten, and Smit (2006) conducted a meta-analysis that included the comparison of 17 randomized controlled interventions for late-life depression. They found an overall mean effect size of .72 (95% CI: .57–.87). Furthermore, Scogin and colleagues (2005) have identified specific psychological treatments as evidence-based, using criteria defined by a taskforce of the US-based Society of Clinical Psychology, in the treatment of late-life depression. These treatments include cognitive behavioral, behavioral, problem-solving, reminiscence and brief psychodynamic psychotherapies, as well as cognitive bibliotherapy (Scogin, Welsh, Hanson, Stump, & Coates, 2005).

Because cognitive behavioral therapy (CBT) is one of the most extensively researched psychotherapies, a bit more detail on this specific intervention will be provided. The primary goal of CBT is to improve the skills of older adults in examining and modifying maladaptive thoughts and belief systems. Some techniques used in CBT include recognizing and altering irrational thoughts, changing the way in which individuals process information, increasing positive self-statements and experiences, and countering mistaken belief systems. Behavioral techniques are also incorporated into treatment, including behavioral activation, relaxation training, problem-solving, and communication skills. The desired outcome of CBT is to diminish depression by developing reinforcing and
rewarding experiences and perceptions (Kennedy & Tanenbaum, 2000). CBT has been modified for use with older adults through the protocols developed by Thompson, Gallagher-Thompson, and colleagues (Thompson, Dick-Siskin, Coon, Powers, & Gallagher-Thompson, 2010). Meta-analyses of cognitive behavioral therapies for older adults indicate effect sizes ranging from .70 to .85 (Cuijpers et al., 2006; Scogin & McElreath, 1994). These effect sizes are comparable to those of younger adults in meta-analyses, using cognitive-behavioral approaches ($d = .72$; Michael & Crowley, 2002).

Another approach to treating geriatric depression is a combination treatment involving both psychotherapy and antidepressant medication. This approach might be indicated in cases of moderate to severe depression as indicated by scores in the 15–30 range on the GDS. Client preferences and financial resources also become considerations in such recommendations. In one study, the efficacies of a specific antidepressant medication (desipramine), a specific psychotherapy (cognitive behavioral therapy), and a combination approach (cognitive behavioral therapy and desipramine) for treating late-life depression were evaluated (Thompson, Coon, Gallagher-Thompson, Sommer, & Koin, 2001). The results showed that while patients in all three treatment conditions experienced substantial improvement of their depressive symptoms, the patients receiving the combination treatment and participants receiving CBT alone improved to a greater degree than did those receiving desipramine alone. The authors concluded that CBT is an effective treatment for depression in older adults, and that the combination of pharmacotherapy and CBT may be particularly effective for patients who do not respond completely to antidepressants or psychotherapy. The medication component of treatment can be provided by the patient’s primary care physician or a psychiatrist, while psychotherapy can be delivered by a qualified psychotherapist.

For persons evidencing mild depression (GDS of around 10 or less), another option is a self-administered treatment, such as cognitive bibliotherapy. Cognitive bibliotherapy has been effective for the treatment of depression in older adults (Floyd, Scogin, McKendree-Smith, Floyd, & Rokke, 2004; Landreville & Bissonnette, 1997). Some results indicate that the effectiveness of cognitive bibliotherapy is comparable to psychotherapy, with both showing significant improvement compared to control (Scogin, Hamblin, & Beutler, 1987; Scogin, Jamison, & Gochneaur, 1989). Floyd et al. (2004) found that although individual psychotherapy was superior to cognitive bibliotherapy at immediate posttreatment, the bibliotherapy participants continued to improve after treatment, showing no differences from psychotherapy at the 3-month follow-up evaluation. In another group of studies, treatment gains from cognitive bibliotherapy were maintained through a 2- and 3-year follow-up period (Floyd et al., 2006; Smith, Floyd, Scogin, & Jamison, 1997). Finally, a meta-analysis of nine studies of cognitive bibliotherapy for older individuals yielded an average effect size of $d = 0.57$ (CI = 0.37–0.77; Gregory, Canning, Lee, & Wise, 2004).

**Consideration of Client’s Treatment Preferences**

An important part of evidence-based practice is to include client preferences in the treatment decision-making process (American Psychological Association Presidential Task Force on Evidence-Based Practice, 2006). A meta-analysis found that incorporating patient preferences improved treatment outcomes and reduced drop-out rates (Swift & Callahan, 2009). As most older adults seek treatment for their depression from their primary care physician, it is imperative that these providers be aware of the treatment preferences of their depressed older patients. While most physicians primarily treat geriatric depression with antidepressants, this does not necessarily mean that older adults prefer pharmacological treatment over psychotherapy. In fact, depressed older adults tend to have a treatment preference for psychotherapy over pharmacotherapy (Gum et al., 2006; Rokke & Scogin, 1995).
These findings have important clinical implications for the treatment of geriatric depression. While many older adults prefer consulting their primary care physician when they are experiencing mental health problems, this does not mean that they necessarily consider psychotherapy to be less acceptable than pharmacotherapy. Providers need to be aware of not only the effectiveness of psychotherapy for treating geriatric depression, but also older adults’ acceptance of these treatments. Being able to provide a complete and accessible description of psychotherapy (i.e., that it is more than just “talk”) is vitally important. When designing a treatment plan for a depressed client, these preferences must be considered, as treatment acceptability influences adherence to treatment (Landreville & Guérette, 1998). The more acceptable an individual considers a treatment to be, the more likely that treatment will be completed and will be effective (Swift & Callahan, 2009). As treatment acceptability can affect adherence to treatments as well as their effectiveness, it is important to consider older adults’ preferences when treating geriatric depression.

**Limitations in the Use of the GDS for Treatment Planning**

In our opinion, the GDS should not be used in isolation to make treatment planning decisions. The GDS score is a useful indicant of depression severity, but like most instruments cannot provide contextual information. For example, do the depressive symptoms arise from a clear psychosocial stressor or do they represent a lifelong pattern of dysphoria? Does the present score represent increasing or decreasing depressive severity? What forms of treatment, if any, have been used by the respondent previously? Data provided by these types of questions and many more form the scaffolding from which treatment planning decisions can be diligently made. The GDS thus provides useful but necessarily limited information for treatment planning.

A specific potential limitation of the GDS is also one of its strengths, namely, its yes/no format. Patients may not feel comfortable with this format, as a “yes” or “no” response may not adequately describe the ambiguous or partial presence of some of the depressive symptoms. Patients may state that they experience the symptoms “sometimes.” In this case, it is necessary for the assessor to encourage the client to indicate one of the two choices, keeping in mind that their response should indicate the presence or absence of the symptoms the majority of the time. Thus, items should be individually discussed if time permits.

**Provision of Feedback Regarding Findings From the Instrument**

After the clinician has scored the patient’s completed GDS, attention should then turn to the individual items that were endorsed. For example, for patients endorsing GDS items that inquire about memory or concentration problems (i.e., “Do you feel you have more problems with your memory than most?”, “Do you have trouble concentrating?”), further feedback is needed. While depression often exacerbates age-related deficits in memory, it is nevertheless important to determine whether the patient may be experiencing memory problems severe enough to warrant an assessment of cognitive functioning. Additionally, for patients who do not endorse the question “Do you feel it is worthwhile to be alive now?,” it may be necessary to inquire further about potential suicidal ideation.

Feedback in the form of a score and a categorical descriptor is suggested. For example, a person with a score of 15 on the 30-item version might be told “You received a score of 15 on this depression scale, which probably means you are experiencing mild to moderate depression.”
USE OF THE INSTRUMENT FOR TREATMENT MONITORING

In addition to screening for depressive symptoms in older adults, the GDS can also be used to monitor a patient’s progress in treatment. Moreover, the GDS can be used to assess treatment outcomes. Additional information is provided in this section about the use and limitations of the GDS for treatment monitoring and treatment outcomes assessment.

Purpose of Treatment Monitoring

After an appropriate treatment plan has been developed, it is necessary to monitor the patient’s progress. Monitoring a patient’s progress on a regular basis will allow the provider to determine whether the treatment approach is effectively decreasing depressive symptoms. If a treatment approach doesn’t appear to facilitate improvements or the client’s symptoms worsen, alternative treatment approaches can be considered. Conversely, when a patient does respond favorably to an individual or combination treatment approach, it becomes necessary to monitor the patient’s progress so as to prevent the relapse or recurrence of depression. This becomes increasingly important in consideration of the finding that less than one-third of older adults will maintain a positive outcome (i.e., no relapse) after 1- to 3-year follow-ups (Murphy, 1994).

How to Use the Instrument for Treatment Monitoring

The GDS can serve as a beneficial tool for monitoring a patient’s progress. At the beginning of every weekly therapy session, the GDS can be administered to the patient. Various short form alternatives can be used if need be. After the therapist scores the GDS, he or she can discuss the results with the patient. Over time, the patient’s scores can be graphed showing the potential change over time. Endorsements of certain GDS items (i.e., suicidal ideation) should serve as an impetus for further discussion, particularly if there is a marked increase in depressive symptoms endorsed. Furthermore, symptoms that are consistently endorsed can be targets of intervention. The goal, to state the obvious, should be to decrease depressive symptoms (i.e., as evidenced by GDS scores) to as near zero as possible to decrease the likelihood of relapse or recurrence. Residual symptoms are risk factors for such negative long-term outcomes.

Limitations of the Use of the GDS for Treatment Monitoring

In addition to time constraints, other potential limitations of the GDS for treatment monitoring may include limits associated with the use of self-report measures and the yes/no response format. Particularly salient may be social desirability pressures on the patient. For example, as treatment continues and a therapeutic relationship develops between the patient and the clinician, patients may feel a need to under-report depressive symptoms. They may feel that this is necessary, as treatment “should” be working, in addition to not wanting to “disappoint” their clinician.

USE OF THE INSTRUMENT FOR TREATMENT OUTCOMES ASSESSMENT

The GDS can be used to assess treatment outcomes in different ways. First, it can be used for follow-up purposes after treatment is complete and the client’s depression is determined to be in remission.
In this scenario, the GDS can be used to monitor the maintenance of the patient’s remission status. Second, the GDS can be used at the predetermined end of treatment (e.g., 16–20 sessions of cognitive behavioral therapy), in addition to planned posttreatment evaluations. In this case, the GDS can be used to determine how effective the treatment was in decreasing the client’s depressive symptoms. Finally, treatment may be continued for an indeterminable amount of time until clients reach a predetermined low cutoff score on the GDS and are able to maintain these results.

The GDS is a popular research instrument that has been used extensively in different treatment outcome studies. The instrument’s ability to detect depressive symptomatology parallels other depression measures. It demonstrates sensitivity to change that is desired in outcomes studies.

**Use of Findings From the GDS With Other Evaluation Data**

The GDS can be used in conjunction with other evaluation data to assess treatment outcomes. For example, findings from measures such as the HAM-D, Symptom Checklist-90, Outcome Questionnaire, and the MMPI-2 may be compared with the GDS. These provide other evaluation data to determine whether changes in depressive symptomatology on the GDS corroborate similar findings on these different measures.

** Provision of Feedback Regarding Outcomes Assessment Findings**

As stated before, it is important to provide feedback regarding the client’s GDS scores at predetermined points throughout the course of treatment in a graphical format. When it has been determined that treatment is complete, it is suggested to hold a final feedback session regarding the client’s progress throughout the course of treatment. The patient’s progress may be mapped graphically by detailing the patient’s GDS score for each treatment session. Most clients’ scores will not decrease in a linear fashion, instead fluctuating over the course of treatment. Thus, the concept of successive approximation should be discussed with clients, explaining that these “setbacks” are expected as part of treatment, but do not mean that treatment wasn’t successful. When mapping the client’s progress using GDS scores, the change in GDS scores since the onset of treatment to the completion of treatment should be emphasized.

**Potential Limitations for Use of the Instrument for Outcomes Assessment**

The aforementioned limitations in using the GDS for treatment planning and monitoring (time restraints, social desirability, limits of self-report measures, and the instrument’s yes/no format) apply as well to treatment outcomes assessment. Thus, if there is concern about the validity of the client’s responses because of some of these factors, the use of a semi-structured interview, such as the HAM-D, may be helpful. For example, this interview would allow the patient to detail their mood in a more open-ended fashion that is not allowed with the yes/no format of the GDS. In addition, social desirability may be more difficult to maintain with a longer clinical interview, with more extensive questioning.

Newman, Ciarlo, and Carpenter (1999) detailed 11 guidelines for selecting instruments to be used in treatment planning and outcome assessment. The GDS meets the following guidelines, in that it has: relevance to a target group; simple, teachable methods; low measure costs relative to its uses; compatibility with clinical theories and practices; usefulness in clinical services; information
regarding the processes by which treatments may produce positive treatment outcomes; and psychometric strengths. In addition, the instrument is clearly understood by nonprofessional audiences, with the scoring procedures and results easily understood to all. The GDS does not meet the criterion for the use of measures with objective referents and multiple respondents. An objective reference is one in which concrete examples are provided at each level of an instrument, or at key points on the measure (Newman et al., 1999). As we have previously discussed, collateral sources may be indicated when a patient has moderate to severe cognitive impairment. In conclusion, the GDS meets most of the guidelines presented by Newman et al., further indicating its usefulness as a depression screener.

CASE STUDY

S.C. is a 64-year-old Caucasian male who presented to his primary care physician with complaints about memory loss, decreased energy, and low motivation. S.C. has been actively followed by his primary care physician for over six years to manage his hypertension, chronic hyponatremia, and gout. He also had squamous cell carcinoma of the right oropharynx and tongue, which was successfully treated with chemotherapy and radiation therapy. However, S.C.’s diet is now restricted to oral liquids or nutritional supplements through a feeding tube.

To explore factors for S.C.’s presenting concerns, the physician completed a physical examination; ordered a comprehensive metabolic panel, complete blood count and urinalysis; and screened his vitamin B-12, folic acid and thyroid-stimulating hormone levels. Labs did not reveal any significant findings. Next, the physician administered the GDS-15 to assess for depressive symptoms. S.C. scored an 8/15, which is suggestive of probable depression. The physician explained the GDS results to S.C., educated him about depression, explained the potential treatment options, and assessed his treatment preference. Because S.C. is currently taking several medications, he preferred to engage in psychotherapy. In an integrated care model, S.C.’s physician referred him to be seen by a psychologist embedded within the primary care clinic.

During S.C.’s initial visit with the psychologist, he completed a brief cognitive screener (St. Louis University Mental Status examination [SLUMS]), the GDS-30, and a diagnostic clinical interview. S.C.’s SLUMS score of 29/30 indicated normal cognitive functioning. On the GDS, he scored a 14/30, which is suggestive of mild to moderate depression. Specifically, S.C. endorsed feelings of worthlessness, unhappiness, sadness, and dissatisfaction with his life. He also endorsed cognitive symptoms such as poor concentration and memory problems. S.C. reported that he generally feels a lack of motivation and has dropped many activities and interests. During the clinical interview, S.C. discussed his current emotional stressors, which included physical health problems limiting his day-to-day functioning. For example, he reported that he is unable to do many activities around the house because of his energy level and physical weakness. He denied current or past suicidal or homicidal ideation, plan, or intent. It was determined that S.C. met DSM-5 criteria for major depressive disorder, moderate.

S.C. had no prior mental health treatment but endorsed a preference to engage in cognitive behavioral therapy (Thompson et al., 2010). This protocol involves 16–20 sessions of cognitive behavioral therapy, modified for use with older adults. S.C. was an eager and conscientious participant in treatment, completing homework assignments and engaging in therapeutic activities willingly. A productive working alliance was rapidly developed between him and his therapist.

Each week prior to the session, S.C. completed the GDS-15. The purpose of the weekly GDS was threefold: first, it gave the therapist an idea of S.C.’s current depression level; second, it
produced a means to measure S.C.’s improvement; third, it provided a scan of the symptoms or problem areas to target in treatment. His scores were graphed and used to help S.C. see his progress in therapy over time. The overall trend was a decrease of symptoms, but there were weeks throughout treatment when his score on the GDS-15 increased as well as several weeks (weeks 4 and 7) that evidenced dramatic decreases. This is often referred to as “sudden gain” in the psychotherapy literature (Tang & DeRubeis, 1999).

Seeing the decline in depression severity helped S.C. increase his hopefulness that he was capable of recovering from his depression. When S.C. was discouraged because of setbacks, the GDS was used as a tool to encourage him to continue working toward his treatment goal. The graphic presentation served as a reminder that, even when his score did increase slightly from the week before, overall he was moving in the right direction and was functioning better than when he first presented for treatment. By the end of the twentieth session, S.C. was significantly improved and evidenced a very low GDS score.

SUMMARY AND CONCLUSIONS

Geriatric depression affects a significant percentage of older adults, and when left untreated it can have detrimental consequences. Thus, accurate screening for depression is crucial. The GDS is an inexpensive, relatively brief, easy-to-use depression measure for older adults, with well-established validity and reliability. The instrument is able to assess the appropriate level of depressive symptomatology among patients (mild, moderate, severe), which then informs appropriate treatment approaches. Psychotherapy, antidepressant medication, or a combination of the two treatment modalities may be indicated. Once a treatment plan has been implemented, the GDS can be used to monitor the treatment’s progress throughout the course of treatment as well as assess treatment outcomes. In addition, the instrument can be used for follow-up purposes during posttreatment assessments. This allows monitoring of the maintenance of a client’s remission status. Thus, considering the wide-reaching benefits of the use of the GDS, it is important that the instrument be considered for incorporation into primary care and other clinical settings, as screening for depression in older adults should be a significant mental health priority.

REFERENCES


Assessment of Suicidal Risk
Julia A. Harris, Erika M. Roberge, Kent D. Hinkson Jr, and Craig J. Bryan

Of the many clinical settings in which suicide risk assessments are conducted, primary care has been identified as especially key, prompted in large part by research findings that approximately 45% of those who die by suicide visit their primary care doctor in the month before they die (Luoma, Martin, & Pearson, 2002) and 22% visit their general practitioner one week before dying by suicide (Pirkis & Burgess, 1998). In comparison, only 20% of suicide decedents visited a mental health professional during that same time frame. Similar trends have been noted among high-risk subpopulations such as the military. Between 2001 to 2010, for instance, 45% of military personnel who died by suicide, and 75% of those who made a suicide attempt or otherwise intentionally injured themselves, kept an outpatient medical appointment at a military treatment facility in the 30 days prior to their suicidal behavior, of which the most frequently visited clinic type was primary care and family medicine (Trofimovich, Skopp, Luxton, & Reger, 2012). In many ways, primary care often serves as a “last stop” for individuals who die by suicide.

The reasons underlying these trends are not fully understood, but some have suggested that primary care settings often serve as a “one-stop shop” for patients seeking either medical and/or psychological support and treatment. In fact, research has found that psychosocial and behavioral issues are the most common presenting issues during primary care visits, even when the chief complaint entails a physical issue or “purely” medical condition (Gatchel & Oordt, 2003; Kroenke & Mangelsdorff, 1989). Related to this, in the US healthcare system, primary care often serves as the gateway to the full spectrum of medical services. When an individual decides to “go to the doctor” because he or she is feeling unwell for any reason, primary care is very often their destination. Thus, primary care providers are not only the first health professional an individual will typically see when seeking support or health services related to behavioral and psychosocial factors; they are also very often the last health professional to provide care to suicide decedents.

In light of these trends, many have come to view primary care as an essential part of effective and comprehensive suicide prevention programs (The Joint Commission, 2016). For example, the role of primary care in suicide prevention was noted in the 1999 Surgeon General’s Call to Action to Prevent Suicide (Davidson, Potter, & Ross, 1999). Unfortunately, specific strategies and recommendations for suicide prevention in primary care settings were not clearly articulated until just recently as part of the National Strategy for Suicide Prevention published by the Office of the Surgeon General and the National Action Alliance for Suicide Prevention (2012). Specifically, Strategic Direction #2 of this document notes that “suicide assessment and preventive screening by primary care and other health care providers are crucial to assessing suicide risk and connecting individuals at risk for suicide to available clinical services and other sources of care” (p. 40) and that “assessment of suicide risk should be an integral part of primary care” (p. 58). Despite considerable consensus in favor of integrating suicide risk screening and assessment methods into primary care, the US Preventive Services Task Force (USPSTF) recommends neither for nor against it, noting that there is insufficient evidence to support its accuracy or effectiveness as a strategy to reduce or prevent suicidal behavior (LeFevre, 2014).
GENERAL CONSIDERATIONS

Clinicians across all healthcare settings have long struggled with identifying patients who might be at imminent risk of taking their own life (Galynker, Yaseen, & Briggs, 2014). Despite the wealth of empirical attention that suicide risk assessment has received over the past several decades (Berman & Silverman, 2014; Phillips, Leardmann, Gumbs, & Smith, 2010; Sheehan, Giddens, & Sheehan, 2014; Silverman & Berman, 2014), clinicians continue to struggle with this task, especially in primary care settings, due to the relatively low base rate of suicidal behavior as compared to other healthcare settings where suicide risk screening and assessment are common, such as inpatient and outpatient psychiatric settings and emergency departments. Although many factors likely contribute to the limited accuracy and efficacy of suicide risk screening and assessment in primary care, one particularly salient limitation is that suicide risk screening in primary care is frequently restricted to “high-risk” subgroups with one or more identified risk factors for suicide.

Limitations of Indicated Screening

In primary care, the high-risk subgroup for whom suicide risk is most frequently conducted is patients who screen positive for depression, consistent with an indicated approach to suicide risk screening. Indicated screening entails the targeting of “high-risk” individuals who identified as having minimal but detectable signs or symptoms suggesting the presence (or likely emergence) of a psychiatric or behavioral disorder (Institute of Medicine, 1994). In primary care, suicide risk screening and assessment very frequently occurs within the context of depression screening and/or treatment, such that patients with an established depression diagnosis and/or who screen positive for depression may also be explicitly screened for suicide risk. The indicated approach in which suicide risk screening occurs within the context of depression screening and treatment aligns with USPSTF recommendations (LeFevre, 2014). It is distinguished from the universal approach to suicide risk screening, in which an entire population is screened for risk regardless of their individual risk profiles. Primary care clinics using a universal approach might, for instance, screen all patients at every visit, regardless of their presenting complaint, demographics, or known risk factors.

A core assumption of the indicated approach to suicide risk screening is that suicidal behaviors are best conceptualized as a symptom of certain psychiatric conditions, especially depression. If suicidal behavior occurs primarily within the context of depression, limiting suicide risk screening to only those patients who screen positive for depression makes sense. Unfortunately, research clearly demonstrates that suicidal thoughts and behaviors cut across all psychiatric conditions and many medical conditions, not just depression (Harris & Barraclough, 1997). Accumulating evidence further suggests that depression and other variables that have traditionally been seen as key suicide risk factors (e.g., hopelessness) may not actually be risk factors for suicidal behavior after all. A recent meta-analysis (May & Klonsky, 2016) found, for instance, that although depression and hopelessness were statistically significant and useful correlates of suicidal thoughts, they were very poor correlates of suicidal behavior. Depression and hopelessness also performed poorly in differentiating between individuals who merely think about suicide from those who act upon their suicidal thoughts (May & Klonsky, 2016). Similar findings have been reported in large-scale epidemiological studies conducted in the United States (Kessler, Borges, & Walters, 1999; Nock et al., 2013) and 21 nations worldwide (Nock, Borges, & Ono, 2012). These findings are not limited to depression, however; with few exceptions, psychiatric disorders are relative poor predictors of suicide attempts and suicide death despite being robust correlates of suicide ideation (Kessler et al., 1999; May & Klonsky, 2016; Nock et al., 2012; Nock et al., 2013). An indicated approach to suicide risk screening that is contingent upon positive
responses to depression screening, or is limited to certain subgroups of patients with identified psychiatric conditions, is therefore inadequate.

**Limitations of Screening for Suicide Ideation**

A second limitation of suicide risk screening and assessment is an almost exclusive focus on suicide ideation. Arguably the most common method for suicide risk screening is to ask patients if they have been experiencing suicidal thoughts or urges during a particular timeframe (usually the past 1–2 weeks). In many cases, suicide screening tools ask respondents to indicate if they have experienced any suicidal thoughts at all (i.e., presence of ideation), how often they have experienced suicidal thoughts (i.e., frequency of ideation), or how severe these suicidal thoughts have been (i.e., intensity of ideation). The assumption underlying this approach is that suicide ideation must precede suicidal behavior: in order to attempt suicide, one must mentally reflect upon it first. A related assumption is that more severe or intense suicide ideation signals more imminent suicide risk. Both of these assumptions are predicated on a model of suicide that presumes a continuous spectrum of suicide risk in which there is a linear progression from lower risk states (i.e., absence of or mild suicide ideation) to moderate risk states (i.e., severe suicide ideation and/or suicide planning) to higher risk states (i.e., suicide attempts). Contrary to this continuum model, however, recent research suggests that the emergence of suicide ideation during the time leading up to suicidal behavior is often discontinuous in nature, and is characterized by fluctuations in risk states and suicide ideation over time (Wyder & De Leo, 2007).

Thus, although suicide ideation is an empirically supported and robust predictor of suicidal behavior (Brown, Steer, Henriques, & Beck, 2005), the vast majority of individuals who think about suicide do not attempt suicide or die by suicide. Suicide ideation therefore has much less positive predictive value for predicting suicidal behavior than is often assumed. This may be due in large part to the fact that suicide ideation and other commonly assessed risk variables (e.g., hopelessness, depression) tend to fluctuate over time (Bryan & Rudd, 2016), meaning that the results of screening depend to a large degree on the timing of the screening and the context within which the screening occurred. Because suicide ideation fluctuates over time, detecting and tracking risk across time in clinical and/or medical settings can be especially difficult. The difficulty of this task increases as the frequency of screening and assessment declines. Because suicide risk screening in primary care occurs fairly infrequently, the ability to detect temporal patterns and trends that might signal emerging risk is inherently limited.

The fluctuating nature of suicide ideation may explain why so many individuals who attempt suicide or die by suicide are missed via existing screening methods. Studies indicate that anywhere from one-half to two-thirds of individuals who die by suicide failed to disclose or denied suicide ideation or suicidal intent during the last medical visit preceding their death (Busch, Fawcett, & Jacobs, 2003; Coombs et al., 1992; Hall, Platt, & Hall, 1999; Kovacs, Beck, & Weissman, 1976). Screening for suicide ideation therefore misses over half of individuals who die by suicide. In light of these limitations, what is the primary care provider to do? An important first step is adopting an empirically supported and practical model of suicidal behavior.

**The Fluid Vulnerability Theory of Suicide**

One theory of suicide that captures the dynamic, temporal nature of suicidal thoughts and behaviors is the fluid vulnerability theory (FVT; Bryan & Rudd, 2016; Rudd, 2006). The FVT posits that suicide risk is best conceptualized on two dimensions. The first dimension, often referred to as baseline risk,
refers to the stable properties of suicide risk that persist over time (e.g., demographics, personality traits, emotion regulation deficits), whereas the second dimension, referred to as acute risk, refers to the dynamic properties of suicide risk that fluctuate or change over time. In some cases, these fluctuations can happen very rapidly on a moment-to-moment basis (e.g., depression, hopelessness, suicide ideation). This is depicted in Figure 11.1, which plots suicide risk over time for two hypothetical cases. As can be seen in this figure, both individuals experience fluctuations in suicide risk over time; this corresponds with the acute dimension of risk. In addition, the individual depicted by the solid black line has a higher baseline level of risk than the individual depicted by the dashed line. As a result, the person depicted by a solid line has a slightly higher overall risk for suicide over time than the person depicted by the dashed line, even when the former is at his or her relative best.

Embedded within the FVT is the suicidal mode, which serves as the structural basis for organizing risk and protective factors and conceptualizing an individual's suicide risk (see Figure 11.2). Within the suicidal mode are four domains including both stable and dynamic factors that may contribute to the onset of an acute suicidal crisis: emotional, cognitive, physiological, and behavioral. The stable aspect of each domain entails those properties of risk and protective factors that increase an individual's propensity for becoming suicidal and, when experiencing a suicidal crisis, for transitioning to suicidal behavior. The dynamic aspect of each domain, by contrast, entails those properties of risk and protective factors that change over time, often in response to life events and/or aversive internal experiences (e.g., nightmares, mood states, agitation). For example, in terms of the cognitive domain, an individual might have a higher level of cognitive reactivity to events that occur in the environment and relatively entrenched negative self-schemas (i.e., stable dimension). In response to stressful life events, he or she might also experience situational hopelessness and suicide ideation (i.e., dynamic dimension). The combination of the individual's stable cognitive style and their dynamic thought processes thereby influence whether or not they will become suicidal in response to a given life stressor, and how that suicidal crisis will unfold over time. The stable properties of risk and protective factors thereby contribute to an individual's baseline risk for suicide, and the dynamic properties of risk and protective factors influence the individual's acute risk.

Suicidal episodes often occur when an individual with sufficient baseline risk experiences a sufficiently intense triggering event, thereby activating the acute suicidal mode. Triggering events can be external environmental events (e.g., relationship problem, financial strain) or they can be internal experiences (e.g., trauma-related memories, negative mood state, physiological arousal). Individuals with elevated baseline risk are more easily triggered than individuals with fewer predisposing vulnerabilities. Individual differences in baseline risk can help to explain why some individuals become suicidal but others do not. Although individuals with relatively low baseline risk are less likely to become suicidal and to engage in suicidal behavior, it is possible for them to experience suicidal episodes if
they experience sufficiently extreme or severe triggers. Understanding who will become suicidal and/or engage in suicidal behavior therefore requires the consideration of triggers within the context of baseline risk.

**Implications for Suicide Risk Screening and Assessment**

Because suicide risk is a fluid and ever-changing construct, suicide risk screening and assessment in primary care should, wherever possible, seek to capture these properties. This is admittedly easier
said than done, as most patients do not visit primary care clinics on a regular or frequent basis. As a result, suicide risk screening and assessment in primary care is often based on information available at a single point in time and typically entails the weighing of a given patient's identified risk (and protective) factors (Borges et al., 2010; Glenn & Nock, 2014; Kessler et al., 1999; Nock et al., 2008). As described by Kraemer et al. (1997), a risk factor is defined as a “measurable characterization of each subject in a specified population that precedes the outcome of interest and which can be used to divided the population into two groups (e.g., the high-risk and the low-risk groups that comprise the total population)” (p. 338). As applied to the issue of suicide, risk factors must occur or be present prior to the onset of suicidal thoughts and behaviors, and must be able to distinguish between high and low levels of risk for suicide. Although risk factors are by definition correlated with the outcome or condition of interest, they may not necessarily be causally related to the condition. Risk factors that influence or change the condition of interest to the degree with which it is manipulated are therefore referred to as causal risk factors. For example, suicide ideation is generally assumed to be a causal risk factor for suicidal behavior in that increased severity and/or frequency of ideation are associated with increased likelihood of making a suicide attempt (Brown et al., 2005).

Distinguishing between a suicide risk factor and a causal suicide risk factor is important for suicide risk screening and assessment because causal risk factors should be considered more informative for clinical decision-making (Kraemer et al., 1997). Numerous risk factors for suicide have been empirically identified and disseminated to include variables such as prior history of suicidal thoughts/behaviors, male gender, under the age of 50 years old, unmarried relationship status, unemployment, lower education/income, parental psychopathology, and the presence of a DSM mental disorder (Borges et al., 2010). Because the number of suicide risk factors is so extensive, clinicians can maximize the accuracy of risk assessment in an efficient manner by prioritizing certain variables over others (e.g., Bryan & Rudd, 2006). Adaptations of this approach have been developed and applied to primary care settings (Bryan & Rudd, 2011; Bryan, Corso, Neal-Walden, & Rudd, 2009). For example, Bryan et al. (2009) recommended prioritizing the assessment of prior suicidal behavior, the presence and specificity of suicide planning, intensity of suicide ideation, and preparatory or rehearsal behaviors, because these risk factors have been shown to predict suicidal behaviors better than other risk factors. Notably, this cluster of variables is composed of causal risk factors in that they each influence the likelihood that suicidal behavior will occur in the future. Furthermore, this cluster of variables captures both baseline (i.e., prior suicidal behavior) and acute (i.e., ideation, planning, preparatory behavior) aspects of suicide risk.

Although causal risk factors like suicide ideation are especially useful for suicide risk screening and assessment, clinicians must remember that these constructs are inherently dynamic. The results of suicide risk screening and assessment should therefore be seen as a snapshot of the patient's risk at any given point in time. Consider, for example, the two patients depicted in Figure 11.1. The patient depicted by the solid line (Patient A) is very clearly a higher risk individual as compared to the patient depicted by the dashed line (Patient B). If these two patients were screened between time points 4 and 12, our obtained scores would likely reflect this considerable risk differential. If, however, these two patients were instead screened between time points 13 and 17, they would seem fairly similar to each other and the higher risk status of Patient A might be missed or underestimated. Finally, if these two patients were instead screened at time point 25, we might conclude that Patient B is the higher risk individual. Primary care providers must remember that the preponderance of existing screening and assessment methods are designed to assess causal risk factors that change over time. The value of suicide risk screening and assessment tools are therefore maximized when they are used repeatedly and are considered within the context of the patient's history.
The Problem of False Negatives

The FVT also provides a useful framework for understanding and approaching another challenge of suicide risk screening and assessment: false negatives. False negatives refer to patients who screen negative for suicide risk and then subsequently make a suicide attempt or die by suicide. As discussed above, the current “state-of-the-art” approach for suicide risk screening and assessment almost always entails directly asking a patient about suicidal thoughts or intentions (e.g., “Do you have any intent or plan to kill yourself?”; Glenn & Nock, 2014). The accuracy of a patient’s positive response to this question (or related questions) depends on several factors: (1) the actual presence of suicidal thoughts and intentions, (2) the patient’s awareness of such thoughts or intentions, and (3) the patient’s willingness to honestly disclose these thoughts and intentions to someone else (e.g., a primary care provider). False negatives (also referred to as “misses”) are influenced by each of these factors, although the precise amount that can be attributed to each remains unknown. If a patient endorses suicidal thoughts or intentions, it is reasonable to conclude that all three conditions have been met: suicidal thoughts and intentions are present, the patient is aware of this, and he or she is willing to communicate this to the provider.

If a patient denies suicidal thoughts or intentions, however, it is not always clear which of these three conditions have or have not been met. In most cases, a denial of suicidal thoughts and intentions reflects the absence of these experiences; in this case, the patient is honestly and accurately denying overt suicide risk. In other cases, patients may be experiencing suicidal thoughts and intentions but may not recognize these thoughts as suicidal in nature. The patient is therefore being honest but not accurate, although they do not realize that they are providing an inaccurate response. In many cases, patients in this category are in a prodromal stage of emerging suicide risk. Finally, there are cases in which patients who are experiencing suicidal thoughts and intentions recognize their risk, but choose not to disclose this to the provider. The patient is therefore concealing his or her risk from the provider. Patients in this group often express concern about the perceived negative consequences of honest self-disclosure. For example, patients may be afraid that a provider will hospitalize them, and prevent them from carrying out their intended suicidal behavior.

Primary care providers should therefore keep in mind that a negative screen for suicide risk does not necessarily mean that a patient will not become suicidal or engage in suicidal behavior at a later date. As noted before, over half of individuals who make a suicide attempt or die by suicide screen negative or deny suicidal intent during their most recent medical appointments (Busch et al., 2003; Coombs et al., 1992; Hall et al., 1999; Kovacs et al., 1976). Suicide risk screening and assessment methods should therefore be complemented with other relevant screening and assessment methods. For example, a patient who denies suicide ideation but reports severe insomnia, mood disturbance, and agitation, and makes statements that he or she “can’t take it anymore” may be at elevated risk for suicide despite the absence of explicit disclosure of suicidal thoughts and intent. Because suicide risk is a fluid process, repeated suicide risk screening and assessment is an important strategy for catching and monitoring at-risk patients such as these who may have been missed in previous screenings and/or who may be unwilling to disclose suicidal intentions.

Suicide Risk Screening Versus Assessment

Before proceeding further, it is important to distinguish between suicide risk screening and suicide risk assessment. As noted by Bryan and Rudd (2006), suicide risk screening is designed to maximize sensitivity at the expense of specificity because the primary purpose is to identify as many patients
as possible who might be suicidal. Screening methods are therefore typically brief and result in high false positive rates (i.e., nonsuicidal patients being identified as suicidal). Suicide risk assessment, by contrast, is designed to maximize specificity at the expense of sensitivity because the primary purpose is to develop a more fine-grained estimation of the patient’s overall risk level. Suicide risk assessment is therefore often more time- and resource-intensive. Suicide risk screening and assessment are not mutually exclusive; in primary care, suicide risk screening often leads to a more detailed risk assessment. For example, suicide risk screening in primary care often consists of administering a broad, self-report questionnaire to patients that covers a variety of suicide risk factors such as suicide ideation. A positive screen serves as an indication that a full risk assessment should be conducted, which typically entails a clinical interview by a healthcare provider.

The remainder of this chapter will provide an overview and description of three commonly used measures used in primary care settings for suicide risk screening and assessment. In addition, strategies for effectively implementing these tools and their associated challenges will be discussed.

**SUICIDE RISK SCREENING INSTRUMENTS AND TOOLS**

Choosing an effective screening tool to assess for suicide risk can be challenging given the fluctuating nature of suicidal thoughts and behaviors. It is also important to consider whether the screening instrument addresses suicidal thoughts only, or also specifically addresses suicidal behaviors. Finally, it is worth considering the duration and frequency of these thoughts/behaviors, recognizing that an individual who screens negative on certain screening instruments might have screened positive had they taken the measure a few days prior. In this section, three potential suicide risk screening tools will be presented, as well as their corresponding strengths and weaknesses.

**Patient Health Questionnaire 9-Item Depression Subscale (PHQ-9)**

The Patient Health Questionnaire (PHQ) is a three-page self-report questionnaire that was developed to screen for and assess a wide range of mental health conditions that are commonly experienced by medical patients. The PHQ consists of several subscales, each of which is designed for the purpose of aiding health professionals in making mental health diagnostic decisions in a manner that aligns with DSM-IV diagnostic criteria (American Psychiatric Association, 2000). The depression subscale of the PHQ consists of nine items that correspond with each DSM-defined symptom of major depressive disorder (Kroenke, Spitzer, & Williams, 2001). Because the depression subscale entails a total of nine items, it is frequently referred to as the PHQ-9. The ninth item of the scale assesses for thoughts about death and self-harm (i.e., “Thoughts that you would be better off dead or hurting yourself in some way”); it is therefore widely used for suicide risk screening. A briefer version of the scale, which is comprised of the scale’s first two items and the suicide screening item (referred to as the PHQ-3) is also widely used. Because of the scale’s brevity, considerable psychometric strength as an indicator of depression, and low cost (i.e., it is in the public domain), it is arguably one of the most widely used screening tools in primary care settings.

**Administration and Scoring**

The PHQ-9 is a self-report questionnaire that directs patients to rate the frequency with which they have experienced each symptom during the past two weeks on the following 4-point Likert scale: not at all (0), several days (1), more than half the days (2), and nearly every day (3). The scale typically takes
less than three minutes to complete; it can therefore be easily and flexibly administered to patients in primary care clinics. For example, patients can be asked to complete the scale in the waiting room or an examination room prior to or during their appointment in either paper or electronic format. To receive a diagnosis of major depressive disorder (MDD), five of the nine items need to be positively endorsed and present for more than half of the days within the past 2 weeks. The exception to this is item 9, which addresses suicidal thoughts/behaviors and counts regardless of duration. Additionally, one of the five symptoms (i.e., items) is required to be feeling down/depressed or decreased interest in activities (i.e., anhedonia). When the PHQ-9 is used as a measure of MDD severity, symptoms are rated on a scale of 0 (not at all) to 3 (nearly every day). The PHQ-9 can also be administered in a binary yes/no format; however, results are more accurate when administering in the frequency format (see “Relevant Psychometric Support” for more information). Scores range from 0 to 27, with higher scores indicating more severe depressive symptoms. The PHQ-9 is available in paper-and-pencil format and can be easily converted to electronic format for integration into electronic medical records and/or other patient tracking systems.

Available Norms
The PHQ-9 was initially validated in a sample of approximately 6,000 individuals presenting to eight primary care clinics and to seven obstetrics-gynecology clinics within the United States. In these samples, high scores on the PHQ-9 were significantly related to functional status, symptom-related difficulty, number of sick days taken, and an increase in healthcare utilization (Kroenke et al., 2001). A recent meta-analysis of 18 validation studies using the PHQ-9 indicated there is considerable variability in the scale’s sensitivity and specificity, but found that a cutoff score of 11 was likely optimal for a diagnosis of MDD (Manea, Gilbody, & McMillan, 2012). This meta-analysis did not consider the instrument’s psychometrics specific to suicide risk, however.

General Strategy for Interpreting Test Results
As a suicide risk screening tool, item 9 is often interpreted in a binary manner in which a score of 0 ("not at all") is interpreted as a negative screen or “not suicidal” whereas any positive endorsement is often interpreted as a positive screen or “suicidal.” Although item 9 is often referred to as the “suicide item,” it is important to note that the item’s wording includes a much broader range of thought processes than suicide ideation alone. Specifically, the item’s wording directs patients to rate the frequency of their “thoughts that you would be better off dead or hurting yourself in some way.” The item may therefore assess the desire for death without suicidal intent as well as thoughts about nonsuicidal self-injury (discussed in greater detail later). This may account for the item’s relatively high false positive rate, especially among patients who endorse a score of 1, of which only one-third have are estimated to have actual suicide ideation (Corson, Gerrity, & Dobscha, 2004). Although a binary interpretive approach to the PHQ-9’s item 9 may be practical for developing clinical pathways and risk management plans, clinicians are generally discouraged from using a binary interpretive approach for the purposes of suicide risk assessment and clinical decision-making in light of evidence that incremental increases in scores are associated with increased risk for suicide attempts and suicide death (Simon et al., 2013).

Use as a Screener for Suicide Risk
One of the PHQ-9’s many strengths is its brevity; as previously mentioned, the entire scale often takes less than three minutes to complete (Kroenke et al., 2001). In addition to its brevity, the PHQ-9 is publicly available, making it an affordable option for many clinics. Indeed, it is the combination of the scale’s efficiency, cost, and psychometric strength that makes the PHQ-9 one of the most widely administered scales for depression and suicide risk screening. The PHQ-9’s limitations as a suicide risk
screening tool include its emphasis on thoughts of death and self-harm during the past two weeks, which can potentially miss those patients with fluctuating suicide ideation and/or discontinuous suicidal episodes. In addition, the item’s face validity is vulnerable to motivated responding, thereby interfering with the instrument’s accuracy. Despite these limitations, results of empirical studies (e.g., Simon et al., 2013) support the conclusion that positive screens on the PHQ-9 indicate increased risk for suicidal behavior. Positive screens to the entire screener should therefore be followed by a suicide risk assessment interview by a clinician to clarify and obtain a more nuanced understanding of the patient’s suicide risk level. Negative screens on the PHQ-9, by contrast, do not necessarily indicate the absence of elevated suicide risk. If a patient denies thoughts of death or self-harm on the PHQ-9 but is manifesting other indicators of elevated suicide risk (e.g., insomnia, agitation, self-hatred), clinicians may nonetheless want to probe further.

As noted previously, many primary care clinics employ an indicated approach to suicide prevention screening in which screening occurs only among patient subgroups with identifiable characteristics associated with increased risk for suicide. This contrasts with the universal approach to suicide prevention screening in which all patients are screened regardless of their individual risk level. In primary care, suicide risk screening is often tied to the presence of a depression diagnosis and/or the results of depression screening. For example, a commonly used approach in primary care is to screen for the presence of depression using the PHQ-2, which is comprised of the PHQ-9’s first two items (i.e., depressed mood and anhedonia). If a patient screens positive for probable depression, the remaining seven items of the PHQ-9 (including item 9) are then administered. If a patient screens negative for probable depression, however, no further screening is conducted. This two-step approach, which aligns with USPSTF’s recommendation, hinges on the assumption that suicide risk occurs only within the context of acute depressive episodes, despite considerable evidence indicating that suicidal behavior cuts across most (if not all) psychiatric conditions (Harris & Barraclough, 1997). An indicated suicide prevention screening that is contingent upon positive responses to only two depressive symptoms is therefore inadequate. For these reasons, universal suicide risk screening in primary care has gained traction and become more common.

Relevant Psychometric Support
The PHQ-9 was initially developed to align with DSM-IV-TR diagnostic criteria for MDD (American Psychiatric Association, 2000). Because these criteria were minimally changed in the transition to DSM-5, the PHQ-9 largely aligns with current DSM diagnostic criteria for MDD (American Psychiatric Association, 2013). The ninth symptom criterion for major depressive disorder is “recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.” This specific criterion is assessed by item 9. The PHQ-9 is therefore fairly consistent with the DSM-5’s symptom criterion, although its specific wording (i.e., “hurting yourself”) assesses for a broader spectrum of self-injurious thoughts and behaviors than the DSM-5’s wording, which is specific to suicidal thoughts and behaviors.

The newest edition of the DSM includes a new condition for further study termed suicidal behavior disorder, which entails having made a suicide attempt within the past two years. Suicidal behavior disorder excludes nonsuicidal self-injury, preparatory behaviors, and suicide ideation, as well as self-directed violence that occurs during an altered mental state (e.g., delirium or “confusion”) or due to ideological motives (e.g., religious, political). This proposed diagnosis is therefore specific to suicidal behavior, not thoughts, the latter of which is assessed by the PHQ-9’s suicide risk screening item.

Aside from the DSM, the past decade has seen considerable progress in the standardization of language and terminology specific to suicide-related constructs (Crosby, Ortega, & Melanson, 2011; Matarazzo, Clemans, Silverman, & Brenner, 2013; Silverman et al., 2007a, 2007b). These efforts have resulted in the development of specific definitions by the Centers for Disease Control and Prevention
Assessment of Suicidal Risk

(Crosby et al., 2011), which have since been adopted by a growing number of researchers, policy-makers, and agencies. Using this framework, the PHQ-9’s suicide risk screening item appears to assess at least two constructs: self-harm ideation, which is defined as any thought of or communication regarding a person’s desire to engage in self-inflicted potentially injurious behavior without any evidence of intent to die; and suicide ideation, which is defined as thoughts of engaging in suicide-related behavior (cf. Crosby et al., 2011, p. 90). The PHQ-9 also very clearly measures morbid thinking about death that does not necessarily involve suicidal intent. Thus, although item 9 is widely referred to as the “suicide risk” item and researchers often use this item to denote the presence (or absence) of suicide ideation, positive endorsement may not necessarily be specific to suicide ideation. This may explain in part why the PHQ-9’s performance as an indicator of suicide risk and eventual suicidal behavior is not as strong as many clinicians presume.

Although the validity of the PHQ-9’s item 9 has been examined in several studies (Corson, Gerrity, & Dobscha, 2004; Simon et al., 2013; Uebelacker, German, Gaudiano, & Miller, 2011), most studies were designed to assess the scale’s concurrent validity as compared to other measures of suicide ideation, not its prospective validity as a predictor of later suicidal behavior. For example, a study of 166 primary care patients reported that the PHQ-9’s suicide item had a sensitivity of .69 and specificity of .84 for detecting recurrent thoughts of death or suicide as measured by the Structured Clinical Interview for DSM-IV (SCID; Uebelacker et al., 2011), but did not examine how the item predicted subsequent suicidal thoughts or behaviors. A separate study conducted in a Veterans Affairs (VA) hospital found that approximately one-third of primary care patients screening positive on the PHQ-9’s item 9 had suicide ideation as measured by the SCID (Corson, Gerrity, & Dobscha, 2004).

To date, only one study has considered the prospective validity of the PHQ-9 for predicting suicidal behavior. In a large ($n=84,418$) sample of medical outpatients within a large healthcare system, Simon and colleagues (2013) examined the association of PHQ-9 scores with suicide attempt and suicide death during the 4-year period. Results of this study supported a cumulative risk of suicide attempt based on patients’ response to item 9 of the PHQ-9, increasing from a cumulative probability of 0.4% (1 in 250) among patients with a score of 0 (“not at all”) on the item to a cumulative probability of 4% (1 in 25) for those with a score of 3 (“nearly every day”). Suicide attempts were approximately 6.4 times more likely among patients who scored a 3 on the item as compared to those who scored a 0. Simon and colleagues further reported that the cumulative probability of suicide death increased from 0.03% (1 in 3,000) for those with a score of 0 to 0.3% (1 in 300) for those with a score of 3, indicating suicide death was approximately 5.3 times more likely among patients who scored a 3 on the item as compared to those who scored a 0.

Summary

Although limited, the PHQ-9 has value as a suicide risk screener and is widely recommended for this use because of its ability to detect suicidal individuals who would otherwise remain undetected (Bryan & Rudd, 2011; Bryan, Corso, Neal-Walden, & Rudd, 2009). Of note, responses on item 9 are associated with incremental improvement in the prediction of subsequent suicidal behavior (Simon et al., 2013). Very high scores, in particular, seem to be especially meaningful and useful. Understanding the PHQ-9’s limitations can help clinicians to use it in an effective and practical way in primary care. Interested readers are referred to Chapter 9 of this book for more information regarding the PHQ-9.

Behavioral Health Measure-20

The Behavioral Health Measure-20 (BHM-20) is a brief, 20-item measure of common psychological symptoms and associated problems that was originally developed with the intention to serve as
a practical and efficient tool for assessing overall mental health functioning and psychotherapy outcomes (Kopta & Lowry, 2002). The BHM’s construction was influenced by the phase model of psychotherapy, which posits that clinical improvement during the course of psychological treatment begins with improvement in subjective well-being, followed by reductions in symptomatology, and finally enhanced life functioning (Howard, Lueger, Maling, & Martinovich, 1993). Consistent with this model, three items of the BHM-20 assess subjective well-being (e.g., “How distressed have you been?”), 13 assess common mental health symptoms (e.g., “How much have you been distressed by feeling hopeless about the future?”), and four items assess problems in social-occupational domains (e.g., “How have you been getting along in work/school since your last appointment?”). Item responses range in value from 0 to 4, with higher scores indicating better mental health. The mean value of all 20 items is calculated to obtain a Global Mental Health (GMH) score. Although originally designed for outpatient psychotherapy settings, the BHM-20 has since been successfully implemented as a clinical outcomes tracking system in primary care behavioral health settings (e.g., Bryan, Corso, Corso et al., 2012; Bryan, Corso, Rudd, & Cordero, 2008; Cigrang et al., 2011; Ray-Sannerud et al., 2012). The tenth item of the BHM-20 assesses for suicide ideation (i.e., “Thoughts of ending your life”) that can be used for suicide risk screening. Patients who positively endorse this item are directed to also respond to a supplemental item that assesses their subjective level of suicide risk (i.e., “If you answered 0–3 on question #10 above, please check below to indicate your overall risk of suicide”) on a scale ranging from “extremely high risk” to “no risk.”

Administration and Scoring
The BHM-20 is a self-report questionnaire that directs patients to rate the severity or frequency of each symptom or problem during the past two weeks on a Likert scale ranging from 0 to 4, with higher scores indicating better health (i.e., less severe problems or symptoms). The scale typically takes less than five minutes to complete. When used in primary care clinics, patients have been asked to complete the scale in the waiting room upon checking in for an appointment. A web-based administration system is also available from the CelestHealth Solutions, which calculates all scores automatically and provides a color-coded visual graphic of the patient’s scores over time, thereby enabling quick and easy interpretation by the clinician.

Available Norms
The BHM-20 was initially validated in a community sample (n = 380), a college student sample (n = 465), a college counseling center sample (n = 206), and an adult outpatient psychotherapy sample (n = 211). Means and standard deviations for each are reported in Kopta and Lowery (2002), and confirmed that community samples had the highest (i.e., healthiest) scores, followed by the college counseling center sample and the outpatient psychotherapy sample. Reliability and normative data for the BHM-20 in primary care behavioral health settings have since been published (Bryan, Blount et al., 2014) using data from three separate samples of primary care patients who were referred to an integrated primary care behavioral health provider (total n = 2,902). Across the four subscales that emerged from this study, Cronbach’s alpha values ranged from .72 to .93. Overall, the mean BHM-20 score for men was 2.87 (SD = 0.67) and for women it was 2.61 (SD = 0.70). These norms were comparable to those obtained from the college counseling center sample reported by Kopta and Lowery. Clinical cutoff scores originally derived for college counseling centers are therefore most appropriate for primary care behavioral health settings. The BHM-20 manual lists the following cutoff scores to designate clinically meaningful subgroups: distressed/clinical (mean score <2.74), at-risk (mean
score from 2.74 to 2.93), and healthy (mean score >2.93). Cutoff scores were derived based on the recommendations and methods described by Jacobson and Truax (1991) for establishing clinically significant change indices. Scores can therefore be used to track a patient’s progress over the course of treatment.

Norms for the BHM-20’s suicide ideation item have not been published, although one report found that 12.4% primary care patients who were referred to an integrated behavioral health provider endorsed this item (Bryan, Corso, Corso et al., 2012): 4.9% reported “a little bit,” 3.2% reported “sometimes,” 2.9% reported “often,” and 1.3% reported “almost always” thinking about suicide. A separate report found that 7% of military personnel who screened positive for posttraumatic stress disorder endorsed suicide ideation on the BHM-20 when evaluated in primary care (Bryan & Corso, 2011).

General Strategy for Interpreting Results
A score of 4 on item 10 of the BHM-20 can be interpreted as a negative screen for suicide ideation; conversely, scores ranging from 0 to 3 can be interpreted as a positive screen, with lower scores indicating more severe suicide risk. In contrast to the PHQ-9, the BHM-20’s suicide ideation item is much more specific to the construct of suicide ideation, although the specific language used does not explicitly include the word “suicide” (i.e., “thoughts of ending your life”). Patients who endorse suicide ideation on item 10 are then directed to rate their subjective level of risk on a supplemental item ranging from “extremely high risk” to “no risk.” Responses on this item can be interpreted as an indicator of the severity of the patient’s suicidal intent.

Use as a Screener for Suicide Risk
A primary strength of the BHM-20 is its brevity; on average, a patient can complete the scale in less than five minutes (Kopta & Lowery, 2002). The BHM-20 may also be ideal in primary care settings characterized by diversity in presenting complaints and problems, in contrast to diagnosis-specific measures that may not be sensitive enough to detect distress or functional impairment associated with subthreshold psychiatric conditions and/or psychosocial problems more generally. Information about costs for using the BHM-20 is available from CelestHealth Solutions, the instrument’s commercial vendor.

Similar to other scales, the BHM-20 can be used in either an indicated or universal approach. To date, the BHM-20 has primarily been used as a universal screening tool within the context of primary care behavioral health services. When used in this capacity, the BHM-20 has been found to improve the identification of patients with recent suicide ideation sixfold as compared to the indicated approach that is standard practice in many primary care clinics (Bryan, Corso, Rudd, & Cordero, 2008), which may support its use as a universal screening tool. Positive endorsement of the BHM-20’s suicide ideation item should be followed by a suicide risk assessment interview by the clinician to further clarify risk and make appropriate clinical decisions.

Relevant Psychometric Support
The BHM-20 was developed as a tool to track clinical outcomes, not as a diagnostic tool in its own right; its items therefore do not directly map onto any DSM-5 diagnosis. The BHM-20 also does not directly correspond with the DSM-5’s proposed suicidal behaviors disorder due to the items’ focus on suicidal thoughts instead of behaviors. The predictive utility of the BHM-20’s suicide risk items have not yet been empirically evaluated, although Bryan and Corso (2011) reported strong correlations between the BHM-20 item 10 with the PHQ-9 item 9 (.99) and the Suicide Behaviors Questionnaire-Revised (.61; Osman et al., 2001).
Summary
The BHM-20 is a relatively new tool that has only recently been implemented in primary care settings. Although its psychometric strength and utility as a suicide risk screening tool is limited, its implementation aligns with practice recommendations and conceptual models of suicide and risk management. In addition, its use as a universal screening tool has been found to improve the detection of acutely suicidal patients as compared to the indicated screening approach that characterizes the primary care system. Because of its transdiagnostic construction, the BHM-20 may be helpful for quantifying and tracking outcomes that are not specific to any particular diagnosis or condition.

Suicide Cognitions Scale
The Suicide Cognitions Scale (SCS) is an 18-item scale that assesses thoughts and beliefs that are common among individuals who have made suicide attempts (Bryan et al., 2014). Originally constructed for use as a psychotherapy process measure to aid clinicians in identifying and targeting suicidogenic beliefs that contribute to long-term vulnerability for suicidal behavior, the scale's items were developed based on statements commonly made by suicidal patients in treatment. Based on the fluid vulnerability theory of suicide, the scale aims at measuring core beliefs that align with the stable dimension of suicide risk as opposed to the acute dimension of suicide risk. Factor analytic work on the scale in two outpatient psychotherapy samples initially identified two subscales (Bryan et al., 2014): unlovability, which assesses beliefs about worthlessness and self-hatred, and unbearability, which assesses perceptions of one's capacity to tolerate distress. Subsequent work in psychiatric inpatient (Ellis & Rufino, 2015) and chronic pain (Bryan et al., in press) samples have identified three subscales: unlovability, unbearability, and unsolvability, the latter of which assesses perceptions that one is unable to effectively solve problems in life. Recent psychometric analyses have also led to a reduced scale (nine items) that retains the scale's performance (i.e., the Suicide Cognitions Scale-Short Form, or SCS-S; Bryan et al., in press). Although the SCS was not developed as a suicide risk screening tool, its incremental concurrent validity and ability to predict future suicidal behavior relative to other common risk factors and scales (Bryan et al., 2014) has led to new research examining its utility for screening purposes across healthcare settings including primary care. The scale is therefore currently being tested in a large, multisite primary care study.

Administration and Scoring
The SCS is a self-report questionnaire that directs patients to rate the extent to which they agree with 18 statements on a Likert scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). Scores are summed for each of the SCS’s three subscales, with higher scores indicating more severe suicidogenic beliefs and greater likelihood for engaging in suicidal behavior. Response patterns on SCS items tend to have severe positive skew, similar to distributional properties of other suicide-specific variables, with increasing scores providing incremental information about underlying suicide risk (Bryan et al., in press), suggesting that even low scores have clinical meaningfulness. Two items explicitly include the word “suicide” (e.g., “Suicide is the only way to solve my problems” and “Suicide is the only way to end this pain”), but when these items are removed, thereby reducing the SCS’s face validity and potential vulnerability to motivated responding, its performance is unaffected (Bryan et al., in press). The scale therefore holds promise as a method for identifying high-risk patients who are unwilling to disclose suicidal thoughts or intentions to healthcare providers. The scale typically takes less than five minutes to complete, and can be administered via paper survey or electronic administration. The scale is currently in the public domain.
Available Norms

Over 70% of nonpsychiatric medical patients endorse a score of 1 (“strongly disagree”) on most SCS items, and over 80% endorse a score of 1 on over half of its items (Bryan et al., in press). Item scores therefore tend to be very low, and range from 1.12 (SD = 0.47) to 1.99 (SD = 1.32). Mean subscale scores have been found to significantly differ across patient subgroups in clinically meaningful ways: patients who have attempted suicide score the highest (mean >13), followed by patients with a history of suicide ideation (means ranging from 10 to 12) and patients with no history of suicidal thoughts or behaviors (means ranging from 7 to 8; Bryan et al., 2014). In addition, patients with a history of nonsuicidal self-injury have mean scores that do not differ from controls, suggesting the SCS is specific to suicidal forms of self-injury. No differences between men and women have been identified or reported on the SCS.

General Strategy for Interpreting Results

In general, higher scores on the SCS indicate more severe suicide risk and increased likelihood for making a suicide attempt during the next two years, because this reflects a stronger underlying suicidal belief system. From a risk management perspective, higher scores on the unlovability scale (e.g., “There is nothing redeeming about me”) may suggest interventions that target interpersonal relationships and social support; higher scores on the unbearability scale (e.g., “I can’t stand this pain anymore”) might suggest emotion regulation or distress tolerance skills training; and higher scores on the unsolvability scale (e.g., “Nothing can help solve my problems”) might suggest problem-solving skills training.

Use as a Screener for Assessing Suicide Risk

The SCS’s primary strengths include its considerable specificity to suicide risk despite its minimal emphasis on acute suicide ideation, intent, and psychological symptoms. The SCS might therefore be useful for detecting at-risk patients who do not disclose suicide ideation, whether due to the desire to conceal suicidal intent or the absence of acute suicidal intent at the moment of screening. As noted in the previous section, the SCS may also provide clues for indicated management steps and interventions. Because of the newness of the SCS, it is not yet known how the scale should be optimally used in primary care (i.e., universal versus indicated use); this issue is currently being researched in a multisite study.

Relevant Psychometric Support

The SCS does not align with any DSM-5 diagnosis, to include the proposed suicidal behaviors disorder, due to its focus on measuring beliefs and schemas that theoretically contribute to long-term vulnerability to suicidal behavior. The SCS has been examined in four separate clinical samples: active duty military personnel recently discharged from an inpatient unit for suicide risk (Bryan et al., 2014), active duty military personnel in outpatient psychotherapy (Bryan et al., 2014), nonmilitary psychiatric inpatients (Ellis & Rufino, 2015), and nonpsychiatric chronic pain patients (Bryan et al., in press). Across all four samples, the SCS predicted severity of concurrent suicide ideation beyond the effects of other robust risk factors (e.g., depression, hopelessness). In the only prospective study published to date, the SCS predicted suicide attempts during a two-year follow-up period among recently discharged military patients (Bryan et al., 2014). Cutoff scores have not been developed or investigated; sensitivity and specificity are therefore unknown.

Summary

The SCS is a new assessment instrument that is still undergoing psychometric evaluation and development, but holds considerable promise as a suicide risk screening tool across healthcare settings.
including primary care. Additional research is needed to improve the scale's practicality for primary care and to determine how it compares to other screening tools currently in widespread use, such as the PHQ-9.

**IMPLEMENTING SUICIDE RISK SCREENING METHODS IN PRIMARY CARE**

The specific manner in which suicide risk screening tools are selected and implemented will necessarily differ across different primary care clinics based on the unique characteristics and needs of each clinic. For example, different tools and screening methods might be selected for primary care clinics with embedded behavioral healthcare providers as compared to clinics without such providers. Likewise, volume of patient flow might dictate how and when suicide risk screening is conducted. Another consideration relates to the costs of various screening methods.

**Barriers to Implementation**

Because primary care clinics differ in important ways from other healthcare specialties (e.g., high volume of patients, brief appointment durations), implementing suicide risk screening and assessment procedures in primary care can be uniquely challenging. Commonly identified barriers to suicide risk assessment among physicians include limited time, concerns about privacy availability, patient inability or unwillingness to participate in the assessment, and communication barriers with other providers (Petrik, Gutierrez, Berlin, & Sanders, 2015). Results of other surveys have similarly found that time constraints, limited access to mental health resources, and insufficient training in responding to suicide ideation serve as leading barriers to effective suicide risk assessment in primary care (Baraff, Janowicz, & Asarnow, 2006; Betz et al., 2013; Chisholm, Weaver, Whenmouth, & Giles, 2011). These barriers, which can vary across individual clinics, may explain why adherence to suicide risk assessment protocols varies considerably across clinics. The proportion of patients who screen positive for suicide risk and subsequently receive follow-up contact therefore ranges from 36% to 90% (Dobscha et al., 2014; Hooper et al., 2012).

**Provider Variables**

Provider-related factors have been shown to have a stronger impact on the likelihood of completing a suicide risk assessment than patient factors (Hooper et al., 2012). For example, a recent study found that young, Caucasian physicians were more likely to screen for suicide risk than older, non-Caucasian primary care physicians (Hooper et al., 2012). In addition, a physician's specialty may influence their perceptions about when a suicide risk assessment is indicated (Williams et al., 1999). It is common for clinicians and professionals who do not work with suicidal patients on a regular basis to feel uneasy and be unsure of how to appropriately react when faced with an at-risk individual. Many physicians, for instance, fear that asking about suicidal thoughts and behaviors may trigger a suicidal episode or a suicide attempt (Stoppe, Saandholzer, Huppertz, Duwe, & Staedt, 1999), a persistent myth that leaves physicians unaware of how best to proceed with at-risk patients as well as increasing their legal liability. Contrary to this perspective, research suggests that thorough suicide risk assessments may actually prevent suicidal behavior by connecting the at-risk patient with the correct services (Schulberg et al., 2004). Studies further suggest that suicide risk screening is actually associated with decreased suicide risk over time, especially among individuals with more suicide risk factors (Gould et al., 2005).
Although primary care physicians often feel that they are unable to effectively communicate with at-risk patients, research suggests their performance is typically acceptable, thereby enabling patients to comfortably disclose suicide ideation (Vannoy et al., 2010; Vannoy, Tai-Seale, Duberstein, Eaton, & Cook, 2011). Even when they are able to effectively conduct a suicide risk assessment, many physicians remain unsure about the appropriate “next steps” for clinical decision-making (e.g., treatment planning). Taken together, these results suggest that additional training, support, and resources are needed to ensure that primary care providers are adequately informed of existing protocols for when an assessment is warranted, do not fear engaging a patient in a risk assessment, and are aware of appropriate treatments and/or next steps if a patient is suicidal. These issues speak to the benefits of routine suicide risk screening in primary care, especially when they are combined with prompts, templates, and decision-making aids (Petrik et al., 2015). In short, routine use of suicide risk screening tools and measures help to increase providers’ familiarity and comfort with suicide risk screening and assessment, and can enhance adherence to suicide risk assessment protocols.

**Patient Variables**

Little is known about patients’ perspectives on suicide risk screening and assessments, although preliminary research suggests that at least some patient groups (e.g., veterans) acknowledge its importance (Ganzini et al., 2013). Studies focused on patient-related barriers to endorsing mental health complaints indicate that patients are unlikely to volunteer suicide-related thoughts or intentions to their physicians without prompting (Ganzini et al., 2013; Isometsa et al., 1995). Older adults, especially men, may be less likely to disclose suicidal thoughts (Schulberg et al., 2004); men, overall, are less likely to acknowledge any physical or mental struggles to a professional (Schulberg et al., 2004). Czyz and colleagues (2013) have additionally reported that individuals with more severe suicide ideation may be less likely to seek help than those with less severe suicide ideation. These researchers found that stigma, self-reliance, concerns about confidentiality, lack of time, cost, negative experiences, uncertainty about benefit, and preference to reach out to family and peers are among the most commonly cited reasons for not seeking help for suicidal thoughts.

**Cultural Variables**

Suicide risk assessment protocols often ignore cultural differences, which may lead to improper management of patients (Wendler & Matthews, 2006). Because the interpersonal dynamics between a suicidal patient and his or her clinician is believed to be a critical determinant of their willingness to disclose suicidal thoughts and intentions, cultural matters could influence how a patient responds to suicide risk screening and assessment. Culture can also influence a patient’s understanding of suicide-related terms and language, their moral and social beliefs about suicide, and their willingness to honestly disclose this sensitive information (see Chu et al., 2013 for a review). Ethnic minorities, for instance, are less likely than Caucasians to disclose suicidal thoughts (Morrison & Downey, 2000). Suicide risk assessment should therefore be conducted in a thoughtful and compassionate manner that reduces shame and embarrassment (Chu et al., 2013). Contrast, for example, the following two ways to ask about suicide ideation:

1. “You’re not thinking about suicide, are you?”
2. “It’s fairly common for individuals with similar problems and symptoms to also think about suicide or consider how they might kill themselves. I’m wondering if you’ve had any thoughts like this as well, even if only a little bit?”

The first question is worded in a manner that is unlikely to elicit an honest disclosure from a patient who may feel uncomfortable endorsing such thoughts. The second question, by comparison, is
worded in a manner that a positive disclosure of suicidal thoughts would be considered acceptable and possibly even expected.

Similar issues have been raised with respect to military personnel and veterans, a high-risk subgroup with distinct cultural norms and values that depart from the US general population (e.g., mental toughness, self-sufficiency, emotional stoicism). Ganzini and colleagues (2013) have posited that system-wide changes that establish a climate of trust may result in more honest disclosures of suicidal thoughts and intentions by military veterans. Although trust is an important factor across diverse cultural groups, trust may play a more significant role for military personnel or other careers in which disclosure of mental health symptoms or problems could have an impact on their careers (e.g., law enforcement, first responders). Less face-valid suicide risk screening and assessment tools may therefore be particularly useful in these populations.

CASE STUDY

Jane was a 59-year-old, Caucasian female who was referred to the integrated primary care psychologist for evaluation subsequent to a new diagnosis of diabetes, which was related to obesity. Jane’s primary care provider wanted the psychologist to help Jane develop and follow a new diet and exercise plan, and to assess her motivation for regular glucose monitoring and insulin management. On top of these issues, the primary care provider wanted to screen Jane for possible depression given its known comorbidity with the condition and implications for negative diabetes-related outcomes. The primary care provider also noted that Jane “appeared sad” when he told her the diagnosis. Jane was escorted from the primary care provider’s exam room to the front desk to check in for a walk-in appointment with the psychologist. She was handed a clipboard with a paper copy of the BHM-20 and was asked to complete the scale while waiting for the psychologist.

When the psychologist called Jane back to his office, he quickly scanned her responses on the BHM-20 and noticed that she endorsed several symptoms of depression and anxiety, as well as problems in daily functioning. The psychologist also noted that she had skipped the item assessing severity of suicide ideation in the past two weeks. Once in his office, the psychologist began by noting that Jane had been referred for assistance with diabetes management, and asked Jane if that is what she understood to be the reason for their visit. Jane responded, “Yeah, that’s right; I guess so.” The psychologist conducted a brief functional assessment focused on diet, exercise, and motivation, then grabbed the clipboard and noted, “I see here that you’re having some difficulty with mood, energy, anxiety, and family life; can you tell me more about that?” Jane’s account of her symptoms suggested she had recently experienced the onset of a major depressive disorder. She additionally noted that the new diabetes diagnosis “hit me hard” and “makes me feel even worse.”

The psychologist then probed about the skipped item regarding suicide ideation on the BHM-20: “As I was looking over your scores, I also noticed that you skipped this item here that asks about thoughts of ending your life.” The psychologist handed the clipboard to Jane and pointed to the item. “I was wondering what your response to that item might be.” Jane remained silent and looked away from the psychologist. The following is a partial transcript of their subsequent conversation:

Psychologist: Given all that’s been going on lately, I could see how someone might have such thoughts. Thinking about suicide is actually common for many people experiencing the same stress you’ve been facing lately. I can also see why you might not want to talk about it, as it can sometimes be really hard to talk about these things.
Assessment of Suicidal Risk

Jane: I don’t know; I just don’t like to talk about it.

Psychologist: That makes sense. Given your answer, it sounds like you have been thinking about suicide, even if only a little bit.

Jane: Yeah, I have been.

Psychologist: If we were to use the scoring system on this page, how often would you say those thoughts have been in the past two weeks?

Jane: Probably “often.”

Psychologist: Okay, thanks for sharing that with me. Can you tell me more about what, specifically, you think about when you have these thoughts?

The psychologist subsequently conducted a full suicide risk assessment interview, consistent with the approach recommended by Bryan et al. (2009). During the course of this assessment, Jane reported that she had made five previous suicide attempts, in all cases via overdose on sleep medication. These suicide attempts typically occurred within the context of a major depressive episode, of which she had experienced several during her lifetime. Jane additionally reported that the diabetes diagnosis “was further proof about how disgusting I am; I got this disease because I’m fat and can’t even control my eating.” The psychologist was able to help Jane create a crisis response plan and recommended a referral to outpatient cognitive behavioral therapy. Jane was scheduled for a follow-up in one week due to her elevated risk status. Over the course of the next few appointments, Jane was able to identify a new outpatient psychologist and resumed psychological treatment focused on depression. The BHM-20 was administered at each of these appointments to track Jane’s risk over time; her overall scores, as well as her scores on the suicide risk item, gradually improved.

In this case study, Jane’s reason for referral was for a reason that might not seem to warrant suicide risk screening or assessment: diabetes management. However, because a universal suicide risk screening process was in place, the psychologist was able to identify a high-risk patient who might otherwise have been missed. Jane was experiencing suicidal thoughts on a regular basis, but was reluctant to disclose these thoughts on either the paper screener or in response to initial questioning. Nonetheless, the use of a suicide risk screening method enabled the psychologist to follow up in an appropriate and clinically useful way to elicit greater information about Jane’s level of risk, thereby enabling the development of a better risk management and treatment plan. In the absence of suicide risk screening, it is possible that Jane would have deteriorated over time and potentially made another suicide attempt.

**SUMMARY AND CONCLUSIONS**

Although our ability to accurately predict suicidal behavior remains limited, recent efforts focused on improving suicide risk screening and assessment tools are yielding promising new methods that will improve the options available to primary care providers. Despite current limitations, with the proper systems and tools in place, suicide risk screening and assessment can be relatively simple and effective in primary care. For example, the Department of Veterans Affairs and the Department of Defense have implemented universal suicide prevention screening and assessment in primary care that have led to early identification of at-risk patients who can be connected to effective treatments. Primary care clinicians should select the suicide risk screening tool that is most appropriate and practical for the needs of their specific clinical population. In general, technology should be used to the greatest extent possible to streamline the efficiency and acceptability of suicide risk screening methods (Boudreaux & Horowitz, 2014). Because of the noted limitations of an indicated approach to suicide risk
screening and assessment, suicide risk screening should be conducted across as wide a segment of the primary care population as possible, even with patient groups for whom it may seem unnecessary because they are not disclosing mental health symptoms or demonstrating clear signs of elevated suicide risk. Finally, primary care clinics should be prepared to respond appropriately to positive suicide screenings in a coordinated manner that ensures such patients receive more detailed assessment that guides appropriate risk management decisions and interventions.

REFERENCES


Anxiety disorders are among the most common presenting problems in primary care settings. In an epidemiological study of the general population, approximately 46% of respondents endorsed a lifetime history of a mental health disorder. Of these, anxiety disorders represented the most common category, with 12-month prevalence estimates of approximately 29% lifetime and 18%, respectively (Kessler & Wang, 2008). In a study of primary care clinics, nearly 20% of patients met current diagnostic criteria for at least one anxiety disorder (Kroenke, Spitzer, Williams, Monahan, & Löwe, 2007). Individuals with anxiety disorders are more likely to present to primary care clinics than any other treatment setting, including mental health specialty clinics (McDaniel & deGruy, 2014; Wittchen et al., 2002). For example, results of one study indicated that approximately one-third of patients with an anxiety disorder received care from a primary care provider within the past year (Stein et al., 2004). Further, research suggests that identification of anxiety disorders is suboptimal. According to one study, patients who were eventually diagnosed with an anxiety disorder had been seen in a primary care clinic an average of six times before being identified as having an anxiety disorder (Deacon, Lickel, & Abramowitz, 2008). Despite the high prevalence, detection and intervention for anxiety disorders remains low. Studies have found that the majority of patients presenting in primary care settings with an anxiety disorder did not receive appropriate psychotherapeutic or pharmacological treatment (e.g., Stein et al., 2004).

Taken together, the research literature indicates that anxiety disorders are extremely common, patients with anxiety disorders are likely to present in primary care clinics, and that the current standard of care is suboptimal for patients with anxiety disorders. Accordingly, primary care represents an ideal setting for improving the population-based detection and management of anxiety disorders. Assessing anxiety, however, can be difficult in this setting given the multifaceted nature of anxiety, fast-paced environment of primary care, and high comorbidity rates with medical illness (e.g., Roy-Byrne et al., 2008). One of the greatest advantages of integrating behavioral health providers into primary care, therefore, is the ability to assess for anxiety measures and subsequently provide appropriate evidence-based interventions.

The aim of the current chapter is to provide an overview of assessment measures for anxiety that are compatible with a primary care model. Physicians and psychologists working in primary care settings typically have more limited time with patients, compared to mental health practitioners working in traditional settings (Mori et al., 2003). Accordingly, assessment of anxiety in primary care needs to be performed in a manner congruent with the setting, which requires measures that are both efficient and valid. Furthermore, unlike in traditional outpatient mental health settings, providers across a variety of specialties may use these measures to help identify clinically significant anxiety, necessitating clear and easy scoring and interpretation. Mori and colleagues (2003) suggest the following criteria for evaluating the usefulness of psychological screening instrument in primary care: (1) ease of administration, (2) practicality of the information provided, (3) ability to identify commonly encountered psychological symptoms, and (4) accuracy of identification of these symptoms.
among various medical populations. Consistent with these guidelines, this chapter will focus on commonly used measures of anxiety that have demonstrated usefulness in a primary care setting.

The primary focus of this chapter will be on the seven-item Generalized Anxiety Disorder (GAD-7) questionnaire, with a secondary focus on three additional measures that are appropriate for a primary care setting and offer unique aspects to screening for anxiety, including the Overall Anxiety and Severity and Impairment Scale (OASIS), the Patient-Reported Outcomes Measurement Information System-Anxiety Test Bank (PROMIS Anxiety), and the Beck Anxiety Inventory-Primary Care (BAI-PC). Finally, this chapter offers a tertiary focus on additional assessment measures for specific forms of anxiety (e.g., panic disorder, social anxiety), which are included in a table at the end of this chapter for easy reference.

THE SEVEN-ITEM GENERAL ANXIETY DISORDER (GAD-7)

The GAD-7 is one of the most frequently used assessment measures for anxiety across primary care settings, and in fact, has been the standard of practice for assessing anxiety in primary care across the VA Healthcare System, a leader in primary care mental health integration, among other major organizations for the last several years. Although the GAD-7 was originally designed for use as a screening tool in primary care, it has gained substantial prominence in a variety of research studies, as evidenced by the original research article being cited over 2,050 times (results obtained via Google Scholar search, April 18, 2016). Specifically, the original authors Spitzer and colleagues (2006) recognized a paucity of anxiety measures used in primary care settings at the time because of their length, proprietary nature, lack of usefulness as a diagnostic and severity measure, and requirement of clinician administration rather than self-report. As such, the authors developed the GAD-7 as a brief measure for generalized anxiety disorder (GAD) to meet the fast-paced nature in primary care settings. The GAD-7 was constructed from GAD criteria in the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV; American Psychiatric Association, 1994), although it remains consistent with DSM-5 given that the criteria for GAD did not change in the fifth edition of the manual (American Psychiatric Association, 2013). The GAD-7 was designed as a self-report questionnaire that can be filled out by the patient directly or read out loud by the provider, and it assesses how bothered the patient has been by a spectrum of anxiety symptoms over the previous two weeks. Specifically, the items inquire about the degree to which the patient has been bothered by the following symptoms during the past two weeks: feeling nervous, anxious, or on edge; not being able to stop or control worrying; worrying too much about different things; having trouble relaxing; being so restless that it is hard to sit still; becoming easily annoyed or irritable; and feeling afraid as if something might happen. If patients endorse any of these symptoms, they then indicate the degree to which these problems made it difficult for them to work, take care of home responsibilities, or get along with people. The GAD-7 consists of seven items, with each symptom rated on a Likert-type scale with the following:

0 = Not at all
1 = Several days
2 = Over half the days
3 = Nearly every day.

Total scores can range from 0 to 21, with higher scores indicating higher levels of GAD severity. The GAD-7 is considered to be a simple tool that is easy to administer, without placing undue burden on the patient or clinician—an important quality of any measure intended for use in primary care.
Written administration is generally less than five minutes, whereas oral administration is generally 10 minutes. The general interpretation of scores is guided by the original anxiety disorder sample used in establishing test norms (Spitzer, Kroenke, Williams, & Lowe, 2006) and is outlined below:

0–4 = Minimal range
5–9 = Mild range
10–14 = Moderate range
15–21 = Severe range
General rule: >10 = Probable diagnosis of GAD; confirm by further evaluation

**Summary of Development**

In the original study (Spitzer et al., 2006), the GAD-7 was developed and validated based on responses from 2,739 patients (65% female; 80% White, non-Hispanic; 62% had attended some college; mean age = 47.4) across 15 primary care clinics located in 12 states. The initial item pool consisted of nine items that reflected all of the DSM-IV criteria for GAD and four items on the basis of review of existing anxiety scales. From this, a 13-item questionnaire was created that asked how often, during the previous two weeks, they were bothered by each symptom. The developers’ goal was to determine the number of items necessary to achieve good reliability and procedural, construct, and diagnostic criterion validity.

In the first phase of scale development \((n = 2,149)\), the goal was to select the scale items and cutoff scores to be used for making a GAD diagnosis. Before seeing their physicians, patients completed a four-page questionnaire that included the 13 items being tested for use in the GAD scale, as well as the Medical Outcomes Study Short-Form General Health Survey (SF-20; Stewart, Hays, & Ware, 1988), and either the 12-item anxiety subscale from the Symptom Checklist-90 (SCL-90; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974; first study phase only) or the Beck Anxiety Inventory (BAI; Beck, Epstein, Brown, & Steer, 1988; second study phase only). Depression was assessed with the eight-item Patient Health Questionnaire (PHQ; Kroenke, Spitzer, & Williams, 2001), which includes all items from the PHQ-9 except for the item about suicidal ideation.

In the first study phase, 1,654 subjects also agreed to a telephone interview, and of these, a random sample of 965 were interviewed within one week of their clinic visit by one of two mental health professionals—a PhD clinical psychologist or a senior psychiatric social worker. The mental health professionals were blinded to the results of the self-report research questionnaires. Interviews consisted of the GAD portion of the Structured Clinical Interview for DSM-IV (First, Spitzer, Williams, & Gibbon, 1995). The resulting DSM-IV GAD diagnosis was used as the criterion standard for assessing the validity of the new scale. The interview also included the 13 potential GAD scale items to test agreement between self-report and clinician administration. The best items for the GAD scale were selected by rank ordering the correlation of each item with the total 13-item scale score in the sample of 1,184 patients who did not undergo the clinical interview. The seven items of the GAD-7 had the highest rank correlations in the developmental sample.

In the second phase of development \((n = 591)\), the goal was to determine the scale’s test-retest reliability. In this phase, the 591 subjects who had completed the research questionnaire were sent a one-page questionnaire that consisted of the 13 potential GAD scale items. Of these, 236 subjects returned the completed one-page questionnaire with no or minimal missing data within a week of completing the research questionnaire at the clinic. In all, 2,982 subjects were approached and 2,739 (91.9%) completed the study. Psychometric findings of this sample and other samples reported elsewhere in the literature are discussed later in the “Basic Psychometric Information” section.
Available Norms

The GAD-7 was originally developed for use in primary care settings, although it has recently been studied in a variety of other populations given its brief format, simple scoring, and public availability (e.g., Beard & Bjorgvinsson, 2014). The GAD-7 has been validated for use as a screening tool and severity measure in primary care (e.g., Ruiz et al., 2011; Spitzer et al., 2006), general populations (e.g., Lowe et al., 2008), addictions treatment (e.g., Delgadillo et al., 2012), heterogeneous psychiatric samples (Beard & Bjorgvinsson, 2014; Kertz et al., 2013), acute psychiatric samples (Kertz et al., 2013), pregnant and postpartum samples (e.g., Barthel, Barkmann, Ehrhardt, & Bindt, 2014; Simpson, Glazer, Michalski, Steiner, & Frey, 2014; Zhong et al., 2015), people with epilepsy (e.g., Seo et al., 2014), Chinese people with epilepsy (Tong, An, McGonigal, Park, & Zhou, 2016), patients with migraine headaches (Seo & Park, 2015), patients with multiple sclerosis (Terrill, Hartoonian, Beier, Salem, & Alschuler, 2015), elderly people (e.g., Wild et al., 2014), Canadian samples (Henderson, Antony, & Koerner, 2014), and Malaysian samples (Sidik, Arroll, & Goodyear-Smith, 2012). The GAD-7 has also been translated into several languages (Chinese; Zeng et al., 2013; Dutch; Donker, van Straten, Marks, & Cuijpers, 2011; Finnish; Kujanpaa et al., 2014; Portuguese; Sousa et al., 2015; Turkish; Konikan, Senormanci, Guclu, Aydin, & Sungur, 2013; Spanish; Garcia-Campayo et al., 2010; British Sign Language; Rogers et al., 2012). Lastly, the GAD-7 has also been validated for web-based (e.g., Donker et al., 2011) and telephone use (Ryan, Bailey, Fearon, & King, 2013).

Basic Psychometric Information

The GAD-7 has good psychometric properties based on a number of published studies on various clinical and nonclinical samples. The following is a review of GAD-7 reliability and validity research.

Internal Consistency Reliability

Studies reporting the GAD-7’s internal consistency have been highly favorable. Reliability coefficients reported (Cronbach’s coefficient alpha) have generally ranged from .82 to .92, indicating that at least 82% of variability in GAD-7 scores among test takers is due to true score variability. In the original test sample, Spitzer et al. (2006) reported an excellent alpha of .92. Similarly, Zhong and colleagues (2015) reported an alpha of .89 with their sample of 2,978 women who attended their first prenatal care visit. Delgadillo et al. (2012) reported robust internal consistency with a Cronbach’s alpha value of .91 in their sample of 102 patients engaged in outpatient addictions treatment. Kertz et al. (2013) reported an excellent alpha value of .91 for their sample of 232 acute psychiatric patients. Lowe et al. (2008) reported an alpha value of .89 in their sample of 5,036 individuals randomly selected from the general population. Wild et al. (2014) reported an alpha of .82 among a subsample of 438 elderly patients. Beard & Bjorgvinsson (2014) reported an alpha coefficient of .88 for their heterogeneous sample of 1,082 participants receiving brief, intensive cognitive behavioral therapy (CBT) treatment in a partial hospital setting. Seo et al. (2014) reported an alpha value of .92 for their sample of 243 patients with epilepsy. Tong et al. (2016) reported an alpha value of .89 among their sample of 213 Chinese patients with epilepsy. Zeng et al. (2013) reported an alpha value of .91 for the Chinese version of the GAD-7 among their sample of 2011 outpatients selected from the traditional Chinese internal department. Dear and colleagues (2011) reported an alpha value of .79–.91 for their sample of 195 adults who met DSM-IV criteria for GAD and who participated in two randomized treatment controlled trials.

In a study examining the psychometric properties of the English and Spanish language versions of the GAD-7 for Hispanic Americans in the United States (n = 436), Mills et al. (2014) reported a strong alpha value of .93 for the total sample. Rogers et al. (2012) reported an alpha value of 0.88 for the
British Sign Language version of the GAD-7 in their sample of deaf signing individuals in the United Kingdom \((n = 113)\). Lastly, Henderson et al. (2014) reported excellent internal consistency reliability with an alpha level of .90 in a Canadian student/community sample \((n = 41)\).

**Test-Retest Reliability**

GAD-7 test scores have been demonstrated to remain stable over time. Because it is a self-report measure of symptoms within a two-week time frame, the GAD-7 tends to be better suited as an instrument assessing anxiety states rather than anxiety as a trait that is stable over long time frames. In the original article, the authors reported a test-retest reliability coefficient of .83 over a seven-day period (Spitzer et al., 2006). Similarly, Sousa et al. (2015) reported seven-day test-retest coefficients between .60 and .93 for individual items of the Portuguese version of the GAD-7 among their sample of 100 patients diagnosed with GAD. Konikan et al. (2013) examined the test-retest reliability of the Turkish version of the GAD-7 by retesting 39 of the 110 patients diagnosed with GAD in their study 3 weeks later. Results of their analysis showed no significant differences between the two measurements in this comparison, suggesting that the Turkish version of the GAD-7 was consistent across time points. Delgadillo et al. (2012) also found the GAD-7 to have adequate temporal stability at repeated measurements (intra-class correlation = .85) among their sample of 102 patients engaged in outpatient addictions treatment. In summary, the GAD-7 demonstrates good stability over time, but is also sensitive to changes in reported anxiety symptoms, an important factor when considering the GAD-7 as a measure of outcome.

**Item-to-Total Score Correlations**

With regard to other indexes of reliability, Spitzer et al. (2006) reported corrected item-to-total score correlations ranging from .75 to .85 in the original test sample. Lowe et al. (2008) reported corrected item-to-total score correlations ranging from .63 to .74 among their sample of 5,030 individuals from the general population. Likewise, Zhong et al. (2015) reported item-to-total score correlations ranging from .67 to .72 among their sample of 2,978 pregnant women. Tong et al. (2016) reported corrected item-to-total score correlations ranging from .67 to .82 among their sample of Chinese individuals with epilepsy. Dear and colleagues (2011) reported corrected item-to-total correlations ranging from .38 to .68 at pretreatment and between .60 to .84 at posttreatment among their sample of 195 individuals diagnosed with GAD. Lastly, Kertz et al. (2013) reported corrected item-to-total score correlations ranging from .64 to .81 among their sample of acute psychiatric patients. As such, the available data suggest that the GAD-7 has good item-scale correlation and tends to measure a consistent cluster, or anxiety construct, with regularity across clinical and nonclinical samples.

**Convergent and Discriminant Validity**

If the GAD-7 is a valid measure of anxiety, it should correlate more strongly with tests of anxiety (convergent validity) than with tests of distinct constructs (discriminant validity). By and large, this has been the case. Spitzer et al. (2006) reported strong GAD-7 correlations with each of two anxiety scales: the Beck Anxiety Inventory (BAI; \(r = 0.72\)) and the anxiety subscale of the Symptom Checklist-90 (SCL-90; \(r = 0.74\)). Consistent with results of previous studies of anxiety and depression (e.g., Carter, Wittchen, Pfister, & Kessler, 2001), Spitzer et al. (2006) found that the GAD-7 and SCL-90 anxiety scales also strongly correlated with a measure of depression, the Patient Health Questionnaire-8 (PHQ-8; \(r = 0.75\) and \(r = 0.74\), respectively). However, these researchers concluded that measuring anxiety and depression was complementary rather than duplicative. Specifically, in the original sample, of the 2,114 patients who completed the GAD-7 and PHQ-8, there were 1,877 (88.8%) patients with neither high anxiety nor high depression, 99 (4.68%) with high anxiety only, 68 (3.2%) with...
high depression only, and 70 (3.31%) with high anxiety and high depression. As such, more than half (99/169) of patients with high anxiety scores did not have high depression scores (Spritzer et al., 2006). This suggests that the GAD-7 and PHQ-8 measure separate phenomena despite correlated total scores.

Other studies are relatively consistent with these findings. Among their sample of 513 patients with multiple sclerosis, Terrill et al. (2015) found that the GAD-7 was significantly correlated with the anxiety subscale of the Hospital Anxiety and Depression Scale (HADS-A; r = 0.70), and was also significantly correlated with a measure of depression (PHQ-9, r = 0.70), with approximately half of the participants meeting the cutoff criteria for GAD also meeting the criteria for depression. Wild et al. (2014) found a strong correlation between the GAD-7 and the mental component scores of the 12-item Short-Form General Health Survey (SF-12; r = −.48), although a moderate but still significant correlation between the GAD-7 and the physical component scores of the SF-12 (r = −.20) among elderly patients. Barthel et al. (2014) found significant correlations between the GAD-7 and the PHQ-9 among mothers and children in Côte d’Ivoire (r = .41) and Ghana (r = .39), as well as between the GAD-7 and a measure of disability, the World Health Organization Disability Assessment Scale (WHO-DAS II; r = .28 and r = .34, respectively). Among their sample of 103 patients in addiction treatment, Delgadillo and colleagues (2012) found a strong correlation between the GAD-7 and the gold standard Revised Clinical Interview Schedule (CIS-R; r = .76), indicating strong convergent validity. Conversely, these researchers found a weak association between the GAD-7 and the Severity of Dependence Scale (SDS; r = .39), indicative of acceptable discriminant validity (Delgadillo et al., 2012).

Taken together, these results indicate that while convergent validity is well established, future studies with better measures of discriminant validity are needed.

The GAD-7 was not designed to discriminate between psychiatric diagnoses, but there is evidence that individuals with anxiety disorders will score significantly higher on the GAD-7 than individuals without an anxiety disorder. Beard and Bjorgvinsson (2014) found that patients with one of the four anxiety disorders (i.e., GAD, PTSD, panic disorder, social anxiety disorder) scored significantly higher on the GAD-7 than patients without those anxiety disorders, t(1,080) =11.32, p < .001. Indeed, the capacity of the GAD-7 to detect other anxiety disorders (excluding specific phobias) has long been established (e.g., Kroenke et al., 2007).

**Factorial Validity**

The GAD-7 demonstrates overall good factorial validity. The bulk of factor analytic data gleaned to date suggests that the GAD-7 is a solid, internally consistent measure of anxiety and not confounded with measurement of other affective states. There are two broad conclusions drawn regarding factor analytic research of the GAD-7:
1. Most factor analyses of the GAD-7 yield one-factor solutions.
2. Factors uncovered by analyses generally do not correlate with factors reflecting depressive symptoms when depression measures are included in analyses.

These findings are briefly reviewed.

In the original validation sample, factor analysis that included the GAD-7 anxiety items and the PHQ-8 depression items confirmed two distinct dimensions, with all depression items having the highest factor loadings on one factor (0.58–0.75) and all anxiety items having the highest factor loadings on the second factor (0.69–0.81; Spitzer et al. 2006). Lowe et al. (2008) tested for the unidimensionality of the GAD-7 among their general population sample and found that all seven items were specified as indicators of a single factor, with high factor loadings ranging from 0.76 to 0.90 with a mean on 0.83. Results of their study also indicated gender invariance of the GAD-7 factor structure as well as age invariance of the GAD-7 factor structure at the level of a congeneric measure (Lowe et al., 2008). Likewise, Garcia-Campayo et al. (2010) corroborated the original structure and the existence of one single factor where all items loaded for the Spanish Version of the GAD-7. Confirmatory factor analysis also suggested that the unidimensional factor structure of the GAD-7 adequately matched observed face-to-face and telephone data among a large (n = 23,672) heterogeneous sample in London (Ryan et al., 2013). Zhong et al. (2015) reported a one-factor structure for the GAD-7 as confirmed by exploratory and confirmatory factor analyses among their sample of pregnant women. Conversely, Beard and Bjorgvinsson (2014) examined the factor structure of the GAD-7 among a heterogeneous psychiatric sample and found that a one-factor structure did not fit the data as well, in that items 4, 5, and 6 warranted a separate factor, suggesting that these items may reflect a somatic tension/autonomic arousal factor compared to the other items, which are more cognitive in nature. Overall, more research is needed to see if a one-factor solution fits best across various samples.

**Use as a Screener**

As a screening instrument for the detection of anxiety disorders in primary care, the GAD-7 is an unparalleled practical instrument. It is brief, easily and quickly scored, and there is a host of research data (most of which is discussed in this chapter) that supports its validity and reliability. It allows for immediate feedback to both patients and practitioners. One point of caution that should be kept in mind, however, is that the GAD-7 is a face-valid self-report instrument, making the validity of its results subject to the patient’s motivation to honestly and accurately report their experience. As such, the accuracy with which the GAD-7 will detect anxiety disorders in primary care screening is contingent on practitioner knowledge of patient motives for self-report, and his or her knowledge of DSM-5 diagnostic criteria for anxiety disorders.

According to the original validation study, using the threshold score of 10, the GAD-7 has sensitivity of 89% and specificity of 82% for detecting GAD, with higher scores associated with worsening functional impairment (Spitzer et al., 2006). A clinical cutoff of 8, however, has also recently been proposed for the GAD-7 on clinical significance grounds and is being used in large-scale treatment studies (e.g., Clark et al., 2009; Dear et al., 2011). Indeed, studies investigating the psychometric properties of the GAD-7 among different populations have found that different cutoff scores may be warranted. Among their sample of patients seeking outpatient addictions treatment, for example, Delgadillo et al. (2012) found that the GAD-7 as a case-finding tool for anxiety disorders had a significant area under the curve (AUC) value of .88 (.81–.95) and the best trade-off between sensitivity (80%) and specificity (86%) at a cut score of 9 points and above. Among their sample of patients with migraine headaches, Seo and Park (2015) found that at a cutoff score of 5, the GAD-7 had a sensitivity
of 78.1%, a specificity of 74.6%, a positive predictive value (PPV) of 46.3%, and a negative predictive value (NPV) of 92.4%. These researchers acknowledge their small sample size, however, and how this may have caused a difference in the sensitivity and specificity compared with other studies. Likewise, Kertz and colleagues (2013) found that a cutoff score of 10 resulted in good sensitivity (83%) but poor specificity (46%) and a high false positive rate (.54) among their sample of acute psychiatric patients, although they acknowledge their limited sample size as well. In their sample of pregnant ($n = 155$) and postpartum women ($n = 85$), Simpson et al. (2014) reported good sensitivity (76.0%), poor specificity (51.5%), poor PPV (41.6%), good NPV (82.5%), and poor kappa (.22) as a screening tool for GAD when using the previously established cutoff score of 10. A cutoff score of 13 yielded the best fitting model, with a sensitivity of 61.3% and specificity of 72.7%, with a PPV of 50.5%, NPV of 80.5%, and kappa of .32 (Simpson et al., 2014). Conversely, among their sample of Chinese patients with epilepsy, Tong et al. (2016) found that at a cutoff score of 6, the GAD-7 achieved the largest Youden index of .854 with a sensitivity of 94%, a specificity of 91.4%, a PPV of 77%, and a NPV of 98%. Taken together, it appears that it is important to consider the population one is working with when determining the appropriate cutoff score when using the GAD-7 to screen for anxiety.

A lower cutoff point appears to be especially warranted when using the GAD-7 to screen for GAD in geriatric populations. Wild and colleagues (2014) evaluated the validity of the GAD-7 and GAD-2 for detecting GAD in elderly people. Their sample included 438 elderly persons ages 58–82, wherein the GAD-7 was administered to each person as part of a home visit and the Structured Clinical Interview for DSM-IV (SCID) was subsequently administered over the telephone by a blind interviewer. Using the SCID interview diagnosis of GAD as the criterion standard to determine sensitivity and specificity, a cutoff score of 5 on the GAD-7 appeared to be optimal for detecting GAD, with a sensitivity of .63 and specificity of .90. Overall, Wild et al.’s (2014) study suggested that the recommended cutoff scores for the GAD-7 for detecting GAD should be lowered for the elderly general population. Correspondingly, Vasiliadis, Chudzinski, Gontijo-Guerra, and Preville (2015) also set out to establish an appropriate cutoff score for detecting anxiety in an older adult population using the GAD-7 and found that when screening for anxiety with the GAD-7, receiver operating characteristic (ROC) analysis yielded an AUC of .695; a cutoff of 5 was found to balance sensitivity (.709) and specificity (0.568). No significant differences were found between subgroups of age and gender. The GAD-7 was able to discriminate between cases and noncases when screening for anxiety in an older adult population of primary care service users (Vasiliadis et al., 2015).

Although originally developed for GAD, the GAD-7 has also proved to have good sensitivity and specificity as a screener for panic disorder (sensitivity 74%, specificity 81%), social anxiety disorder (sensitivity 72%, specificity 80%), and posttraumatic stress disorder (PTSD; sensitivity 66%, specificity 81%), with area under the curve by ROC greatest for GAD (.91), but also quite good for panic disorder (.85), social anxiety disorder (.83), and PTSD (.83; Kroenke et al., 2007; Kroenke, Spitzer, Williams, & Lowe, 2010). In general, a cutoff score of 8 appears to be optimum for identifying symptoms of panic disorder, social anxiety disorder, or PTSD (Kroenke et al., 2007). According to Beard and Bjorgvinsson (2014), it should be noted that the GAD-7 performed worst as a screener in the social anxiety disorder group and that these individuals showed less improvement as measured by the GAD-7. The authors speculate that this may be due to the GAD-7’s focus on physical symptoms of anxiety, which may not be as relevant to individuals with social anxiety disorder because many of these individuals may not regularly experience these symptoms unless they are in a social situation (Beard & Bjorgvinsson, 2014).

**GAD-2**

A simpler two-item scale, the GAD-2, is also commonly used as a screener in primary care. Although it has nearly the same accuracy as the GAD-7, the latter provides additional information that can guide management. Most researchers have found that a cutoff score of 2 yields the best trade-off for
sensitivity and specificity. For example, at a cutoff score of 2, the GAD-2 had a sensitivity of 86% and a specificity of 83% for detecting GAD (Kroenke et al., 2007). Likewise, among their geriatric sample, Wild et al. (2014) found that the optimal cutoff score for the GAD-2 was 2 or greater, with a sensitivity of 0.67 and a specificity of 0.9. Overall, as with all screening instruments, the GAD-7 and GAD-2 detect only probable psychiatric diagnoses; positive scores should be followed by more extensive interviewing (e.g., using the DSM-5 criteria) accompanied by appropriate management and referral as described in the next section.

Guidelines for Decision-Making

Use With Other Instruments/Procedures

If problematic anxiety or an anxiety disorder is expected (e.g., a GAD-7 score >10), there are several lines of clinical procedure to take. First, the primary care physician or other clinician should inquire about a history of anxiety disorders or treatment by other physicians, psychologists, or other healthcare practitioners (e.g., see Ferguson, 2000). This question will narrow down the diagnostic possibilities. If the patient reports no history of anxiety treatment, the clinician should ask the patient to describe the anxiety problem. Clinicians can assist the patient by using the questions presented in Table 12.1. These questions are aimed at five general problem areas associated with various anxiety

<table>
<thead>
<tr>
<th>Anxiety Problem Area</th>
<th>Clinical Question(s)</th>
</tr>
</thead>
</table>
| Panic Disorder                       | • Have you had sudden unexpected periods of intense physical symptoms and a feeling you were going to die? Have you experienced symptoms such as a pounding heart, shortness of breath, and dizziness? I’m talking about a panic or anxiety attack.  
• How long did the symptoms take to become intense? [Look for <10 mins.]  
• Do these attacks come on out of the blue, in situations where you wouldn’t expect for them to occur? |
| Agoraphobia                           | • Do you experience intense fear in certain situations, including when using public transportation, being in open (or enclosed) spaces, or just being outside of the home alone?  
• Have you changed your behavior because of this fear, such as avoiding places outside of your home or needing someone to accompany you when you leave your home? |
| Generalized Anxiety Disorder         | • Are you a worrier?  
• Do you worry excessively, or more than you believe you need to?  
• About what percent of the day are you worried, nervous, or tense? (>50%?)  
• Do you have difficulty controlling the worry once it starts? |
| Social Anxiety Disorder              | • Do you ever have strong feelings of anxiety or shyness in social situations where you could be evaluated by others?  
• Do you have strong feelings of anxiety in situations where you believe you could do something to embarrass or humiliate yourself?  
• Have you changed your behavior because of this anxiety, such as avoiding social situations altogether? |
| Substance-Induced Anxiety Disorder   | • How often do you use substances including alcohol, tobacco, or marijuana? Do you use any other illicit substances? Do you ever take more of your prescription medications than prescribed?  
• Does your anxiety (or panic attacks) develop during or soon after exposure to a substance?  
• Has there been a period of time where you experienced anxiety (or panic attacks) when not using substances? How long was this period of sobriety? |
## Table 12.2 Anxiety Self-Report Measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Primary Reference</th>
<th>What It Measures</th>
<th>Description</th>
<th>Psychometrics</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anxiety Scale of the Four-Dimensional Symptoms Questionnaire (4DSQ)</strong></td>
<td>Terluin et al. (2006); Terluin, Van Rhenen, Schaufeli, and Deltaan (2004)</td>
<td>Irrational fears, panic, avoidance</td>
<td>50 items distributed over 4 scales. Anxiety scale contains 12 items rated on a 5-point scale (no to very often or consistently) Total administration time: 5–10 minutes</td>
<td>Generally α &gt; .80 Good factorial, criterion and concurrent validity</td>
<td>Longer time to administer Available for public use</td>
</tr>
<tr>
<td><strong>Anxiety Sensitivity Index-3 (ASI-3)</strong></td>
<td>Reiss, Peterson, Gursky, Richard, and McNally (1986); Taylor et al. (2007)</td>
<td>Fears of anxiety symptoms</td>
<td>18-item version of the original ASI (Reiss et al., 1986) rated on a 5-point Likert scale ranging from 0 (very little) to 4 (very much) Administration time: 2–3 minutes</td>
<td>Generally α ≥ .80 Good convergent, discriminant, and criterion-related validity</td>
<td>Superior to its predecessors (ASI and ASI-R), as it measures the construct more precisely and has higher reliability and construct validity Internal structure stable across diverse samples Cost to purchase</td>
</tr>
<tr>
<td><strong>Beck Anxiety Inventory (BAI)</strong></td>
<td>Beck and Steer (1993); Beck, Epstein, Brown, and Steer (1988)</td>
<td>Cognitive and somatic symptoms of anxiety</td>
<td>21 items rated on a 0 (not at all) to 3 (severely) scale Administration time: 5–10 minutes</td>
<td>α ranges .90 to .94; Good test-retest reliability (r = .75–.83); moderately correlated with anxiety (r = .36-.69) Good content, construct, and factorial validity</td>
<td>Extensively studied across a variety of populations Longer administration time Cost and purchase is restricted to qualified individuals</td>
</tr>
<tr>
<td><strong>Beck Anxiety Inventory-Primary Care (BAI-PC)</strong></td>
<td>Beck, Steer, Ball, Ciervo, and Kabat (1997)</td>
<td>Subjective dimension of anxiety</td>
<td>7 items rated on a 4 point scale (0 = not at all, 3 = severely) Administration time: 2–5 minutes</td>
<td>α = .90 Item-total correlations .59–.81, highly correlated with the full version of the measure (r = .74)</td>
<td>Derived from the 21-item BAI that has been extensively studied Does not assess avoidance behavior or functional impairment Somatic symptoms of anxiety not included to screen out medical patients without anxiety and increase specificity Cost to purchase</td>
</tr>
<tr>
<td>Measure Primary Reference</td>
<td>What It Measures</td>
<td>Description</td>
<td>Psychometrics Considerations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>-----------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Daily Assessment of Symptoms Anxiety (DAS-A)</em>&lt;br&gt;Morlock et al. (2008)</td>
<td>Detects reduction of anxiety symptoms in patients with GAD during the first week of treatment</td>
<td>8 items on a 0 (not at all) to 10 (extremely) point scale&lt;br&gt;Administration time: 2–5 minutes</td>
<td>α ranges .77 to .91&lt;br&gt;Good test-retest reliability (.77 to .91), construct validity, predictive validity&lt;br&gt;DAS-A items are strongly interrelated and support a one-dimensional scoring structure&lt;br&gt;α = .71–.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fear Questionnaire (FQ)</strong>&lt;br&gt;Mark and Mathews (1979)</td>
<td>Phobia, related anxiety</td>
<td>24-items rated on a 0 to 8 scale&lt;br&gt;Administration time: 5 minutes</td>
<td>De-emphasis on rare phobic items, emphasis on agoraphobia and social fears&lt;br&gt;Weak support for utility of the FQ in older adults; consider using revised version&lt;br&gt;Stanley, Novy, Bourland, Beck, &amp; Averill, 2001&lt;br&gt;No fee, but permission is required</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Generalized Anxiety Disorder Inventory</strong>&lt;br&gt;Argyropoulos (2007)</td>
<td>Generalized anxiety</td>
<td>18 items rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely)&lt;br&gt;Administration time: 5 minutes</td>
<td>α ≥ .84, good convergent and divergent validity&lt;br&gt;Limited predictive utility of the somatic and sleep subscales&lt;br&gt;α ranges .82 to .92, .84, good convergent and discriminant validity&lt;br&gt;Specificity: 82%&lt;br&gt;α ranges from .68 to .93&lt;br&gt;Good psychometric support, brief, public domain&lt;br&gt;Does not incorporate avoidance behavior or specify functional impairment</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Generalized Anxiety Disorder 7-item scale (GAD-7)</strong>&lt;br&gt;Spitzer, Kroenke, Williams, and Lowe (2006)</td>
<td>Generalized anxiety</td>
<td>7 items rated on a 4-point Likert-type scale ranging from 0 (not at all) to 3 (nearly every day)&lt;br&gt;Administration time: 2–3 minutes</td>
<td>α ≥ .84, good convergent and divergent validity&lt;br&gt;Limited predictive utility of the somatic and sleep subscales&lt;br&gt;α ranges .82 to .92, .84, good convergent and discriminant validity&lt;br&gt;Specificity: 82%&lt;br&gt;α ranges from .68 to .93&lt;br&gt;Good psychometric support, brief, public domain&lt;br&gt;Does not incorporate avoidance behavior or specify functional impairment</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hospital Anxiety and Depression Scale (HADS)-Anxiety Subscale</strong>&lt;br&gt;Bjelland, Dahl, Haug, and Neckelmann (2002); Zigmond and Snaith (1983)</td>
<td>Generalized symptoms of anxiety and fear</td>
<td>7 items on the anxiety scale (14 items total), scored on a 4-point scale ranging from 0 to 3&lt;br&gt;Administration time: 2–5 minutes</td>
<td>Created HADS to avoid reliance on somatic symptoms of anxiety to better detect anxiety in patients with physical health problems&lt;br&gt;Obtain from National Foundation for Educational Research&lt;br&gt;Current, this is the only GAD-specific tool validated to assess symptom improvement sooner than one week following treatment initiation&lt;br&gt;Does not assess avoidance behavior or functional impairment&lt;br&gt;α ranges .68 to .93&lt;br&gt;Good discriminant, concurrent, factorial validity, sensitivity and specificity in the .70–.90 range&lt;br&gt;Good retest reliability (r &gt; .80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hospital Anxiety and Depression Scale (HADS)-Depression Subscale</strong>&lt;br&gt;Bjelland, Dahl, Haug, and Neckelmann (2002); Zigmond and Snaith (1983)</td>
<td>Generalized symptoms of depression</td>
<td>7 items on the depression scale (14 items total), scored on a 4-point scale ranging from 0 to 3&lt;br&gt;Administration time: 2–5 minutes</td>
<td>Currently, this is the only GAD-specific tool validated to assess symptom improvement sooner than one week following treatment initiation&lt;br&gt;Does not assess avoidance behavior or functional impairment&lt;br&gt;α ranges .68 to .93&lt;br&gt;Good discriminant, concurrent, factorial validity, sensitivity and specificity in the .70–.90 range&lt;br&gt;Good retest reliability (r &gt; .80)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Measure</th>
<th>Primary Reference</th>
<th>What It Measures</th>
<th>Description</th>
<th>Psychometrics</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall Anxiety Severity and Impairment Scale (OASIS)</strong></td>
<td>Norman, Cissell, Means-Christensen, and Stein (2006)</td>
<td>Continuous measure of anxiety-related severity and impairment across anxiety disorders</td>
<td>5-item scale assesses symptoms over the past week, and each item is rated on a 5-point scale (0 = little or none: anxiety was absent or barely noticeable; 4 = extreme: anxiety was overwhelming)</td>
<td>$\alpha \geq .80$</td>
<td>Test-retest reliability .82 Good convergent, and discriminant validity Good sensitivity and specificity with cutoff score of $\geq 8$ Incorporates avoidance behavior and functional impairment Easy to score Public domain</td>
</tr>
<tr>
<td><strong>My Mood Monitor (M-3) Check-list</strong></td>
<td>Gaynes et al. (2010)</td>
<td>Generalized anxiety disorder, panic disorder, social anxiety disorder, obsessive-compulsive disorder, agoraphobia, and PTSD</td>
<td>27-item checklist ranging from “not at all” to “most of the time” Administration time: 5 minutes</td>
<td>Anxiety module has sensitivity of .88 and specificity of .76 Newer measure, less research Screens for multiple common psychiatric illnesses and differentiates PTSD and bipolar disorder Questionnaire can be completed on line or by phone using a private service called What’s My M3</td>
<td></td>
</tr>
<tr>
<td><strong>PROMIS-Anxiety</strong></td>
<td>Cella et al. (2010)</td>
<td>Self-reported fear, anxious misery, hyperarousal, and somatic symptoms related to arousal</td>
<td>Short forms available with four items, six, seven, and eight items; items are rated on a 5-point scale ranging from 1 (never) to 5 (always)</td>
<td>On 4-item short form: $\alpha = .87$ Sensitivity: 64.6% Specificity: 86.3% Availability of CAT format Lack of research that evaluates the use of the PROMIS Anxiety for assessing treatment outcomes Public domain</td>
<td></td>
</tr>
<tr>
<td><strong>Pain Anxiety Symptoms Scale (PASS-20)</strong></td>
<td>McCracken and Dhingra (2002)</td>
<td>Fear and anxiety responses specific to pain</td>
<td>20 items rated on a scale from 0 (never) to 5 always Administration time: 6 minutes</td>
<td>Mean alpha = .81, good convergent and divergent validity (intercorrelations $r = .95$); good predictive validity; adequate construct validity Appears equivalent to the longer 40-item version in most important respects Public domain</td>
<td></td>
</tr>
</tbody>
</table>

Table 12.2 (Continued)
<table>
<thead>
<tr>
<th>Measure</th>
<th>Reference</th>
<th>Psychometric Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Anxiety Severity and Impairment Scale (OASIS)</td>
<td>Norman, Cissell, Means-Christensen, and Stein (2006)</td>
<td>Continuous measure of anxiety-related severity and impairment across anxiety disorders. 5-item scale assesses symptoms over the past week, and each item is rated on a 5-point scale (0 = little or none: anxiety was absent or barely noticeable; 4 = extreme: anxiety was overwhelming). Administration time: 2–3 minutes. α ≥ .80. Test-retest reliability .82. Good convergent, and discriminant validity. Good sensitivity and specificity with cutoff score of ≥ 8. Incorporates avoidance behavior and functional impairment. Easy to score. Public domain.</td>
</tr>
<tr>
<td>My Mood Monitor (M-3) Check-list</td>
<td>Gaynes et al. (2010)</td>
<td>Generalized anxiety disorder, panic disorder, social anxiety disorder, obsessive-compulsive disorder, agoraphobia, and PTSD. 27-item checklist ranging from “not at all” to “most of the time”. Administration time: 5 minutes. Anxiety module has sensitivity of .88 and specificity of .76. Newer measure, less research. Screens for multiple common psychiatric illnesses and differentiates PTSD and bipolar disorder. Questionnaire can be completed online or by phone using a private service called What’s My M3.</td>
</tr>
<tr>
<td>PROMIS-Anxiety</td>
<td>Cella et al. (2010)</td>
<td>Self-reported fear, anxious misery, hyperarousal, and somatic symptoms related to arousal. Short forms available with four items, six, seven, and eight items; items are rated on a 5-point scale ranging from 1 (never) to 5 (always). On 4-item short form: α = .87. Sensitivity: 64.6%. Specificity: 86.3%. Availability of CAT format. Lack of research that evaluates the use of the PROMIS Anxiety for assessing treatment outcomes. Public domain.</td>
</tr>
<tr>
<td>Pain Anxiety Symptoms Scale (PASS-20)</td>
<td>McCracken and Dhingra (2002)</td>
<td>Fear and anxiety responses specific to pain. 20 items rated on a scale from 0 (never) to 5 always). Administration time: 6 minutes. Mean α = .81, good convergent and divergent validity (r = .95); good predictive validity; adequate construct validity. Appears equivalent to the longer 40-item version in most important respects. Public domain.</td>
</tr>
<tr>
<td>Panic Disorder Self-Report Questionnaire (PDSR)</td>
<td>Newman, Holmes, Zuellig, Kachin, and Behar (2006)</td>
<td>Panic disorder. 7 items rated on a 5-point Likert scale ranging from 0 (none) to 4 (extreme). Administration time: under 5 minutes. 100% specificity and 89% sensitivity. Good test-retest reliability, convergent and discriminant validity, and kappa agreement of .93 with a structured interview. PDSR should be used only as a measure to identify the likely presence or absence of panic disorder, not designed to ascertain dimensional severity. Public domain.</td>
</tr>
<tr>
<td>Penn State Worry Questionnaire (PSWQ)</td>
<td>Meyer, Miller, Metzger, and Borkovec (1990)</td>
<td>Measure the trait of worry. 16 items scored on a 5-point Likert scale ranging from 1 (not at all typical) to 5 (very typical). Administration time: 5 minutes. α ranges .83 to .93; Good test-retest reliability (r’s = .74–.93) across a 1-month period; good discriminant and convergent validity. Specificity: 71%. Sensitivity: 71%. Widely used self-report tool for pathological worry and GAD. Reverse keyed items make the scale cumbersome to score. Some research suggests that a two-factor structure may provide a better fit than the intended one-factor structure. Public domain.</td>
</tr>
<tr>
<td>Penn State Worry Questionnaire 8-item version (PSWQ-A)</td>
<td>Crittendon and Hopko (2006); Hopko et al. (2003)</td>
<td>Measure the trait of worry. 8 items scored on a 5-point Likert scale ranging from 1 (not at all typical) to 5 (very typical). Administration time: 2–3 minutes. Internal consistency ranging from .87 to .94 and test-retest reliability ranging (r = .63–.95) Moderate to strong convergent validity with measures of worry and anxiety (r = .46–.83); one-factor solution in younger and older adults. Shows promise for use as brief screen. Easier to score than full version. Developed in older adult sample to improve psychometric properties and factor structure of the PSWQ in older adults.</td>
</tr>
<tr>
<td>Penn State Worry Questionnaire 3-item version (PSWQ-3)</td>
<td>Berle et al. (2011)</td>
<td>Core features of worry in GAD. 3 items scored on a 5-point Likert-type scale ranging from 1 (not at all typical) to 5 (very typical). Administration time: under 2 minutes. α &gt; .80. Specificity: 73%. Sensitivity: 71%. Only moderate convergent and discriminant validity. Developed to avoid problems related to reverse scored items and increase clinical utility. Easier to score than full version. Quick to administer.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Measure</th>
<th>Primary Reference</th>
<th>What It Measures</th>
<th>Description</th>
<th>Psychometrics</th>
<th>Considerations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Social Interaction Anxiety Scale (SIAS)</strong></td>
<td>Mattick and Clarke (1998)</td>
<td>Measures distress when meeting and talking with others</td>
<td>20 items on a 5-point scale ranging from 0 (not at all characteristic) to 4 (extremely characteristic) Administration time: 5–10 minutes</td>
<td>α ranging from .88 to .94; good test-retest reliability (.92) Good discriminate and construct validity</td>
<td>Research lacking among black population Difficult to discriminate between the fears of someone who experiences social anxiety and the more general worries of a patient with GAD, as patients who suffer from both disorders score higher on the SIAS than those with just social phobia</td>
</tr>
<tr>
<td><strong>Social Phobia Diagnostic Questionnaire (SPDQ)</strong></td>
<td>Newman et al. (2003)</td>
<td>Social anxiety</td>
<td>27 items consisting of questions assessing presence or absence of social fears, symptom severity ratings, and a list of social situations for which patients rate their degree of fear Administration time: 10 minutes</td>
<td>Specificity: 85% Sensitivity: 82% Good internal consistency reliability (α = .95), good split half (r = 0.90) reliability; strong 2-week test-retest reliability and good convergent and discriminant validity</td>
<td>Needs more research to confirm psychometric properties in a clinical population Public domain</td>
</tr>
<tr>
<td><strong>State Trait Anxiety Inventory (STAI)</strong></td>
<td>Spielberger, Gorsuch, Lushene, Vagg, and Jacobs (1983)</td>
<td>State and trait anxiety</td>
<td>20 items for state anxiety, 20 items for trait anxiety. All items are rated on a 4-point scale (e.g., from “Almost Never” to “Almost Always”) Administration time: 10 minutes</td>
<td>α ranges from .86 to .95 Test-retest reliability coefficients have ranged from .65 to .75 over a 2-month interval Considerable evidence attests to the construct and concurrent validity of the scale</td>
<td>Some research suggests a correlation with depression Sensitive to change during treatment Reverse scoring required Cost to purchase Attempts to develop a 6-item version of this scale (Marteau &amp; Bekker, 1992)</td>
</tr>
</tbody>
</table>
disorders: panic disorder, agoraphobia, GAD, social anxiety disorder, and substance-induced anxiety disorder. The reason for this line of inquiry is that these disorders and their variations have been associated with high rates of medical utilization (e.g., Katon et al., 1990). Again, knowledge of the DSM-5 anxiety disorders criteria is critical to determine the appropriate treatment course.

Another method of obtaining additional clinical information on the nature and intensity of problem anxiety involves the use of other anxiety measures, some of which are detailed later in this chapter (i.e., OASIS, PROMIS Anxiety, and BAI-PC). These and other anxiety measures have also been summarized in Table 12.2 and encourage clinicians to consider use of these measures and their associated psychometric properties when assessing anxiety in primary care. It should be noted that when supplementing the GAD-7 in primary care practice with other anxiety measures, the measures should be properly implemented under the supervision or guidance of those licensed to administer them and interpret the results, usually a licensed psychologist. The psychological consultant can also help make decisions of parsimony and identify only the most essential and pragmatic measures. The goal in primary care is to use brief and effective detection measures and intervention methods, not to provide comprehensive assessment and treatment—as that is the purview of specialty mental health.

**Determination of Need for Behavioral Health Intervention**
Using the guidelines just detailed in the screening process, the GAD-7 score can help determine the need for behavioral intervention. To activate behavioral intervention, two criteria should be considered: (1) a GAD-7 score above the threshold criteria for the specified population, and/or (2) patient endorsement of other anxiety complaints as outlined in Table 12.1. In general, if a patient meets one or both of these criteria, a behavioral health consultation or intervention for anxiety is warranted. Many individuals may not meet either criterion, but this does not rule out behavioral methods as a means of addressing the presenting problem.

**Determination of the Most Appropriate Intervention**
Determining the most appropriate treatment intervention largely depends on the anxiety problem. Suffice it to say that an elevated score on the GAD-7 alone cannot determine what the most appropriate clinical intervention should be, nor can a GAD-7 score alone predict treatment outcome. For example, individuals with social anxiety disorder tend to score lower on the GAD-7, presumably because the GAD-7 is more sensitive to the physical symptoms of anxiety. As a result, socially anxious individuals may be overlooked in the fast-paced primary care setting if the GAD-7 is the only screen used for problem anxiety. This may present a problem, as social anxiety disorder can be debilitating if left untreated. The recommendation, therefore, is to verbally inquire about other anxiety problems (as detailed in Table 12.1) in addition to the GAD-7 results. If anxiety appears to be an acute reaction to transitory stressors, then a full behavioral health consultation may not be necessary at first visit. The primary care provider can consult behavioral health at follow-up should the anxiety problems persist. If, however, a patient scores extremely high on the GAD-7 and reports a greater than 5-year history of problematic anxiety with associated poor functioning, for example, then a behavioral health consultation is likely warranted, and may even result in a referral to specialty mental health thereafter. Whether patients are seen for follow-up by the behavioral health consultant for continued psychotherapy in primary care (or if the consultant is a psychiatrist, for medication management) largely depends on the primary care mental health integration model in place.
Provision of Feedback to the Patient

Because the GAD-7 can be quickly scored, results can be discussed with patients in real time. It is recommended that the scoring ranges (i.e., minimal, mild, moderate, severe) be discussed with patients as well. An example of feedback on an initial visit might be as follows:

Thanks for completing the questionnaire. It appears that you scored a 16 on the 7-item General Anxiety Disorder scale, our standard measure of anxiety symptoms we frequently administer to patients seen in primary care. This score suggests that you are experiencing anxiety symptoms in the severe range, and that these symptoms are impacting your functioning and quality of life. This indicates that we should look more closely at your anxiety so that we may help you manage it. How does this assessment compare to your experience?

It is also helpful to point out to patients that although the GAD-7 is a valid and highly reliable measure of clinical anxiety, it will not capture entirely the patient’s experience. However, in an imperfect world, it is one of the best and most practical methods of establishing an objective index of anxiety symptomatology. Furthermore, patients should be told that the GAD-7 score at the first appointment can act as a measure to which subsequent administrations of the GAD-7 can be compared.

Use as an Instrument for Treatment Monitoring and Outcomes Assessment

Given the brevity and ease in scoring, the GAD-7 is ideal for anxiety treatment monitoring in primary care. Indeed, studies suggest that the GAD-7 is responsive to treatment changes in individuals with anxiety disorders (e.g., Dear et al., 2011, Mewton, Wong, & Andrews, 2012; Robinson et al., 2010; Ruiz et al., 2011). Successive GAD-7 scores, after an initial baseline score is obtained in the screening phase, can help determine treatment effectiveness. Results of subsequent GAD-7 administrations should be communicated back to patients in real time in order for patients to view progress, which can be a therapeutic process in and of itself.

There are several considerations, however, that cannot be overlooked when implementing the GAD-7 as an outcome measure in primary care. First, as previously mentioned in this chapter, the GAD-7 may be less sensitive to social anxiety and avoidance symptoms. As such, benchmarking improvement in such symptoms should be done by methods other than the GAD-7. Second, while the GAD-7 does have an item that assesses global functional impact, it does not specify which domain of function. Thus, the GAD-7 cannot provide outcome data about the impact of anxiety symptoms on role function in social versus occupational versus familial domains and so forth. Third, when implementing any psychometric instrument in the primary care setting, it must be made clear from the outset why using outcome measures benefits the practice and patient care. Too often, psychometric measures are applied with no forethought into why measuring a particular psychological phenomenon (whether it be behavioral, cognitive, neurocognitive, etc.) is important to the organization or to those it serves. The result of poor planning is a poor organizational attitude toward the measure and an unwillingness among staff and patients to value data that, in turn, can yield poor and unreliable outcome data. Overall, using the GAD-7 in practice should be well planned and its goals well articulated before it is implemented. When this is done, the implementation will become routine and smooth, and will likely result in increased patient and provider satisfaction.
Relevant Support
As previously mentioned, the GAD-7 has been demonstrated to be sensitive to change as affected by treatment, making it a good choice as an anxiety treatment outcome measure. Indeed, Dear et al. (2011) used the GAD-7 as an outcome measure for patients getting treatment for GAD, and found that the GAD-7 was more sensitive to change compared to the Penn State Worry Questionnaire (PSWQ). Ruiz et al. (2011) examined the validity of the GAD-7 as an outcome measure of disability in patients with GAD in primary care, and found that the GAD-7 highly correlates with disability outcome measures. Several trials of Internet CBT for GAD have also used the GAD-7 as an outcome measure and found high sensitivity to change over time (Mewton et al., 2012; Robinson et al., 2010). Further, researchers have used the GAD-7 as an outcome measure not only in mixed anxiety samples but depression samples as well (e.g., Hammond et al., 2012; Richards & Borglin, 2011; Richards & Suckling, 2009). The GAD-7 is also sensitive to change in heterogeneous psychiatric samples (Beard & Bjorgvinsson, 2014). Taken together, more research is needed to determine the responsiveness of the GAD-7 in treatment trials and other types of longitudinal studies where anxiety is the primary outcome.

Implementation Issues
One frequent barrier to implementation of a psychological measure in primary care is the cost. Fortunately, the GAD-7 is public domain and available for free with no permission required to reproduce, translate, display, or distribute.

Guidelines for Decision-Making
When deciding to continue, discontinue, or change treatment based on repeated GAD-7 scores, three key points may prove useful: (1) specifying the treatment goals, (2) identifying a specific time frame for the treatment, and (3) using the Reliable Change Index (RCI; Jacobson & Truax, 1991). With regards to specifying a treatment goal, one must consider not only a symptom goal as measured by the GAD-7 but also a functional goal. A functional goal is one that involves the reduction or elimination of an observable and measurable behavior that impairs occupational, social, or other function. Identifying treatment goals prior to treatment should be a collaborative effort between the patient and provider. Such collaboration will enhance treatment adherence and help clarify when satisfactory resolution of the anxiety problem is achieved. A symptom goal can be reaching a minimal GAD-7 score (i.e., < 5). An example of a functional goal can be the reduction of emergency room visits for panic. If both symptom and functional goals are achieved, then treatment can be discontinued. If one goal but not the other is achieved, treatment may be changed to focus more solely on the unachieved goal or discontinued if the patient finds results unsatisfactory.

With respect to identifying a specific time frame for treatment, the patient and provider should consider a long enough interval to see satisfactory results. Most anxiety disorders do not onset rapidly (with the exception of acute stress disorder, or adjustment disorder with anxiety), and thus treatment will not necessarily reverse the effects of anxiety rapidly. As the old saying goes, “it’s a mile in, and a mile out.” A realistic, achievable time frame should therefore be emphasized. Because the GAD-7 is an assessment of symptoms over the previous two weeks, administration of the GAD-7 should be done at a recommended frequency no greater than once every two weeks.

Lastly, the RCI is a statistic that was devised as a method to determine if differences in pre- and posttreatment test scores are due to a true change in the psychological construct being measured (e.g., anxiety), or if the change is simply due to chance (Jacobson & Truax, 1991). In other words, it determines if the change in pre- to posttreatment test scores is statistically significant. This has
important treatment implications for treatment outcome, in that it provides a guideline for determining whether to stop or continue treatment. If the patient’s pre–post treatment difference in GAD-7 scores exceeds the magnitude of the RCI, then the change in GAD-7 scores is significant. Based on previous research using the original validation study statistics (e.g., Gyani, Shafran, Layard, & Clark, 2013), the RCI value computed is 3.5. That is, a change of 3.5 or more GAD-7 points is needed in order for the change to be considered significant. Overall, using the goal of a GAD-7 score within the minimal range in conjunction with the RCI is advised when making treatment decisions.

Cross-Cultural Applicability
As previously discussed, the GAD-7 has been studied in large, nationally representative samples (e.g., Lowe et al., 2008) and has numerous language adaptations (Chinese; Zeng et al., 2013; Dutch; Donker et al., 2011; Finnish; Kujanpaa et al., 2014; Portuguese; Sousa et al., 2015; Turkish; Konikan et al., 2013; Spanish; Garcia-Campayo et al., 2010; British Sign Language; Rogers et al., 2012). It also has been used in many different cultures (Donker et al., 2011; Garcia-Campayo et al., 2010; Konikan et al., 2013). Until recently, however, little was known about the psychometric properties of the GAD-7 across cultural groups. Specifically, because the GAD-7 was developed using a sample comprised primarily of Caucasian patients (80%; Spitzer et al., 2006), an examination of whether items functioned differently across cultural groups was precluded. As such, Parkerson, Thibodeau, Brandt, Zvolensky, and Asmundson (2015) set out to assess the factor structure of the GAD-7 in Caucasian (from Regina, n = 610; from Houston, n = 170), Hispanic (n = 103), and black (n = 174) young adults and tested for cultural-based biases in response pattern in the form of differential item functioning (DIF). Results indicated that the original unitary model as proposed by Spitzer and colleagues (2006) provided a relatively poor fit to the data. Instead, a revised unitary model was tested (Kertz et al., 2013), allowing for covariance of error terms among items 4 (Trouble relaxing), 5 (Being so restless that it is hard to sit still), and 6 (Becoming easily annoyed or irritable). Similar to previous results (Kertz et al., 2013), all fit indices improved for all groups when testing the refitted model. In other words, Parkerson and colleagues’ (2015) results aligned with findings of Kertz et al. (2013), who originally suggested that shared variance across items 4–6 may reflect a somatic tension/autonomic arousal subtype of GAD. In addition, Parkerson and colleagues (2015) found that participants who self-identified as black reported high levels of the latent GAD trait and typically endorsed lower options on items 1 (Feeling nervous, anxious, or on edge), 5 (Being so restless that it is hard to sit still), and 6 (Becoming easily annoyed or irritable) compared to other participants with similar latent levels of GAD. From this, the researchers proposed that the lower prevalence rate for the black sample in their study could potentially reflect cross-cultural measurement biases related to the diagnostic instrument rather than true differences in GAD symptoms. Taken together, the results of Parkerson et al.’s (2015) study brings to light the possibility that the GAD-7 scores reported by black individuals may reflect a lower severity of GAD symptoms than are experienced. As such, providers are encouraged to be aware of this potential bias when screening for anxiety in primary care.

OTHER ANXIETY MEASURES FOR PRIMARY CARE
Although the GAD-7 is established and commonly used for assessing anxiety in primary care settings, there are several other measures that are well suited for this purpose. Next we discuss the use and psychometric properties of the Overall Anxiety Severity and Impairment Scale (OASIS), Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety, and the Beck Anxiety Inventory-Primary Care (BAI-PC).
Overall Anxiety and Severity and Impairment Scale (OASIS)

The Overall Anxiety and Severity and Impairment Scale (OASIS) was developed in an effort to assess for anxiety as a construct that exists across anxiety disorder diagnoses, as well as subsyndromal levels of anxiety (Norman, Cissell, Means-Christensen, & Stein, 2006). Given that the majority of anxiety measures are designed to assess disorder-specific symptoms (e.g., social anxiety, panic disorder), their ability to serve as a population-based screener for the presence of anxiety symptoms across domains may be limited (Campbell-Sills et al., 2009). In busy primary care settings, it may be impractical to conduct routine assessments with multiple anxiety disorder-specific measures given time limitations and burden to patients and providers (Campbell-Sills et al., 2009; Norman et al., 2011). Accordingly, the OASIS, a brief transdiagnostic anxiety measure, is ideal for this purpose. Specifically, the OASIS assesses anxiety on a continuum of severity of anxiety and related impairment (Norman et al., 2006). Although it was developed based on DSM-IV-TR guidelines for assessing the severity and related impairment of anxiety symptoms, it remains consistent with DSM-5 criteria and includes items reporting the frequency of anxiety symptoms, perceived intensity of the symptoms, and life-related interference due to anxiety (Norman et al., 2006). Furthermore, unlike most brief measures of general anxiety symptoms, the OASIS captures one of the most prominent features of anxiety disorders: avoidance. This is very important given that phobic avoidance and associated impairment are highly clinically relevant (Campbell-Sills et al., 2009)—perhaps arguably even more so than frequency or intensity of symptoms.

Administration and Scoring

This 5-item scale assesses symptoms over the past week, with each item rated on a 5-point Likert-type scale from 0 to 4 (e.g., “Little or None: Anxiety was absent or barely noticeable,” to “Extreme: Anxiety was overwhelming. It was impossible to relax at all. Physical symptoms were unbearable.”). When completing this measure, individuals are asked to consider a variety of anxiety-related experiences and to answer questions in a manner that reflects their experience with all of these symptoms (e.g., panic attacks and flashbacks). Adding the item scores, with higher scores representing greater anxiety severity and impairment, generates a total score ranging from 0 to 20. The OASIS is estimated to take 2–3 minutes to complete (Norman et al., 2011) and is available for public use.

Scale Development and Psychometric Properties

The OASIS was initially developed using a sample of 711 undergraduate college students enrolled in an introductory psychology course (mean age = 18, 72.6% female, 54.3% Caucasian; Norman et al., 2006). Participants completed other well-established measures of depression and anxiety to assess validity, and a subset completed surveys 1 month later to assess reliability of the scale. Results from this initial sample yielded adequate internal consistency (Cronbach’s $\alpha = .80$). Inquiry into the factor structure of the measure yielded a single factor when assessed in each half of the sample separately, accounting for 55%–56% of the variance (Norman et al., 2006). Next, the OASIS demonstrated high convergence with other measures of anxiety, including the Brief Stress Inventory-18 (BSI-18), Fear Questionnaire (FQ), Spielberger Trait Anxiety Questionnaire (STAIT), the Neuroticism subscale of the Neuroticism Extraversion Openness Five-Factor Inventory (NEO-FFI), the Anxiety Sensitivity Inventory (ASI), and the Mini Social Phobia Index (Mini-SPIN). There was also a high convergence with depressive symptoms on the Beck Depression Inventory (BDI). Evidence for discriminant validity of the OASIS was supported by a high negative correlation with the Connor-Davidson Resilience Scale (CD-RISC), and low correlations with the Barratt Impulsivity Scale (BIS-II) and the NEO openness and agreeableness subscales (Norman et al., 2006). Furthermore, 1-month test-retest reliability ($n = 75$) was high (kappa = .82).
In a follow-up study (Norman et al., 2011) using a separate college student sample \((N = 171; \text{mean age} = 19.0, 78.4\% \text{ female}, 50.3\% \text{ Caucasian})\), participants were selected based on screening in the “high anxiety” (upper 15th percentile) and “normal anxiety” (40th–60th percentile) ranges on the BSI-Anxiety subscale. Consistent with the prior evaluation, results of a factor analysis confirmed a single-factor structure (Norman et al., 2011). Results also indicated moderate significant correlations with measures of anxiety (BSI-18, ASI, STAIT, Social Interaction Anxiety Scale [SIAS]) and neuroticism (NEO-Neuroticism). As before, there was a high negative correlation with resiliency (CD-RISC), and low or no significant correlations between the OASIS and impulsivity (BIS), sensation-seeking (Zuckerman Sensation-Seeking Scale; SSS-V), agreeableness (NEO-Agreeableness), or physical health-related quality of life (SF-12).

Although the initial psychometric investigation of the OASIS was promising, a major limitation was the use of a nonclinical sample. In order to study the psychometric properties in a clinical sample, Campbell-Sills and colleagues (2009) conducted an evaluation with a large sample \((N = 1,036)\) of primary care patients who had been referred for anxiety treatment \((\text{mean age} = 42.8, 71.2\% \text{ female}, 61.9\% \text{ Caucasian})\). A subsample \((n = 623)\) enrolled in a treatment study and completed a variety of measures used to establish convergent and discriminant validity. Participants engaged in a structured clinical interview (Mini-International Neuropsychiatric Interview), and results indicated a high proportion of patients meeting criteria for GAD (63.1%), panic disorder with or without agoraphobia (46.4%), social anxiety disorder (34.7%), PTSD (15.9%), agoraphobia without panic (9.7%), and obsessive-compulsive disorder (8.3%). A small proportion of the sample did not meet criteria for any anxiety disorder (10.7%). There were high rates of comorbidity with depressive disorders (60.6% depression, 4.3% dysthymia) and substance use disorders (alcohol use 10.6%, drug use 3.9%). Results of the psychometric evaluation indicated that OASIS scores were associated with both anxiety and depressive disorder diagnoses. Similar to the initial validation study, factor analyses yielded a one-factor solution, with items loading on a single latent factor. As before, internal consistency was good (Cronbach’s \(\alpha = .84\)). Finally, results supported strong convergence between the OASIS and other measures of anxiety, including general symptoms (BSI-18 anxiety subscale), social anxiety disorder (SPIN), PTSD (Posttraumatic Stress Disorder Checklist-Civilian Version; PCL-C), and GAD (Generalized Anxiety Disorder Severity Scale; GADSS). Likewise, there was a high association between the OASIS and depressive symptoms (PHQ-9). The OASIS also correlated highly with a measure of disability, and negatively correlated with measures of mental health and quality of life. Additionally, the OASIS did not correlate significantly with alcohol use, and there were low correlations with measures of social support and physical health (Campbell-Sills et al., 2009). Overall, these results further support the evidence of good psychometric properties for the OASIS, including in a primary care sample.

The OASIS has also been evaluated among specific anxiety populations. In sample of women with and without trauma histories \((N = 84)\), results further supported the construct validity of the OASIS (Norman et al., 2013). The OASIS demonstrated significant positive correlations with measures of anxiety (BSI-18-Anxiety, ASI, STAIT), neuroticism (NEO-Neuroticism), PTSD (CAPS), and depression (BSI-18-Depression). Consistent with previous research, the OASIS did not correlate significantly with measures assessing dissimilar constructs, including agreeableness (NEO-Agreeableness) and physical health-related disability (SF-12).

More recently, the psychometric properties of the OASIS were explored among a naturalistic sample of individuals seeking specialty outpatient treatment for anxiety disorders (Moore et al., 2015). Participants included 347 patients \((\text{mean age} = 34.4, 61\% \text{ female}, 86\% \text{ Caucasian}, \text{ the majority of whom met DSM-IV-TR criteria for an anxiety disorder} \ ([n = 281])\). Among those with an anxiety disorder, there were high rates of comorbidity with at least one other anxiety disorder (28%) or depressive disorder (41%). Results indicated significantly higher mean OASIS scores for participants with an
anxiety disorder (9.82 ± 4.28) compared to those without an anxiety disorder (7.33 ± 4.25), further supporting the validity of the instrument. As in prior studies, high convergent validity was found between the OASIS and other related measures, including the Outcomes Questionnaire (OQ) distress measure, PSWQ, Yale-Brown Obsessive Compulsive Scale (YBOCS), and the PHQ-9 (Moore et al., 2015).

**Use as a Screener**
Overall, the literature suggests good specificity and sensitivity with a cutoff score of 8 or greater indicating clinically significant anxiety. In a primary care sample, cutoff scores between 7 and 9 maximized sensitivity and specificity, with a cutoff score of 8 or greater having the most favorable balance between sensitivity (.89) and specificity (.71). Further, this score correctly classified 87% of the study sample. In a nonclinical sample of undergraduate students (N = 48; Norman et al., 2011), a score of 8 or greater correctly identified 78% of those meeting criteria with an anxiety disorder on the SCID. This cutoff score was associated with 69% sensitivity and 74% specificity. Furthermore, this cutoff score was identified as optimal for classifying a sample of patients presenting for outpatient anxiety treatment (N = 357; Moore et al., 2015). This score resulted in correct classification of 67% of patients, with a favorable balance of sensitivity (.69) and specificity (.55).

Other samples have yielded evidence for different cutoff scores. In a sample of individuals with and without trauma histories (N = 84), a cutoff score of 5 or greater was reported to have optimal sensitivity (.93) and specificity (.90), and resulted in correct identification of 91% of the sample with clinically significant PTSD symptoms (Norman et al., 2013). In a Japanese sample of individuals participating in an online survey marketing program, a cutoff score of 9 was found to be optimal (Ito et al., 2014).

**Use for Assessing Outcomes**
There have been promising initial results for the OASIS as a potentially useful measure for assessing treatment outcomes, although research in this area is limited. An RCI of 4 is recommended as indicating clinically significant change (Moore et al., 2015). Among patients presenting for anxiety treatment, an RCI of 4.41 was found for the OASIS. Accordingly, researchers recommended a decrease of 4 be considered as reliable improvement, whereas an increase of 4 or more points indicates reliable deterioration (Moore et al., 2015). In a Dutch translated version of the OASIS, a pre- to posttreatment change score of 4 was found to present clinically significant improvement (Hermans, Korrelboom, & Visser, 2014).

The OASIS has also been demonstrated to converge well with other measures assessing treatment outcomes. In a recent study, change in the OASIS was highly related to corresponding changes on other anxiety and measures (i.e., YBOCS, PSWQ, OQ), with a moderate correlation with changes in depressed mood on the PHQ-9 6 months after the intake assessment (Moore et al., 2015). In a study of individuals with PTSD (Norman et al., 2013), OASIS change scores were highly correlated with clinical improvement as assessed by the Clinician-Administered PTSD Scale (CAPS) and self-reported anxiety (BSI-Anxiety).

**Cross-Cultural Applicability**
Given that the OASIS is a relatively new instrument, there have been few studies as of yet directly evaluating the psychometric properties of the OASIS across a diverse range of participants. In the United States, few associations have been found between OASIS scores and demographic characteristics. Specifically, no relationship between OASIS scores and gender, age, or income level/socioeconomic status was found in studies of college students (Norman et al., 2006; Norman et al., 2011) or
patients presenting for outpatient anxiety disorder treatment (Moore et al., 2015). In one college sample, Latino/a participants had lower OASIS scores (6.39 ± 3.39) on average compared to Asian individuals (8.38 ± 2.69; Norman et al., 2006). This finding has not been replicated in other studies (Moore et al., 2015; Norman et al., 2011). These studies, however, have been limited by small sample sizes for non-Caucasian participants. Additional research exploring potential between-group differences for OASIS scores is warranted.

The OASIS has also been evaluated in Dutch- and Japanese-translated forms. In the Netherlands, the OASIS was administered to patients presenting for outpatient mental health treatment (anxiety disorders, \( n = 155 \); other disorders, \( n = 102 \)). Researchers found similar psychometric properties for the OASIS as in prior samples, including a single-factor structure, high internal consistency (Cronbach’s \( \alpha = .91 \)), and strong correlations with measures of anxiety, distress, and neuroticism (Hermans et al., 2014). However, unlike prior studies in clinical samples in the United States, a lower cutoff score of 5 was found to yield a favorable balance between sensitivity (.86) and specificity (.77) and to correctly identify 82.5% of the sample (Hermans et al., 2014). There were no associations between OASIS scores and sex, age, or educational level in this study (Hermans et al., 2014).

The Japanese-translated form (Ito et al., 2014) was examined among individuals who registered with an Internet marketing company and agreed to take part in a survey study (\( N = 2,830 \)). Results supported a one-factor solution for the structure of the OASIS and high internal consistency and test-retest reliability for this version of the measure (Ito et al., 2014). It is expected that additional translated forms of the OASIS will be generated and evaluated in the future, allowing cross-cultural comparisons.

**Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety**

The Patient-Reported Outcomes Measurement Information System (PROMIS) was developed in an effort to assess perceived treatment-related changes in symptoms, distress, or functioning (Cella et al., 2010), consistent with the World Health Organization’s (2007) biopsychosocial framework of health (Pilkonis et al., 2011). Accordingly, a cooperative group of scientists funded by the National Institutes of Health (NIH) was tasked with creating an assessment system for efficient and valid measurement of symptoms, quality of life, and functioning across domains and patient populations (Cella et al., 2010). This is referred to as the PROMIS Cooperative Group. The PROMIS framework includes multiple domains assessing physical functioning, pain, emotional distress, social functioning, fatigue, and general health. The anxiety subscale (PROMIS Anxiety), which is the subscale of relevance for this chapter, includes a test bank of items that was generated out of this initial wave of test development (Cella et al., 2010). Pilkonis and colleagues (2011) describe the development of the anxiety test bank in detail. Each domain was subjected to extensive psychometric analyses using sophisticated statistical techniques, including both qualitative and quantitative methods (DeWalt, Rothrock, Yount, Stone, & PROMIS Cooperative Group. 2007). Specifically, items were pulled from a broad range of well-established measures, and subjected to item response theory (IRT) analysis in a computerized adaptive testing (CAT) format, with the goal of selecting items that best assess the relevant latent trait (i.e., anxiety; DeWalt et al., 2007). Several short forms of the test bank are available that are suitable for delivery in a primary care setting.

**Administration and Scoring**

Scales are available to administer via CAT and are also available in a traditional form with a fixed number of items that can be administered via paper-and-pencil format (Kroenke, Yu, Wu, Kean, & Monahan, 2014). Short forms of the measure are available with four, six, seven, and eight items,
respectively. In order to improve comparison across different measures and to utilize a standardized metric, PROMIS scores can be reported as both raw scores and T-scores, such that a score of 50 represents the mean with a standard deviation of 10 (Schalet, Cook, Choi, & Cella, 2013). Raw scores are generated by totaling the scores for each item. Promis Anxiety instruments, including short forms of the measure and a scoring and interpretation manual, are available via the National Institutes of Health website for public use (http://nihpromis.org).

Scale Development and Psychometric Properties
The PROMIS Anxiety was developed first by reviewing items of existing anxiety measures, which yielded 443 unique anxiety assessment items across 30 facets (i.e., specific symptoms or characteristics, such as fear (Pilkonis et al., 2011). Next, a qualitative review of items was conducted, including focus groups and cognitive interviews (as described in detail by DeWalt et al., 2007). For the anxiety item bank, three focus groups were conducted with medical and psychiatric outpatients (Pilkonis et al., 2011). A resulting bank of items was reviewed and a 7-day time frame and 5-point Likert-type scale were applied (“I felt ________,” never, rarely, sometimes, often, or always). Items were developed at an average of a first-grade reading level (SD = 1.5 grades) to increase the accessibility of the measure to patients at various levels of literacy (Pilkonis et al., 2011).

Next, a subset of patients engaged in interviews to discuss their understanding and interpretation of test items, which resulted in the revising of several items to improve clarity. After an item bank was finalized, a large sample of participants was administered either a full anxiety item bank (N = 1,974) or blocks of seven items across domains (N = 14,128). This sample included a large proportion of patients reporting at least one chronic health condition. A subsample that was demographically representative of the US population was used to generate population-normed T-scores for the scale. Factor analysis confirmed a one-factor solution for the anxiety scale. IRT calibration revealed that two items on the anxiety scale showed differential item functioning related to age. Specifically, the item “I felt like I was going crazy” was endorsed more frequently by younger participants, whereas the item “I felt shaky” was endorsed more frequently by older participants (Pilkonis et al., 2011).

A final review yielded 29 items in the anxiety test bank, of which 11 assess affective symptoms, 8 cognitive, 8 somatic, 1 behavioral, and 1 reflecting need for treatment (Pilkonis et al., 2011). A shorter form of seven items was created based on the results of a CAT simulation. Evaluation of the brief form was found to have good correlation with the full form (.83) and excellent internal consistency (Cronbach’s α = .93). Results indicated high correlations between the anxiety scale and the Mood and Anxiety Symptom Questionnaire (MASQ; r = .80) and CES-D (Center for Epidemiological Studies Depression Scale; r = .75). Test information curves are available via Pilkonis and colleagues (2011).

The PROMIS measures were later updated to include a scale of somatic arousal, developed from the MASQ. This represents a separate scale from the anxiety test bank. Together, the anxiety and somatic arousal scales form the new Fear Test (Pilkonis et al., 2013).

The emotional distress scales, including anxiety, were also subjected to evaluation with a nominal response model, in order to evaluate response category discrimination across the distress item pools (i.e., depression, anxiety, anger; Preston, Reise, & Hays, 2011). In this study, 767 individuals responded to the full set of anxiety test items. Results indicated that each of the distress item pools, including anxiety, had high measurement precision and also contributed unique variance to the overall distress domain (Preston et al., 2011).

In order to assess how the PROMIS Anxiety functions in comparison to other well-established measures of anxiety, Schalet and colleagues (2013) transformed scores on the MASQ-GA and GAD-7 to the PROMIS metric. Accordingly, researchers and clinicians now have the availability of converting scores on other anxiety measures for comparison with PROMIS scores.
Regarding specific populations, the PROMIS Anxiety has been widely studied among patients with various chronic conditions, including cirrhosis (Bajaj et al., 2011), osteoarthritis and chronic pain (Driban, Morgan, Price, Cook, & Wang, 2015; Fenton, Palmieri, Diantonio, & Vongruenigen, 2011; Kroenke et al., 2014; Stone, Broderick, Junghaenel, Schneider, & Schwartz, 2015), HIV (Edwards et al., 2015; Fredericksen et al., 2015), cardiac problems (Fischer et al., 2015), cancer (Flynn et al., 2010; Romero, Flood, Gasiewicz, Rovin, & Conklin, 2015; Wagner et al., 2015; Yost, Eton, Garcia, & Cella, 2011), chronic obstructive pulmonary disease (Irwin et al., 2014), inflammatory bowel disease and irritable bowel syndrome (Kappelman et al., 2013; Taft, Riehl, Dowjotas, & Keefer, 2014), scleroderma (Khanna et al., 2012), spinal cord injury (Kisala et al., 2015), multiple sclerosis (Senders, Hanes, Bourdette, Whitham, & Shinto, 2014), and asthma (Stucky, Sherbourne, Edelen, & Eberhart, 2015).

Use as a Screener
The four-item short form of the anxiety test bank represents the briefest version of the measure, although the six-, seven-, and eight-item versions would also be appropriate for screening purposes in a primary care setting. The four-item short form has been studied in comparison with another brief screening measure, the distress and risk assessment method (DRAM), which consists of the Modified Somatic Perception Questionnaire and the modified Zung Depression Index (mZDI; Hung et al., 2015). This study included 316 patients (55% female) at a large university health center, who were predominately Caucasian with a mean age of 57 years. The short form of the instrument had high internal consistency (.87) and was highly correlated with the DRAM mZDI. Given that the four-item short forms of the PROMIS Anxiety and PROMIS Depression accounted for a significant amount of variance in the DRAM mZDI, the authors concluded that it would be appropriate to use these as brief screeners to lessen patient burden for completing forms (Hung et al., 2015). For this study the authors found that a T-score of 58.1 was the best cutoff score for the anxiety scale (Hung et al., 2015).

The brief form of the PROMIS Anxiety screen (four-item version) was also administered in a study with primary care patients, including 244 predominately white male veterans, presenting with chronic persistent pain (Kroenke et al., 2014). For this population, a cutoff score of 8 or greater was recommended, yielding sensitivity of 64.6% and specificity of 86.3%. The lower than expected sensitivity of this measure was attributed potentially to unique features of the study sample; for example, several patients with clinical anxiety disorders endorsed the lowest possible scores on the PROMIS Anxiety (Kroenke et al., 2014). Schalet and colleagues (2013) found that one standard deviation above the mean (i.e., T-score of 60) for PROMIS Anxiety scores for the full 29-item measure corresponded to percentile ranks of 83 for women and 88 for men, which is consistent with 12-month prevalence rates of anxiety disorder diagnoses. This finding further supports the use of the PROMIS for detecting clinically significant anxiety symptoms; however, these findings have not yet been replicated with the four-item version.

Use for Assessing Outcomes
At present, there is a lack of research that evaluates the use of the PROMIS Anxiety for assessing treatment outcomes. This is likely due to the fact that this measure was developed relatively recently. However, a recent study with a chronic pain population found that a T-score decrease of 3.0–3.55 represents a minimally important difference (MID) on the 29-item PROMIS Anxiety Item Bank (Swanholm, McDonald, Makris, Noe, & Gatchel, 2014). The extent to which these findings would translate to a primary care sample or the more brief forms is unclear and additional research is needed to understand meaningful changes in PROMIS Anxiety scores in this population. Given that the PROMIS Anxiety was developed from existing, well-established measures of anxiety that have been
used to monitor treatment progress, it is likely that this measure may also represent a useful tool for assessing anxiety outcomes.

**Cross-Cultural Applicability**

An important strength of the PROMIS in comparison to the majority of existing symptom measures is that a large standardization sample was used ($N = 21,133; n = 14,836$ for the anxiety scale) and selected to be representative of the distribution of the US population based on the 2000 census (Pilko-nis et al., 2011). Further, translations of the PROMIS Anxiety are available or under development for the following languages: Spanish, Dutch, German, Portuguese, Chinese (Simplified and Traditional), Farsi, Turkish, and Hindi (http://nihpromis.org). Published research studies support the psychometric properties for the Portuguese (Castro, Rezende, Mendonça, Silva, & Pinto Rde, 2014), German (Wahl et al., 2015), and Dutch–Flemish (Terwee et al., 2014) versions of the PROMIS Anxiety.

**Beck Anxiety Inventory-Primary Care (BAI-PC)**

The Beck Anxiety Inventory (BAI; Beck et al., 1988) is a 21-item well-established measure of anxiety symptoms, which includes 14 items assessing physical anxiety symptoms (e.g., dizziness, sweating) and 7 items assessing a subjective domain of anxiety (e.g., fear). In consideration of the scope of the current chapter, the focus will be on the brief form of the instrument, which has been developed for primary care settings. An overview of the use of the full measure in primary care populations is provided by Ferguson (2000). The Beck Anxiety Inventory-Primary Care version (BAI-PC) comprises seven items, which were developed from the Subjective subscale of the original BAI. These items assess difficulty relaxing, fear of worst case scenarios, feeling terrified, nervous and scared, and fear of losing control or dying (Beck et al., 1997). These items represent several of the criteria for generalized anxiety disorder and panic disorder, and are intended to reflect a subjective sense of anxiety (Beck et al., 1997).

**Administration and Scoring**

Patients are asked to rate the extent to which they were bothered by symptoms during the past two weeks, including today, and items are rated on a 4-point scale from 0 to 3 (0 = not at all, 3 = severely). The measure is scored by totaling the items, yielding a total score ranging from 0 to 21 (Beck et al., 1997). It is estimated to take 2–5 minutes to complete (Kearney, Wray, Dollar, King, & Vair, 2014). The BAI-PC is a proprietary measure that is available for purchase via Pearson Education, Inc.

**Scale Development and Psychometric Properties**

In order to evaluate the BAI-PC in a primary care setting, Beck and colleagues (1997) administered the measure to 56 consecutive primary care patients at several locations. The sample was majority female (73%) and Caucasian (92%), with a mean age of 48.5 years. The majority of patients presented with at least one chronic health condition (61%). Patients underwent a diagnostic interview, and results indicated that 48% of the sample met DSM-III-R criteria for an anxiety disorder, the majority of which met criteria for generalized anxiety and/or panic disorder. Results of psychometric evaluation revealed item-total correlations ranging from .59 to .81 and internal consistency of .90, suggesting good reliability for clinical purposes (Beck et al., 1997). Further, scores on the BAI-PC positively correlated with the presence of an anxiety disorder ($r = .58, p < .001$). Although BAI-PC scores converged with depression scores on the BDI-PC (.86), results indicated that the BDI-PC more strongly correlated to a mood disorder diagnosis whereas the BAI-PC more strongly correlated to an anxiety disorder diagnosis (Beck et al., 1997).
In a follow-up study, Mori and colleagues (2003) administered the BAI-PC to 313 veterans presenting at a primary care clinic. Unlike the prior sample, participants in this study were majority male (98%) and Caucasian (77%), with an average age of 64 years. Results of this study supported the single-factor structure of the measure with good internal consistency ($\alpha = .90$). BAI-PC scores were highly correlated with the full version of the measure (BAI, $r = 74$).

**Use as a Screener**

ROC analyses were conducted to identify the optimal cutoff score for the BAI-PC, and results suggested that a total score of 5 yielded a favorable balance of sensitivity (.85) and specificity (.81) for identifying patients with GAD and panic disorder in this setting (Beck et al., 1997). Similarly, in a second study with a larger sample, the cutoff score of 5 resulted in 84.5% sensitivity and 79.5% specificity in identifying patients with clinical levels of anxiety (Mori et al., 2003). This cutoff score also performed well at identifying both depression (sensitivity = .91, specificity = .75) and PTSD (sensitivity = .97, specificity = .73). Consequently, the BAI-PC may be a useful screening measure for anxiety, PTSD, and depression in primary care settings by identifying patients for whom further consultation may be indicated (Mori et al., 2003).

**Use for Assessing Outcomes**

Studies have not yet been published that assess the use of the BAI-PC for evaluating treatment outcomes. However, the measure is potentially useful for monitoring change over time, particularly given that the full form of the BAI has been used to assess treatment outcomes. Additional research is needed using the brief form (BAI-PC) in a primary care sample.

**Cross-Cultural Applicability**

The BAI-PC has not yet been studied in diverse primary care populations. However, the full version of the BAI has been evaluated in a sample of individuals selected to be representative of the demographic characteristics of the US population (Gillis, Haaga, & Ford, 1995), and in older adults in primary care (Wetherell & Arean, 1997). The full BAI is available in a Spanish translated version.

**DISEASE-SPECIFIC CONSIDERATIONS**

Physiological symptoms of anxiety are often associated with general medical conditions or effects of substance use and can inflate the scores of anxiety self-report measures, particularly those that are sensitive to the autonomic symptoms of anxiety. Therefore, when assessing for anxiety, providers must rule out possible medical and substance-related causes. In order to do so, reviewing patient history is crucial. Although the onset of anxiety disorders can vary, typical onset is from late adolescence to mid-thirties (American Psychiatric Association, 1994). Additionally, if atypical symptoms (vertigo, amnesia, headaches, slurred speech, loss of consciousness, and loss of bladder or bowel control) accompany anxiety, a medical or substance-induced cause should be investigated, particularly in cases of onset of autonomic symptoms before mid-adolescence and or after age 45 (DiBartolo, Hofmann, & Barlow, 1995).

Many chronic physical conditions are exacerbated or maintained by high levels of anxiety including gastric ulcers, hypertension, migraines, coronary artery disease, asthma, and pain conditions (e.g., White & Barlow, 2001). Medical conditions that may imitate an anxiety disorder include endocrine disorders (e.g., hyperthyroidism, hypoglycemia, pheochromocytoma, Cushing’s syndrome, menopause), cardiovascular disorders (i.e., myocardial infarction, mitral valve prolapse, arrhythmias,
congestive heart failure, hypertension), neurological disorder (e.g., epilepsy, Huntington’s disease, multiple sclerosis, vestibular disorder) and nutritional deficiencies (folate, pyridoxine, vitamin B12, or vitamin D; Esler, Weisbeg, Sciamanna, & Bock, 2004; Kavan, Elsasser, & Barone, 2009). It is possible that effective, early identification and treatment of anxiety may prevent comorbid conditions that develop later (Noyes, 2001).

An additional cause of anxiety may be pharmaceutical. Therefore, in addition to assessing for the anxiety problem areas identified in Table 12.1, it is essential that a careful medication history be obtained to determine if there may be a link between medication/substances and the patient’s symptoms. A history of all medications and substances, including prescription and over-the-counter medications, illicit drugs, herbal remedies, and diet drugs should be obtained. The following medications or reactions to them may exacerbate (or produce) anxiety: anticholinergic medications; sympathomimetic medications (decongestants, B-2 bronchodilators, weight-reducing agents); thyroid hormone; xanthine-containing medications (bronchodilators with theophylline, many over-the-counter cold and arthritis remedies; Esler et al., 2004; Kavan et al., 2009; Roca, 1991). Additionally, the following medications have anxiety as an adverse psychiatric effect: benzodiazepines, beta blockers, opioid analgesics, SSRIs, sulfonamides, chloroquine (Aralen), ephedrine, interferon alfa, levonorgestrel (Norplant), mefloquine (Larium), methyl dopa (Aldomet), methylphenidate (Ritalin), metronidazol (Flagyl), sumatriptan (Imitrex), and thyroid hormone (Synthroid; Gunning, 2004).

Substances or withdrawal from substances may also result in symptoms of anxiety. Substances that may mimic symptoms of anxiety include sedatives/hypnotics, alcohol, caffeine, nicotine, cannabis, hallucinogens, and stimulant drugs of abuse (cocaine, amphetamines), benzodiazapines, and opioid analgesics (Esler et al., 2004; Kavan et al., 2009). If symptoms persist after 4 weeks of withdrawal from substances, panic disorder or other anxiety problems should be given stronger consideration (Ferguson, 2000). Taken together, if a medical condition or substance-induced etiology is suspected, it is helpful to involve primary care physicians to help identify or rule out nonpsychiatric etiologies that could be contributing to symptoms reported on the anxiety screening measures.

STRENGTH AND LIMITATIONS OF ANXIETY MEASURES IN PRIMARY CARE SETTINGS

As alluded to earlier in this chapter, there are many advantages to using the GAD-7 in primary care. An obvious strength of using the GAD-7 is its simplicity and ease of scoring. It is readily adaptable to the rapid pace of primary care, and is already widely used and researched in this population. Patients can clearly understand its content and instruction, and providers can quickly become familiar with scoring and interpretation procedures. In addition, the GAD-7 is public domain, which is a major strength for many sites where finances are often an issue. Further, the GAD-7 has demonstrated good psychometric properties across various standardization samples including many cultural groups. For these reasons, the VA Healthcare System in addition to many other organizations consider the GAD-7 to be the standard of practice to assess for anxiety in primary care.

It should be noted, however, that the frequency of anxiety symptoms, including somatic and cognitive-affective components, may not fully reflect the clinical significance of anxiety presentations (Ito et al., 2014). Accordingly, one limitation of the GAD-7 is that it does not incorporate avoidance behavior or specify functional impairment in the measure’s total score, and thus providers may need to inquire about these factors separately. Nonetheless, the GAD-7 remains one the most frequently used measures in primary care today, with noted benefit for patients and the providers who serve them.
Similarly, there are advantages and disadvantages to using one of the other measures discussed in this chapter as well. The inclusion of avoidance behavior and functional impairment, for example, is a strength of the OASIS. In addition, like the GAD-7, the OASIS is easy to administer and score and is available free of charge (public domain). The OASIS, however, is a relatively new measure and is in need of further psychometric research across diverse populations. The PROMIS is also less studied in primary care or for evaluating treatment outcomes, but has the advantage of extensive psychometric development research and availability in CAT format, which draws items from a large data pool and administers them in a manner that is tailored to the individual based on their prior responses (Kroenke et al., 2007). However, these are also available to be administered in a standard static format and with brief forms (e.g., four to eight items), which increases their widespread feasibility (Kroenke et al., 2007). This measure is also considered public domain and available online (http://nihpromis.org).

The BAI-PC may be useful for assessing depression and PTSD as well as more general anxiety symptoms; however, this limits the discriminability of the measure. Another benefit of the BAI-PC is that it is derived from the 21-item BAI, which has been extensively studied across a variety of populations. However, it is not known the extent to which these results apply to the short form of the measure. Further, there are important gaps in the psychometric evaluation of the BAI-PC, including its use across different cultural groups and use as an indicator of treatment outcome. The BAI-PC also does not incorporate avoidance behavior or functional impairment in the measure’s total score. Taken together, the OASIS, PROMIS Anxiety, and BAI-PC have been used less frequently as a measure of treatment outcome in published studies compared to the GAD-7. As such, many purport that the GAD-7 remains the gold standard for assessing anxiety in primary care until additional research is conducted.

**CASE STUDY**

A case study illustrating how anxiety screening measures are useful in screening and tracking anxiety problems in primary care is presented here. As the GAD-7 is a well-researched and well-accepted anxiety screening measure in primary care, it was used in this case example.

The patient is a 42-year-old Persian Gulf War veteran with a medical history of hypertension, hyperlipidemia, migraine headaches, tinnitus, and lumbar radiculopathy. He initially presented to the emergency department (ED) with chest pain, palpitations, weakness, complaints of worsening dizziness, and jitteriness. He was concerned that he may be having a heart attack. A medical workup in the ED was completed with normal electrocardiogram (EKG) and normal cardiovascular and neurological findings (normal MRI). The patient was prescribed Ativan and discharged home with a follow-up appointment with his primary care provider. At his follow-up appointment, his primary care provider detected no neurological, endocrine, or cardiovascular abnormalities and consulted primary care-mental health integration (PC-MHI). He had no prior mental health treatment.

PC-MHI psychologist met with the patient during the primary care visit in response to real-time consultation request. He reported an increase in anxiety over the past year related to psychosocial stressors, but reported a worsening of symptoms over the past three months. Stressors included recent abnormal lab results last summer including elevated liver enzymes and glucose, after which he became overly focused on his health and was frequently worrying about worst possible health outcomes. Concerns about his health led him to make positive health-related changes including quitting drinking, increasing exercise, and making healthier nutrition choices, which has resulted in weight loss. Labs were repeated and all were within normal limits. In addition to
health-related stressors, he indicated the presence of family-related stress. He reported that his oldest son suffered from schizophrenia and has made multiple suicide attempts, and that he used illegal drugs and had been treated multiple times through inpatient psychiatric and residential treatment facilities. This reportedly placed a great deal of stress on the patient and his wife.

In order to deter catastrophic interpretations of symptoms, the patient was provided with education on the causes of panic and the physiology underlying anxiety/panic-related symptoms, and why panic attacks can appear to come out of nowhere. He was given a brief instruction on diaphragmatic breathing, was encouraged to increase physical activity (walking four times a week with his wife), and reduce caffeine intake. The PC-MHI psychiatrist was consulted for psychotropic medication recommendations and SSRI treatment for anxiety was initiated. Initial GAD-7 score obtained at this visit was 19, with the patient indicating that symptoms of anxiety have made doing his work, taking care of things at home, and getting along with others very difficult.

The patient presented for follow-up sessions to continue to address anxiety-related symptoms. Consistent with the PC-MHI model, a brief CBT approach was utilized with the patient identifying treatment goals, completing tracking logs to help identify triggers, physiological effects, and maladaptive thoughts/behaviors that perpetuate his anxiety. Brief therapy focused on increasing insight into triggers, challenging maladaptive thoughts, and persisting with physical activities despite producing symptoms resembling panic sensations (i.e., in vivo interoceptive exposure). One month later, he had not experienced any full panic attacks and was following through with exercise, reducing caffeine intake, practicing breathing exercises, and completing tracking logs. At that time, the GAD-7 score was 7. At 2 months from the initial visit (five sessions total), the patient’s GAD-7 score was 5 (a significant decrease from his initial score of 19) and he was panic-free.

This case illustrates that anxiety can be addressed efficiently with brief evidence-based behavioral medicine procedures in primary care. The GAD-7 efficiently captured anxiety symptoms and their change over time, and thus served to guide treatment cessation or continuation decisions. Maintenance of therapeutic change can also be tracked through readministration of the GAD-7 during routine primary care follow-up appointments. Additionally, multiple ED visits for unexplained chest pain are likely for many patients with panic anxiety (Katerndahl & Realini, 1995). Fortunately, in this case, development of full panic disorder was averted, and unknown but likely cost savings were realized in term of future medical utilization and quality of life.

SUMMARY

In this chapter many different assessment tools were examined in depth to aid the provider with the challenges in identifying and diagnosing GAD among other anxiety disorders in primary care settings. Of the four inventories discussed in detail, the GAD-7 appears to have the most research and psychometric support. The goal of the GAD-7 is to identify possible cases of GAD and assess the severity of the disorder. Given the high prevalence rates of GAD in primary care, availability of a quick clinical measure for GAD is imperative. The main strengths of the GAD-7 are its generalizability to primary care, its being validated in a large sample across cultural groups, and its brevity and efficiency in purpose. It is also public domain, which sets it apart from some of its predecessors (e.g., BAI). Chief among its weaknesses are that it does not incorporate avoidance behavior or specify functional impairment in the measure’s total score.

This chapter also details the OASIS, PROMIS Anxiety, and BAI-PC, each with their own associated strengths and weaknesses. While promising, these measures have been used less frequently as a
measure of treatment outcome in published studies compared to the GAD-7 and thus many purport that the GAD-7 remains the gold standard for assessing anxiety in primary care. Lastly, this chapter offers a table (Table 12.2) that includes other anxiety measures deemed appropriate for use in primary care given that each primary care setting will have unique clinical, operational, and fiscal demands on it, and no one method is the best for all primary care practices. Nevertheless, problematic anxiety that goes undetected and untreated contributes to overall poor health and unnecessary healthcare expense. Systematic use of the GAD-7 for both anxiety screening and measurement of treatment outcome is a sound start in addressing problem anxiety in the primary care setting.

REFERENCES


While the prevalence of the US population reporting high stress levels has recently plateaued, 44% report that their stress has increased in the last five years, and 22% report currently experiencing extreme stress (American Psychological Association, 2012). Likewise, patients reporting high stress have been shown to more frequently utilize healthcare resources, incur increased healthcare expenditures, and be at increased risk for a number of adverse physical and psychological consequences, including the development of chronic health problems such as depression, anxiety, cardiovascular diseases, and diabetes (Jackson, Knight, & Rafferty, 2010; Pirraglia, Hampton, Rosen, & Witt, 2011). For example, recent studies have suggested that anxiety, depression, and somatic disorders are 2.5 times more likely to occur in patients who have experienced a psychosocial stressor in the preceding 12 months (Herzig et al., 2012), with both additive and dose-dependent effects reported with the addition of multiple stressors (Haftgoli et al., 2010). In addition, taking into account the prevalence of high-risk health behaviors such as sedentary lifestyle, high rates of chronic illness among primary care patients are not surprising. Given the inability of lower socioeconomic status populations to control the occurrence of common chronic stressors such as poor living conditions and decreased financial resources, chronic stress and its sequelae may be an unavoidable part of their day-to-day existence (Lantz, House, Mero, & Williams, 2005).

Interventions targeting stress reduction have been shown to improve health outcomes, reduce depression and anxiety symptoms, and improve measures of mental health and health-related quality of life (Brantley, Veitia, Callon, Buss, & Sias, 1986; Fjorback, Arendt, Ørnbøl, Fink, & Walach, 2011; Mynors-Wallis, Gath, Lloyd-Thomas, & Tomlinson, 1995; Robinson et al., 1994). However, the ability of physicians to correctly identify those in need of specialized referrals has been called into question (see Chapter 1). Some have suggested that routine assessment of stressful life events may be an effective tool to assisting physicians in making this determination (Campbell, Seaburn, & McDaniel, 1994).

Moreover, although the involvement of psychologists in the management of primary care patients may represent optimal care, this situation may not always be practical. For these reasons, the development of stress assessment techniques may be of great benefit in improving health outcomes, reducing healthcare utilization, reducing levels of chronic stress and resultant physical and psychological comorbidity, and assessing the design of stress management programs specifically targeted for this population.

The Daily Stress Inventory (DSI) and subsequently developed Weekly Stress Inventory (WSI) are two minor life events measures developed to assess the differential contributions of minor stress to the prediction of physical illness and psychological symptomatology. The DSI and WSI contain not only measures of life events but also measures of stress appraisal, thus combining both psychological and environmental perspectives. These two measures are not limited by wide contiguous intervals, are relatively easy to administer, and have direct application to primary care populations. Both instruments will be discussed in this chapter, following an overview of the various theoretical models of stress.
OVERVIEW OF THE CONCEPT OF STRESS

While the origins of the concept of stress date back to Hippocrates, the construct has been marked by broad variations in the physiological, behavioral, and psychological elements actually used to define stress. In the Middle Ages, stress described the social hardship and economic adversity prevalent at the time; by the early 1900s, William Cannon’s (1939) research on biobehavioral survival mechanisms and resultant “fight or flight” theory led to the concept of homeostasis, defined as “the coordinated physiological process which maintains . . . steady states in the organism” (p. 24). In the early twentieth century, Hans Selye began his pioneering research focusing on the behavioral and physiological aspects of stress, eventually leading to an interest in the recognition and systematic investigation of the importance of behavioral factors in the study of stress (Everly, 1989). Selye posited that a “general adaptation syndrome” (GAS) occurs within an organism when confronted by “diverse nocuous agents” and defined the concept as the “nonspecific result of any demand upon the body” (1936). Over time, these demands produce a biological syndrome marked by a triad of physiological changes in the adrenal glands, the lymphatic system, the stomach, and the upper intestinal tract. These biological indicators become evident in a stereotyped fashion subsequent to exposure to any type of somatic or psychological stress, including blood loss, fatigue, pain, ingestion of toxins, emotional arousal, fear, concentration, and great elation (Selye, 1982). The GAS response thus occurs in three discrete stages: “alarm,” “resistance,” and “exhaustion,” and responses to various types of stressors led Selye to distinguish “eustress,” or positive stress, from “distress,” or negative stress. This distinction is evident only in the nature of the stressor itself, however, and not in the body’s response to any particular stressor. Contemporary theorists continue to include Selye’s GAS among the most highly regarded descriptions of the stress response (Everly, 1989).

Building on the foundations created by the basic sciences, social scientists quickly became interested in the stress concept, and adopted the term stress to describe social demands and disruptions (Lazarus & Folkman, 1984a). In 1966, Lazarus suggested that stress be considered a subdiscipline within psychology. Additional developments included the recognition of stress as a contributing factor in psychosomatic illness in the first edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1952), as well as in the emergence of journals dedicated to the study of stress, such as the Journal of Human Stress, Psychophysiology, and the Journal of Traumatic Stress. The diversity of opinion surrounding the definition of stress has prevented “stress” from becoming a universally defined or accepted construct, with some researchers arguing that the concept of stress is too broad and ambiguous to adequately define (Engel, 1985). Despite this criticism, investigators have attempted to define the nature of stress, primarily described in terms of stimulus, response, or interactional theories.

Cannon’s work on homeostasis was the first to identify stress as a stimulus, meaning any event that prepares the organism for the “fight or flight” response (1939). This approach highlights the objective nature of stress, and applies the term “stressor” to the specific internal, external, psychological, and biological events that produce the stress response (Everly, 1989; Lazarus & Folkman, 1984a). While stimulus definitions of the stress response may provide a useful taxonomy, the prevalent view among stress researchers today is that individual differences in stress appraisal are also important considerations (Lazarus & Folkman, 1984b).

In contrast to the stimulus approach, other stress theorists have defined stress as the response an organism makes to environmental changes. Selye (1974) defined stress as the “nonspecific response of the body to any demand.” In a similar vein, Everly (1989) discussed stress in a biological framework, defining it as a “physiological response that serves as a mechanism of mediation, linking any
given stressor to its target organ effect or arousal” (p. 15). Lacey (1950) also noted the importance of specificity in the response mechanism. Specificity refers to the notion that different individuals will respond to the same stressor with differing physiological reactions.

The primary hypothesis of the physiological representation of stress involves the sympatho-adrenomedullary (SAM) and hypothalamo-pituitary-adrenocortical (HPA) systems as mediators of stress responses. Stress-induced activation of the HPA axis results in a series of neuroendocrine changes referred to as the “stress response” or “stress cascade.” This response is described as regulatory in nature, such that it permits the organism to make the physiological and metabolic changes necessary to maintain homeostasis (McEwen & Gianaros, 2010). In humans, this response is initiated with the release of corticotropin-releasing hormone (CRH) from the hypothalamus in response to a discrete stressor. Adrenocorticotropic hormone (ACTH) release by the pituitary gland is then stimulated, which acts on the adrenal cortex to release the glucocorticoids cortisol and corticosteroid into the bloodstream. Under normal conditions, the glucocorticoids act in a negative feedback loop to terminate release of CRH (Miller & O’Callaghan, 2002).

The lack of attention paid by stimulus and response definitions to individual differences subsequently led to interactional descriptions of stress, which focus on the relationship between the individual and the environment in producing the stress response. For example, Wolff (1953) first pointed out that stress is a “dynamic” state dependent on the interaction between an organism and its aversive external environment. Lazarus expanded the interactionist theory, postulating that the perception and management of stressful events are defined by two interrelated processes, appraisal and coping (Lazarus & Folkman, 1984a; Lazarus, 1966). Specifically, situations are appraised in terms of their probable outcomes, and negative situations can be interpreted one of three ways: (1) threat situations, which are anticipated to produce harm; (2) harm-loss situations, which are evaluated as having already produced harm; or (3) challenge situations, which have the potential for either harm or gain (Lazarus, DeLongis, Folkman, & Gruen, 1985).

Coping refers to a variety of methods implemented by the individual in an effort to manage stressful situations. Problem-focused coping includes strategies that enable the individual to prevent stressful events from occurring, or that enable the individual to successfully avoid or resolve any difficulties that do occur. Emotion-focused coping includes strategies that moderate stress-induced emotions and related physiological arousal (Cameron & Meichenbaum, 1982). The transactional model thus suggests that individuals “actively define and shape stressful transactions by means of their cognitive appraisals and their coping responses” (p. 696).

Arousal has been described in theories of personality, performance, motivation, and attention, and has been used to identify changes in the responsiveness of subjects to various types of environmental conditions (Eysenck & Eysenck, 1985; Strelau, 1987; Strelau, 1985). It has been hypothesized that individuals who exhibit higher levels of arousal may also have an increased susceptibility to stress, and measurements of arousal have been associated with the inability to habituate to repeated autonomic stimulation, such as repeated exposure to environmental stressors (Lacey & Lacey, 1958). Stressors are capable of producing both central and peripheral physiological arousal, as evidenced by studies examining the effects of the stress response on the sympathetic nervous, cardiovascular, and neuroendocrine systems via measurements of electrodermal, electromyography, and hormonal indices. Physiological changes occurring during stress-induced arousal include increased heart rate, spleen contraction, glycogen-glucose transfer and release by the liver, increased blood flow to the brain and muscular system, increased respiration, and pupil dilation (Cox, 1978). These physiological adaptations are thought to increase the organism’s resources for responding to threatening stimuli, and are suggestive of Selye’s general adaptation syndrome. Similarly, the effect of arousal on the neuroendocrine system is marked by a variety of hormonal responses, most commonly the increased
secretion of epinephrine and norepinephrine, adrenocorticotropin hormone, cortisol, and corticosteroid. The release of glucocorticoids result in increased production of glucose and urea, release of free fatty acids into the bloodstream, suppression of immune system functions, and increased production of ketones (Everly & Rosenfield, 1981).

Research of life events has stemmed from stimulus theories of stress, a view that suggests that stressful life events impact certain illnesses (Brown & Harris, 1989). Researchers and clinicians have therefore used a variety of measures to assess the impact of stress on psychiatric and medical populations, and research has indicated that the effect of minor stressors on the progression of physical and psychological illness may be greater than the influence of major stressors (Brantley & Jones, 1993; DeLongis, Coyne, Dakof, Folkman, & Lazarus, 1982; Martin & Brantley, 2004). Psychological factors, including life stress and coping, are believed to affect health primarily through direct physiological mechanisms or the alteration of health-related behaviors. Brantley and Garrett (1993) summarize the proposed models of stress and illness that include changes in physiological functioning, increased high-risk behavior, decreased resistance to disease, neurological hypersensitivity, or inadequate coping. Investigations examining the specific relationship between stress and illness have consistently reported correlations between psychological distress and symptom presentation of both acute and chronic illness, with the most consistent evidence found for infectious diseases, cancer, cardiovascular disease, and chronic conditions such as diabetes, asthma, and gastrointestinal disorders (Martin & Brantley, 2004; McEwen & Stellar, 1993).

Chronic illness has been cited as the most prevalent of all the major life stressors (Felton, 1990). Diabetes mellitus, which co-occurs with overweight and obesity in 96% of patients, is a chronic endocrine disease that significantly increases morbidity and mortality and continues to be among the top causes of death in the United States (National Center for Health Statistics, 2016). Glycemic control in diabetics has been shown to be adversely affected by stress via activation of the HPA axis and subsequent secretion of glucose counterregulatory hormones (Sulway, Tupling, Webb, & Harris, 1980). Stress-induced release of growth hormone by the pituitary gland can also cause insulin resistance and sympathetic stimulation of pancreatic hormones (Surwit, Ross, & Feinglos, 1991).

Bjorntorp (1993) has proposed that chronically elevated activation of the HPA axis occurring secondary to psychological stress produces discrete, periodic elevations of cortisol secretion on a daily basis, which over time produces a neuroendocrine perturbation syndrome marked by multiple abnormalities including insulin resistance, hyperinsulinemia, hypertension, dyslipidemia, and accumulation of body fat to central depots—all hallmark features of the metabolic syndrome. Bjorntorp has alternatively referred to this set of symptoms as a “Civilization Syndrome,” highlighting the pressures of modern, competitive lifestyles in the generation of chronic stress (see Figure 13.1), and the

![Figure 13.1 Civilization Syndrome](image-url)
increased prevalence of unhealthy behaviors as contributing factors, such as increased tobacco and alcohol consumption, overeating, and physical inactivity (Bjorntorp, Holm, & Rosmond, 1999, 2000; Bjorntorp & Rosmond, 1999, 2000).

Tobacco and alcohol abuse are often a maladaptive attempt to cope with stressful situations, and stress can maintain the use of these substances (Best, Wainwright, Mills, & Kirkland, 1988; Feverstein, Labbe, & Kuczermierczyk, 1986; Williams, Stinson, Parker, Harford, & Noble, 1987). Nicotine has been shown to potentiate sympathetic arousal, and smoking compromises physiological systems (e.g., cardiopulmonary and immune) susceptible to stress (McGill, 1988; Trap-Jensen, 1988). Evidence also suggests that stress may be an important factor in predicting alcoholism, and individuals who become alcoholics may lack skills in stress management (Brantley & Garrett, 1993; Cotton, 1990).

In addition to the stress-induced physiological changes that contribute to poor health outcomes, there are also stress-induced behavioral changes that appear to impact health behaviors. Ng and Jeffrey (2003) reported a significant association between perceived stress, a higher fat diet, and lower physical activity in a large sample of adults. Smith, Baum, and Wing (2005) prospectively examined the interplay of stress, weight change, and health behaviors in parents of children with cancer. Parents of children with cancer had higher levels of distress, anxiety, and depression; consumed fewer calories; engaged in less frequent physical activity; and subsequently gained more weight compared to parents of children without cancer. Brownell (1982) has posited that the increase in obesity and stress-related disorders over the past century has resulted from sedentary lifestyles, and a renewed interest in exercise during the past several decades has paralleled research suggesting that physical fitness is a significant stress moderator (Brandon & Loftin, 1991; Roth & Holmes, 1985). Moreover, Everly (1989) contends that exercise, more than any other stress management strategy, prevents disease by ventilating the pathophysiological changes associated with the stress response.

MEASUREMENT OF STRESS

Several methods for assessing stress have stemmed from the various theories of stress; however, the primary measures that have been examined in clinical populations are subjective, self-report questionnaires of major and minor life events. Several major life events scales have been developed and tested, including the Schedule of Recent Experiences (Hawkins, Davis, & Holmes, 1957), the Social Readjustment Rating Scale (Holmes & Rahe, 1967), and the Life Experiences Survey (Sarason, Johnson, & Siegel, 1978). These scales measure the impact of relatively infrequent major stressors, such as a divorce, job loss, or the death of a loved one.

Traditionally stress research has focused on major life events; however, more contemporary stress theorists have studied the impact of minor life events, termed daily stressors or hassles, on health and behavior, because of the frequency with which they occur relative to major stressors (Brantley & Jones, 1989; DeLongis et al., 1982). Examples of minor life events include having an argument with a spouse, getting caught in traffic, or running out of spending money. The Hassles Scale, developed by Kanner and colleagues in 1981, was the first scale to assess minor life events by measuring the severity and frequency of minor stressors over the past month on a 3-point Likert-type scale. Similarly, the same research group developed the Uplifts Scale, an index of desirable minor life events (Kanner, Coyne, Schaefer, & Lazarus, 1981).

Dissatisfied with event specific measures, in 1983 Cohen and colleagues constructed the Perceived Stress Scale (PSS) (Cohen, Kamarck, & Mermelstein, 1983). The PSS is consistent with cognitive-based, interactionist stress theories and measures a respondent’s appraisal of the global stress level in his or her life. The developers have reported the PSS to have adequate reliability and validity,
and an abbreviated version is available. However, while the authors purport that the predictive power of the PSS is greater than life events scores, opponents of this approach have argued that PSS scores are often confounded with both psychological and health-related outcome measures (Lazarus et al., 1985). Extending the focus on minor stressors, the remainder of this chapter will review the Daily and Weekly Stress Inventories, two scales developed by Brantley and colleagues to measure the frequency and impact of minor stressors likely to occur on a daily or weekly basis.

**Daily Stress Inventory**

In 1987, Brantley, Waggoner, Jones, and Rappaport published the Daily Stress Inventory (DSI), a 58-item questionnaire measuring the frequency and impact of minor stressors likely to occur on a daily basis. The authors developed the instrument to serve as both an outcome and predictor variable in investigations of etiological and mediating relationships between minor life events, stress-related physical and psychological disorders, and stress management interventions. They theorized that a reliable and valid measure of minor stress and appraisal may be more useful in elucidating these relationships due to their temporal proximity to the onset, exacerbation, and recurrence of symptoms (Brantley & Jeffries, 2000; Brantley, Catz, & Boudreax, 1997; Brantley, Waggoner, Jones, & Rappaport, 1987). Development and validation studies of the scale occurred over a number of years, and were initially based on the senior author's interest in examining the relationships between stress and tension-type headaches.

Preliminary items for the DSI were derived from 85 daily stress diaries completed by patients attending a primary care-based clinical psychology clinic; community volunteers; and psychology doctoral students. The initial 71-item version of the scale was piloted in 164 community adults and 254 undergraduate psychology students. Items were then revised to avoid both confounding items and those overtly referring to either a physical or psychological illness. The final 58-item version of the scale was then created, with items retained being endorsed by at least 15% of the sample, and corrected item-total correlational coefficients of at least .30. The following are example items from the DSI: (22) “Was embarrassed”; (40) “Had car trouble”; (51) “Was late for work/appointment.” Content clusters of DSI items were derived based on outside expert review, and consist of the following categories: “Interpersonal Problems”; “Personal Competency”; “Cognitive Stressors”; “Environmental Hassles”; and “Varied Stressors.”

The DSI assesses minor life events that have occurred over the past 24 hours, and respondents rate each stressor on an 8-point Likert scale ranging from “0 = did not occur” to “7 = caused me to panic.” Three scores are calculated: the “Event Score,” which is the frequency of stressors experienced, and corresponds to the number of items endorsed; the “Impact Score,” which is the respondent’s subjective appraisal of the severity of each stressor, and corresponds to the sum of the ratings for each item endorsed; and the “Impact/Event Ratio Score,” which is the average impact score for any particular day (Brantley & Jeffries, 2000; Brantley, Catz et al., 1997; Brantley et al., 1987).

**Psychometric Properties and Studies Utilizing the DSI**

Initial reliability and validity data were collected by administering the DSI to 433 nonstudent community-dwelling adults. Demographic information indicated that the sample was aged 17 to 77 years (M = 34.9, SD = 12.5), female (56%), married (56%), employed full-time (60%), and had at least some college education (69%). Internal consistency reliability was reported to be .83 for the Event score and .87 for the Impact score. Temporal stability and sensitivity to change over time were examined by administering the DSI to a sample of 35 respondents every day for 28 days. Test-retest reliability
calculated for the Event and Impact scores were .72 and .41, respectively, indicating that individual stressors demonstrated acceptable test-retest reliability (Event Score), but that respondent appraisals of the impact of those stressors were variable over the 28 days (Impact Score). This suggests that the DSI Impact Score demonstrates the variability desired in a measure designed to measure stress on a frequent, daily basis. In addition, the authors reported that single, one-day Event and Impact scores were not predicted by scores reported on past or future days, and consecutive 7-day administrations of the DSI demonstrated expected variability of scores between working days and nonworking days (Brantley, Catz et al., 1997; Brantley et al., 1987).

Convergent validity was demonstrated with endocrine measures of stress by administering the DSI daily for 10 consecutive days to a sample of 18 volunteers. High and low stress days, as measured by the DSI, correlated significantly with changes in both urinary cortisol and vanillylmandelic acid, a metabolite of epinephrine and norepinephrine (Brantley, Dietz, McKnight, Jones, & Tulley, 1988). Likewise, the DSI demonstrated both convergent and concurrent validity by positive associations with the Hassles Scale (Kanner et al., 1981), and a subjective global rating of stress; and divergent validity with the Uplifts scale (Kanner et al., 1981) and a measure of hostility (Brantley & Jeffries, 2000; Brantley, Catz et al., 1997).

The initial studies of the DSI indicated that the scale is sensitive to change, and able to detect daily changes in stress-related symptoms of asthma and anxiety in asthma patients (Goreczny, Brantley, Buss, & Waters, 1988); blood glucose levels in diabetes patients (Goetsch, Wiebe, Veltum, & Van Dorsten, 1990); bowel and pain symptoms in patients with Crohn’s disease (Garrett, Brantley, Jones, & McKnight, 1991); and symptom exacerbations in patients with lupus (Adams, Dammers, Saia, Brantley, & Jones, 1993), migraine headaches, and tension-type headaches (Mosley et al., 1991).

Norms were derived by administering the DSI to the community sample described earlier (N = 433), as well as to a second sample of 233 medical outpatients, and are available in the administration and scoring manual for the DSI (Brantley & Jones, 1989). Demographics for the medical sample indicate that respondents had a mean age of 45.6 years, were 55% Caucasian, 58% female, 73% married, and 51% employed, and 53% had less than a high school education.

Recent studies have reported that DSI Impact scores are associated with dietary restraint and binge eating behavior in a large sample (N = 497) of undergraduate females (Woods, Racine, & Klump, 2010); and that respondents who met criteria for high clinical risk for psychosis (N = 314) in the North American Prodrome Longitudinal Study, Phase 2 (NAPLS-2) reported higher DSI Event and Impact scores compared to healthy controls (Trotman et al., 2014). Smaller studies of chronic migraine and tension-type headache patients found that 2 consecutive days of high stress as measured by the DSI were strongly predictive of headache symptoms, whereas 2 days of low stress provided a protective effect (Houle et al., 2012). In another study, 3-month DSI score reductions were observed compared to normal controls following an 8-week mindfulness-based stress management intervention (Williams, Kolar, Reger, & Pearson, 2001). Also, in a sample of patients diagnosed with pathological gambling, gambling urges were independently predicted by high DSI scores (Williams et al., 2001).

**Weekly Stress Inventory**

In 1997, Brantley and colleagues published the Weekly Stress Inventory (WSI) in order to conduct stress and health-related assessments over intermediate intervals that would bridge the time frames referenced by the Hassles Scales (1 month) and the Daily Stress Inventory (1 day) (Brantley, Jones, Boudreax, & Catz, 1997). The WSI is an 87-item questionnaire assessing the frequency and impact of minor stressors likely to have been experienced in the past week. Items were derived from the Daily Stress Inventory (Brantley et al., 1987), the Hassles Scales (Kanner et al., 1981), the Inventory of Small
Life Events (Zautra, Guarnaccia, & Dohrenwend, 1986), and from items written in by respondents to the aforementioned scales. Items were selected to represent eight stress domains suggested by outside expert reviewers: “Work,” “School,” “Money,” “Transportation,” “Marital/Family,” “Household,” “Personal,” and “Leisure”; and based on whether they occurred during a discreet, observable timeframe, whether they were likely to occur within one week, and whether they were considered undesirable life events. Items were then revised to avoid confounding items and those overtly referring to a physical or psychological illness. The following are example items from the WSI: “(6) Hurried to meet a deadline”; “(16) Ran out of pocket money”; “(51) Argued with a friend.” The final 87-item scale was piloted on 100 community-dwelling adults, with items retained being endorsed by at least 10% of the sample, and corrected item-total correlational coefficients of at least .25.

The WSI assesses minor life events that have occurred over the past seven days, and respondents rate each stressor on an 8-point Likert scale ranging from “0 = did not happen” to “7 = extremely stressful.” Two scores are calculated: the “Event Score,” which is the frequency of stressors experienced and corresponds to the number of items endorsed; and the “Impact Score,” which is the sum of the respondent’s subjective appraisal of the severity of each stressor (Brantley, Jones et al., 1997).

### Psychometric Properties and Studies Utilizing the WSI

Initial reliability and validity data were collected by administering the WSI to 522 community-dwelling adults matched to 1980 census data in terms of age, race, and sex, and to smaller samples of college students, headache patients, coronary heart disease patients, and hemodialysis patients (Brantley, Jones et al., 1997). The instrument has demonstrated good internal consistency ($\alpha = .92–.97$), test-retest reliability ($r = .80–.83$) when administered both over several hours and at the end of one week, and convergent and concurrent validity with the Daily Stress Inventory ($r = .77–.84$), the Hassles Scale ($r = .61–.69$), the Perceived Stress Scale ($r = .57–.54$), and the Hemodialysis Stressor Scale ($r = .53–.50$) (Baldree, Murphy, & Powers, 1982).

Initial studies of the WSI indicated that scale scores are associated with stress-related symptoms of anxiety and depression in hemodialysis and HIV positive patients (Carmack et al., 1995; Thomason, Jones, McClure, & Brantley, 1996); with somatic complaints and treatment noncompliance in hemodialysis patients (Carmack et al., 1995; Hitchcock, Brantley, Jones, & McKnight, 1992); with inflammation in rheumatoid arthritis patients (Thomason, Brantley, Jones, Dyer, & Morris, 1992); and with depression, somatic complaints, and Type A behavior in coronary heart disease patients (Mosley et al., 1996).

Norms were derived for the WSI from the standardization sample of 522 community-dwelling adults matching 1980 census data, and no differences were found in either Event or Impact scores for race and sex. Similar to what has been observed with other stress measures, scores for all age groups declined with age, such that in 15- to 19-year-olds, mean scores for Event = 38.25 (SD = 21.47) and Impact = 189.39 (SD = 99.50); whereas for those 60 years and older, mean scores for Event = 24.65 (SD = 20.75) and Impact = 78.68 (SD = 94.52) (Brantley & Jeffries, 2000; Brantley, Jones et al., 1997).

More recent studies of the WSI have indicated that psychological assessment across a 6-month period provides a stable indicator of minor stress in a sample of adults recruited from primary care medical clinics, and administering the WSI five to seven times over the course of 1 year approximates average stress levels over the same period of time (Scarinci, Ames, & Brantley, 1999). WSI scores have also been shown to predict higher medical utilization in a large sample of low-income African American primary care patients over 5 years (Brantley et al., 2005). In addition, WSI Event and Impact scores have been found to have significant positive correlations with exercise frequency, intensity, and duration in a sample of college-aged females ($N = 95$) determined to be at the maintenance...
stage-of-change for exercise, whereas those in other stages of change had either null or negative associations (Lutz, Stults-Kolehmainen, & Bartholomew, 2010). Finally, in two recent studies utilizing the WSI in the Jackson Heart Study \((N = 4,863)\), high WSI scores have been associated with blood pressure progression in 922 African Americans whose blood pressure increased in severity between 2005 and 2008, prevalence ratio \(= 1.21; 95\% \text{ CI } 1.04–1.40\) (Ford et al., 2016); and with lower odds ratios of both sleep duration and sleep quality (Johnson et al., 2016).

A few authors have used the WSI to develop other measures of minor life events, including stress scales targeting adolescent populations (Self-Brown, LeBlanc, & Kelley, 2004) and individuals with physical disabilities (Rhode, Froehlich-Grobe, Hockemeyer, & Carlson, 2012). In addition, Brantley and colleagues published initial psychometric properties of a shorter, 25-item version of the WSI, called the Weekly Stress Inventory-Short Form (WSI-SF) (Brantley et al., 2007). Items were selected for the WSI-SF based on principal components analysis of two large studies utilizing the WSI \((N = 522\) and 362), and retained based on whether they reflected large percentages of variance accounted for, and endorsement by at least 10\% of each sample. Initial reliability and validity data were reported for the WSI-SF for an independent sample of 171 primary care medical patients. Internal consistency reliabilities for the WSI-SF Event and Impact scores were .91 and .92, respectively. Convergent validity was demonstrated with both the WSI Events and Impact scales \((r = .59–.79)\) and with a measure of general distress \((r = .24–.26)\). No significant relationships were reported between the WSI-SF and sex, race, education, employment, income, marital status, or number of medical conditions; however, like the WSI, WSI-SF scores tended to decrease as age increased \((r = −.16, p \le .05)\) (Brantley et al., 2007).

**IMPLEMENTATION ISSUES RELATED TO USE OF THE DSI AND WSI SCALES**

As noted previously, DSI and WSI scales were developed to serve as both outcome and predictor measures in investigations of etiological and mediating relationships between minor life events, stress-related physical and psychological disorders, and stress management interventions. The separate Events and Impact scores reflect both the stimulus and interactional theories of stress, such that the Event score frequency can be used to assess whether stressful events occurred, and the Impact score can be used to assess appraisal of the stressful events. However, in many of the studies reviewed, a causal relationship between DSI/WSI scores cannot be established; this would require that prospective designs be employed that can establish presence or absence of increased stress prior to the development of the physical or psychological symptoms of interest. Therefore, the DSI and WSI scales cannot be used as diagnostic instruments. Likewise, high stress scores do not suggest any particular etiology of a patient’s symptoms, and also no particular differential diagnosis, prognosis, suggested treatment, or response to treatment. Therefore, these measures cannot be expected to differentiate between various stress-related diagnostic entities, and the use of these scales in primary care must be supplemented with additional relevant information based on each patient’s unique presentation and history. The DSI and WSI scales are appropriate for research involving repeated-measures designs within the referent time frames of each scale, and multiple administrations of each measure can provide estimates of both the range and variations in chronic stress over time.

All versions of the DSI and WSI scales can be quickly and easily administered by nursing or ancillary healthcare staff in 15 minutes or less, have reading levels of seventh grade or less, and can be scored and interpreted by physicians or other treating healthcare providers. The differing time frames reflected by the measures suggest that the DSI be used to assess current, active stress symptoms on patients’ presenting problems, whereas the WSI may be more appropriate for assessing
stress symptoms and appraisal over a longer period of days or weeks. Finally, anecdotal reports of respondents participating in studies utilizing the DSI and WSI measures have suggested that regularly tracking their stressors may provide them with a better sense of how to predict and control their behavioral responses to stress, thus suggesting that the scales may have utility as tools for monitoring treatment outcomes (Brantley & Jeffries, 2000).

**SUMMARY**

The DSI and WSI scales are reliable and valid measures of the both the frequency and appraisal of minor stressful life events, and are appropriate as measures of stress-related physical and psychological symptoms, comorbidity, chronic stress, and interventions designed to specifically target stress in both primary and specialty care populations. While these measures cannot be used to diagnose stress-related psychological disorders, both scales provide clinically relevant scores on both the psychological and environmental aspects of stress, are easy to administer and interpret, and can be used as repeated measures of stress over time in both clinical and research applications. Assessment of stress can be particularly important in low socioeconomic status and primary care populations, where high prevalence rates of stress-related illness and symptomatology have been observed, and stress-related diseases are associated with both poorer health outcomes and higher healthcare utilization.

**REFERENCES**


Engel, B. T. (1985). Stress is a noun! No, a verb! No, an adjective! In M. Field, P. M. McCabe, & N. Schneiderman (Eds.), *Stress and coping* (pp. 3–12). Hillsdale, NJ: Lawrence Erlbaum.


Assessment of Stress


Primary care physicians (PCPs) are usually the first healthcare providers to have contact with a patient. Obviously, that means they have to have broad knowledge about a wide variety of clinical conditions, both those conditions for which the patient has awareness and those conditions for which the patient or family may not have awareness. Additionally, patients with cognitive impairment may think that those problems are superfluous to the presenting complaint. There are some stereotypic complaints related to a general practice. The patient or family may present to the outpatient appointment with a complaint of memory decline or difficulty with attention. It is unlikely that a presenting complaint would involve visual spatial construction deficits, even though that information would be useful in detecting signs of early dementia. Having an armamentarium of screening devices to uncover various forms of cognitive impairment would therefore be very useful to the psychologist working closely with the PCP. However, having only a screening instrument for memory is still helpful.

CONCEPTS IN SCREENING

Cognitive evaluations can be divided into two basic types: screening and standardized comprehensive. A screening device differs from standardized assessment in important ways. The screening device is typically shorter in time requirements. It requires less training to administer, score, and interpret. It typically requires less in the way of equipment. It tends to survey across likely areas of cognitive impairment rather than focusing in on a single area. For example, the Folstein Mini-Mental State Exam or MMSE (Folstein, 2001) can be learned in 20 minutes, whereas the Wechsler Memory Scale-IV (WMS-IV, Wechsler, 2009) requires advanced basic knowledge of psychometrics as well as 20–25 hours of specific training and practice in administration. The MMSE screens for orientation, attention, memory, construction, verbal comprehension, and naming, while the WMS-IV assesses memory in terms of new learning, visual and verbal auditory, modalities, immediate and delayed recall, and recognition. The MMSE can be administered using a single sheet of paper and a writing implement. The WMS-IV requires two wire-bound packets of test stimuli, a response booklet, a writing implement, and a timing device.

The use of a screening procedure will differ when it is used by the specialist, such as the clinical neuropsychologist, compared to when it is used by the generalist, such as a psychologist or other behavioral health clinician integrated into a primary care practice. The clinical neuropsychologist may use a screening procedure to make a decision regarding whether to follow up with a more detailed evaluation of an identified cognitive function. The screening instrument chosen by the clinical neuropsychologist may tap more than one cognitive domain, such as is done with the Mattis Dementia Rating Scale (Jurica et al., 2001), which assesses orientation, memory, construction, and abstraction. On the basis of the results the clinical neuropsychologist may decide to administer the WMS-IV in order to determine whether attentional difficulties are interfering
with memory or whether the rapid forgetting characteristic of the dementias is present. Alternatively, the psychologist may administer a briefer screening test such as the Clock Drawing Test and then decide whether to refer for a more comprehensive evaluation. Although separate scores for different cognitive functions are not available for this test, successful completion of the task requires visuospatial perception, eye–hand coordination and construction, planning, and remote memory. The psychologist is interested in whether cognitive impairment exists and if so, whether the patient should be referred to a specialist for a more detailed diagnostic evaluation. The types of questions addressed by the use of screening procedures can be characterized as being (1) does cognitive impairment exist (for the psychologist) or (2) what type and degree of cognitive impairment is present (for the specialist).

A word might be said here about to whom referrals should be made once the likely presence of cognitive impairment is identified. Disorders that cause cognitive impairment are complex and multifactorial. As such it is likely that more than one healthcare specialty will be involved in the management of the patient’s condition. The treatment of cognitive impairment may require referral to a neuropsychologist, psychiatrist, speech and language therapist, neurologist, or neurosurgeon. In all probability, referral will eventually be made to more than one professional.

The PCP is usually the first healthcare provider to receive complaints about cognitive impairment. When memory starts to fail or when there is an abrupt change in behavior, the patient (or family member) will come to the PCP for consultation. The PCP, by virtue of a continuing relationship with a patient, would know when something is different. For example, a regular patient may begin to forget to fill prescriptions or come to appointments. Alternately, a family member may call and say that the patient “is just not right.”

Following a public awareness program or the disclosure of the diagnosis in a public figure or celebrity, there may be an uptick in complaints of certain disorders such as attention deficit hyperactivity disorder (ADHD), Asperger’s syndrome, or dementia. The patient or family may have completed a brief checklist in a general interest magazine. Regardless of the initial event, the PCP is then left with the task of deciding whether to refer on to a specialist for a more detailed evaluation, to wait and monitor the condition over time to see if it worsens, or to reassure the patient that the suspected condition is actually unlikely.

As well as responding to questions from patients and family, the PCP may engender his or her own suspicions regarding the possibility of cognitive impairment. These may be on the basis of observation and interaction with the patient or on the basis of the PCP’s knowledge regarding the family history, medical diagnosis, or lifestyle behaviors of the patient. Nobody wants to make unnecessary referrals for lengthy and costly evaluations by specialists. The more prudent approach would be to refer the patient to the psychologist who is part of the primary care team for an initial screening. The use of screening tests, along with clinical observations, information from collateral sources, and a careful history can help the psychologist make informed decisions regarding whether the patient should be referred to a specialist for further assessment or treatment.

This chapter will provide a general overview of different screening procedures. With any of these instruments, it is important for the test administrator to be familiar with the procedures so that behavioral observations can be obtained as well as test scores. Therefore, the clinician would be well advised to practice test administration before use with a patient. Because of the busy schedule of the general practitioner and the ability of the cognitively impaired patient to mask and deny symptoms, it is likely that cognitive impairment, especially due to dementia, is under-recognized in this setting. Additionally, it may be helpful to have office clinical staff learn and practice these procedures as long as there is oversight and quality monitoring by the professional.
COGNITIVE IMPAIRMENT

For clinical purposes in this setting, cognitive impairment refers to acquired impairment. While the PCP will encounter cognitive impairment that is developmentally present in patients, the PCP is not called upon to identify its presence. Typically, the presence of cognitive impairment is already documented and the clinician is asked to provide health care to that patient. However, in some cases there can be acquired cognitive impairment, and this can be where the services of the primary care psychologist can be particularly helpful.

Cognitive impairment can be conceptualized in two ways. One way is related to comparisons to population averages. In this paradigm, when a person scores lower than other similar individuals, say in the fifth percentile of skill level for an individual the same age and educational background, then impairment is said to exist. This type of impairment is identified when comprehensive standardized testing is conducted. The other conceptualization of cognitive impairment has to do with the relationship of an individual’s cognitive skill to their capacity to perform a certain activity. This type of impairment may occur even though the individual still possesses a level of skill in the average range compared to other individuals. For example, an individual who owns and manages an apartment building may start making mistakes in the financial aspects of that endeavor. However, that person may still be able to manage their own personal affairs and perform routine maintenance of the building. In these cases, cognitive screening results need to be carefully weighed against the information obtained from the history and collateral sources.

The most likely etiologies for cognitive complaints brought to the attention of the psychologist are ADHD, concussion, dementia, psychiatric disorder, and substance abuse. The detection of dementia-related cognitive decline can be particularly vexatious in the early stages because of the similar effects of normal aging. This fact underlines the importance of using age-appropriate norms when interpreting test results.

AGE- AND PSYCHIATRIC-RELATED DECLINES IN COGNITIVE FUNCTIONING

It’s a fact that cognitive skills decline with age. However, not all skills decline uniformly at a consistent pace, and this decline does not occur in the same age range for different people. Entrepreneurs recognize this face and have developed a wide variety of products and devices purported to slow or reverse age-associated cognitive decline. With our cultural focus on cognitive skills as an indicator of health and youth, the psychologist is likely to be referred patients who come in to be seen because of their worry that their cognitive skills may be slipping. Of course for all of them at some point, cognitive skills will be declining. The question for the clinician is whether the decline is greater than would be expected on the basis of normal aging and whether a referral to a specialist should be made. For that reason, the screening device used should have appropriate age-relevant normative information. Evaluating an 80-year-old patient using an instrument normed on 30-year-old subjects will inevitably result in over-identification of cognitive impairment. Using an instrument with norms for people over the age of 70 to screen for impairment in a 55-year-old patient may result in missed diagnoses.

Although the decline in cognitive functioning with aging is an observation as old as aging itself, some recent work indicates that disease factors (e.g., cerebrovascular risk factors) may account for much of the cognitive change thought to be attendant upon aging (Morra, Zade, McGlinchey, &
Milberg, 2013). Regardless of the relative contributions of normal aging compared to the disease processes associated with aging, the older a patient is, the greater the likelihood that there has been some cognitive decline.

Recently the concept of mild cognitive impairment has been articulated. The American Academy of Neurology describes mild cognitive impairment (MCI) as including the presence of memory complaints by the patient and optimally by an informant as well, as well as objective memory impairment, in the presence of normal functioning in other cognitive areas and relatively intact capacity to perform intact activities of daily life (ADL; Petersen et al., 2001). It is probable that many of the patients presenting for screening will actually have MCI rather than frank dementia. However, screening still has value because in many cases MCI may eventually progress to dementia.

Some psychiatric disorders have associated risks for cognitive impairment. The most extreme effects are seen in the major thought disorders such as schizophrenia, but the most common disorders to be seen in primary care settings are depressive and anxiety disorders. For depression, the most frequently affected cognitive abilities are memory, attention, and to a lesser extent executive functions (Austin, Mitchell, & Goodwin, 2001; Lee, Hermens, Porter, & Redoblado-Hodge, 2012). Although the clinician may be inclined to give lesser import to these complaints because they are not strictly speaking neurological, the effects can be debilitating to the patient and can be significant barriers to adjustment.

Lifestyle variables such as substance abuse or occupational exposure can also be associated with cognitive difficulties. Many of these deficits such as memory impairment associated with alcohol abuse may resolve with abstinence, but there is evidence to suggest that deficits in executive functions and visuospatial skills may persist even after recovery of other functions following abstinence (Sullivan et al., 2000). Although the relation of other forms of substance to cognitive impairment is less well known than for alcohol abuse, it is generally accepted that substance abuse can cause significant cognitive impairment regardless of the substance used (Vik, Cellucci, Jarchow, & Hedt, 2004). Long-term solvent exposure can cause significant cognitive impairment (Sabbath et al., 2012).

For these reasons, it is important that the psychologist conduct a careful medical, family, and social history of the patient with cognitive complaints. Many medical conditions such as hypertension can be associated with cognitive deficits. Obtaining a work and avocation history can help determine whether the patient engaged in any activities (e.g., working around solvents or heavy metals, having hobbies that involve such exposure) that may have resulted in acquired cognitive impairment.

As part of the screening the psychologist can obtain information regarding family history. This information may already be available from the patient’s files, but it is a good idea to revisit this information in a conversation with the patient. Many disorders associated with aging, such as hypertension and Alzheimer’s disease, have a strong genetic component and subjective complaints of cognitive impairment may carry greater import than similar complaints in a person without a family history of those disorders. Furthermore, many psychiatric disorders such as depression have genetic components.

### REALMS OF INFORMATION IMPORTANT FOR SCREENING

The concept of screening requires that a short set of procedures be utilized and that a short amount of time be required. A comprehensive evaluation requires that all important areas of cognitive functioning be evaluated in depth. The requirement of a short set of procedures can be met by having only one or two items that tap a certain skill. The requirement of a short amount of time can be met by tapping only those skill areas implicated in the disorder being screened. Because memory is implicated in most cognitive disorders, memory is almost always included in the screening procedures. However,
it may not be necessary to assess all cognitive skill areas in a given screening procedure. For example, it may not be necessary to evaluate motor coordination when screening for a progressive dementia.

**Types of Psychometric Measures**

Although the process of assessing for cognitive impairment involves what have come to be known as screening instruments, screening should be considered a process with multiple sources of information and not just a single cognitive test. Even in the realm of tests, there are multiple types of information. There are three basic categories of information available to the clinician in screening for cognitive impairment. One are is that of *informant measures*, in which observations regarding the performance of everyday tasks requiring cognitive skill are obtained from individuals in the patient social environment. This type of information has the benefit of providing ecologically relevant information regarding the capacity of the patient to successfully engage in cognitive tasks. The second category of information is the one in which most of the instruments have been designed, namely, *standardized performance measures*. In these, the clinician poses certain tasks such as memory for a list of words or ability to perform mental arithmetic to the patient. Standardized instructions for administration of the items and scoring of the patient’s responses allow for objective comparison of the patient’s performance to normative groups of impaired and unimpaired individuals and enable the grading of the patient’s level of skill, such as average performance or below-average performance. The third category is that of *clinician ratings*, in which the clinician assigns scores based on observations or information derived from interviews. Although these instruments can be helpful in structuring and increasing the objectivity of the clinician judgments, they also require that the clinician obtain training in the method in order to help ensure that the resulting judgments would be in agreement with judgments made by another clinician using the same instrument.

**Domains of Cognitive Functioning**

*Attention* is an early component of the process of taking information from the environment. Attention may become impaired in terms of the capacity to sustain attention, the amount of information to which one can attend, or the ability to withstand distraction. Attention may become impaired after a concussion or a hypoxic event such as a myocardial infarction. Attention can also be impaired as part of a neurodevelopmental disorder such as ADHD. The psychologist may be called upon to make a decision whether to refer for a more detailed evaluation of ADHD. Although ADHD is not a cognitive impairment disorder per se (it is a neurobehavioral disorder with cognitive correlates), there are screening instruments available to aid the clinician. The World Health Organization has published a short self-report questionnaire to aid in the identification of ADHD in adults (Kessler et al., 2005). There are 18 items in the main version, and a short version with six items can also be used. Attention impairment may also be secondary to psychiatric difficulties such as depression or anxiety.

It is important that the psychologist be sensitive to emotional difficulties when patients come in with complaints about cognitive difficulties such as attention or memory deficits. Emotional difficulties such as depression or anxiety may either be the cause of cognitive difficulties or may be the result of the patient coming to awareness of the development of cognitive difficulties. The psychologist should also be sensitive to the fact that most people will conceptualize their difficulties as memory problems. Whether a memory deficit exists or a skills deficit is construed as forgetting how to perform a task begins with careful questioning in the interview leading to the careful assessment of individual skills by the specialist.
**Memory** is by far the most common complaint of individuals with some form of subjective cognitive complaint. Memory is a complex skill that requires initial encoding of the material, learning, immediate recall, and delayed recall. Most patients will describe memory impairment in the form of rapid forgetting, usually for episodic memory, or memory for things that they experience or observe. Most tests of memory evaluate semantic memory or memory for things the person purposefully tries to learn and recall later. Therefore, there may not be agreement between the subjective complaints and the objective test results. Instead, attentional lapses may be interfering with the initial stages of memory processing or the patient may conceptualize a loss of skill, such as aphasic symptoms like losing memory of language. Alternately, a patient’s depression might be coloring their perception of the magnitude of any lapses in memory. It is up to the clinician to parse these distinctions by carefully listening to the patient and the family. Because memory is a ubiquitous complaint, all of the screening instruments discussed later in the chapter have some form of memory assessment as part of the set of procedures. The clinician must also use careful questioning to determine whether a memory deficit does exist. If the complaint involves a lessened ability to perform complex activities the clinician should also consider whether one of the component skills is instead deficient.

**Language** functions involve both receptive and expressive functions. The family of an individual with receptive language deficits may describe the problems as being with hearing acuity. The comprehensive evaluation of language skills requires the expertise of a speech and language clinician or a clinical neuropsychologist, but the primary care psychologist may perform a screen by asking the patient to follow simple commands (receptive) or to name objects shown to the patient (expressive).

**Motor skills** can be assessed by observing the patient’s station and gait. Both speed and accuracy are relevant here. Alternately, there may be a decline in the executive skills. Executive skills includes planning, problem-solving, evaluation of information from the environment, judgment and prediction of likely future events and consequences—in short, all the skills that must be practiced by a good executive of an organization.

**Detection of Suboptimal Test Performance**

The clinician also needs to be attuned to the possibility that some individuals may not be motivated to perform at their optimal level during the administration of cognitive tests. An individual may wish to present with magnified cognitive impairment for multiple reasons. There may be a desire to demonstrate cognitive impairment in order to acquire disability benefits or to evade responsibility for criminal actions. In addition to conscious fabrication of cognitive impairment, patients may not perform at their optimal level for reasons of depression or an effort to call attention to problems that they feel are not getting sufficient attention from healthcare providers or from family members. Clinical neuropsychologists are well aware of this issue and have devoted much time and effort to develop methods to evaluate effort and motivation in cognitive testing. Screening instruments are brief and do not contain validity indices or methods of detecting less than optimal performance. The methods and test procedures developed by researchers and utilized by specialists are not generally in the armamentarium of primary care psychologists. However, there are methods and knowledge that can assist them in making an initial assessment of the likelihood of symptom magnification.

One of the first sets of information that can be helpful in detecting less than optimal effort is information regarding the context of the evaluation. The PCP usually has the advantage of a historical relationship with the patient. Does the patient have a history of multiple subjective complaints that have not been substantiated by objective tests? Are there signs in the history, such as an accumulation of life stressors, which might point to the possible presence of a depressive condition?
Is there evidence that the patient may wish to use a determination of cognitive impairment or to achieve some external gain? Is the degree of subjective cognitive complaint consistent with what is usually found in connection with aspects of the objective history? For example, a patient facing legal difficulties may attempt to avoid responsibility by feigning cognitive impairment. An individual desiring special accommodations in an academic setting or at work may attempt to claim cognitive limitations. As another example, a patient who was involved in a minor motor vehicle collision may describe severe memory impairment or may describe unlikely patterns of impairment by stating that they cannot remember the names of family members.

Another set of information available to the clinician is behavioral observation of the patient during the interview or during administration of the cognitive screening instrument. Is the patient demonstrating exaggerated effort or are they calling attention to their poor performance? The clinician can also enlist collateral information from people in the social environment of the patient. Does a family member provide evidence of consequences of the presumed cognitive impairment in the patient’s natural environment? Can they give examples of times when the patient failed to perform a task adequately because of the cognitive impairment? Alternately, can the family member give examples of intact performance that should not be possible if the impairment were present?

The previous discussion is not meant to suggest that a determination of malingering or suboptimal effort can be accomplished using a brief cognitive screening instrument. However, the primary care psychologist can use that information in making the decision to refer to a specialist or may express their suspicions in the context of the referral.

**SCREENING MEASURES FOR SPECIFIC DISORDERS**

Many instruments have been developed to assist in screening specific types of disorders, many of which present themselves in primary care settings. Three of these types of disorders and related screening measures are discussed in this section. Before going on to consider larger amount of information about screening for dementia, available measures and procedures for the assessment of two other common sources of cognitive impairment—ADHD and concussion—will be considered.

**Attention Deficit Hyperactivity Disorder (ADHD)**

Although ADHD is a developmental disorder—that is, a disorder that is first seen in childhood—adults, especially older adults, may have been previously undiagnosed. This is more likely to be the case if the patient is middle-aged or older and grew up in a time when there was less awareness of the signs of the disorder or the patient had sufficient basic cognitive skill to adjust to and ameliorate the effects of the disorder. These patients may come to the clinician with complaints that they feel they never lived up to their potential, or that they have been unable to hold a job or maintain a long relationship. The patient may state that they don’t finish tasks or that they tend to procrastinate excessively. ADHD is actually a neurobehavioral disorder and not specifically a neurocognitive disorder, even though the disorder has associated cognitive difficulties such as easy distractibility (impaired attention) or impaired planning (executive dysfunction). A detailed diagnostic evaluation will be needed to parse out the presence of current complaints from a history of similar difficulties. However, the psychologist can screen for the presence of current complaints using a few available instruments.

One of the most commonly used questionnaires is the Wender Utah Rating Scale (WURS). Ward (1993) describe the development of this instrument and provide an appendix with the questions.
Cutoff scores are available to classify patients as being likely to have the disorder as adults when compared to adults without the disorder, and compared to adults who have an affective disorder, such as depression, but do not have ADHD. This can be the case particularly because some of the symptoms of ADHD overlap with the symptoms of depression in adults. The WURS is a retrospective method that asks the patient to state whether certain symptoms were present in childhood. Daigre et al. (2015) used the WURS in a sample of adults and validated it against a more comprehensive diagnostic tool, the Conners diagnostic interview (Epstein & Johnson, 2001).

There are also useful methods based on current symptoms of ADHD in adults. One such instrument can be found in a questionnaire developed as part of a World Health Organization project. Kessler et al. (2005) describe a questionnaire that consists of 18 items tied to DSM-IV symptoms. Further refinement resulted in shortening it to a six-item screener (Kessler et al., 2007) without sacrificing the accuracy. Another instrument is described by Semeijn et al. (2013). They call their questionnaire the ADHD Screening List and also validated it against a diagnostic interview, in a sample of adults living in the Netherlands. Hines, King, and Curry (2012) describe a short, six-item instrument, the Adult Self Report ADHD Scale, which was validated in a primary care setting. It has a positive predictive value of .52 and a negative predictive value of 1.0, and may therefore be more useful to rule out ADHD rather than to identify it. They report that it takes less than one minute to administer, and that brevity certainly enhances its value in a busy clinic.

**Concussion**

Concussion results when physical energy is transduced to the central nervous system via an impact or acceleration/deceleration. (Another form of concussion, the blast concussion, has been recently identified, but this is usually applicable to the military population.) There may be immediate effects such as loss of consciousness or vomiting, or there may be delayed effects such as severe headache. Many of the effects of concussion will remit in the immediate period following the concussive injury. The cognitive effects can be subtle and do lend themselves to discovery by brief screening instruments. However, many of the effects, such as dizziness, sleep impairment, or blurred vision are subjective in nature. The Post Concussive Symptom Scale (PCCS; Macciocchi, Barth, & Littlefield, 1998) is a set of 22 items regarding the presence of concussion symptoms in the week prior to completing the instrument. The patient rates the symptoms on a 7-point scale from Absent (0) to Severe (6). The values for each item are summed for a total score. The PCCS probably has greater value for measuring change in a single person across time, and there are some recommendations for assigning level of severity based on the total score (Lovell et al. 2006). Chen, Johnston, Collie, McCroy, and Ptito (2007) found that the PCSS has good agreement with the results of both standardized neuropsychological testing and fMRI following concussion in athletes.

**Dementia and General Cognitive Impairment**

There are several tests and procedures available for screening for dementia and general cognitive impairment. The following list of screening tests is heavily weighted toward detecting cognitive impairment associated with dementia. First, dementia is probably the most likely cognitive condition to be faced by the PCP. Second, for related reasons, much of the professional literature addresses this condition rather than the detection of concussion or ADHD. Because cognitive impairment is more likely in older patients, most research regarding these instruments has been conducted with older patients. The clinician should keep this in mind when choosing an instrument to screen younger
adult patients. Other limitations include the fact that some etiologies of cognitive impairment such as concussion may have effects that are too subtle to be detected by a screening instrument. For example, the MMSE may not be sensitive to the type and degree of cognitive impairment seen with heart failure; instead, the Clock Drawing Test or Montreal Cognitive Assessment (MoCA) may be more suitable (Davis & Allen, 2013).

This list is not meant to be exhaustive. These are only some of the available and more frequently used instruments available. In each case representative publications are cited. When possible, meta-analytic or review articles are cited. Most of these instruments are available from their authors, are described in the citations, or are available online. Some, like the MMSE and NCSE, are available commercially.

**Folstein Mini-Mental Status Exam**

The Folstein Mini-Mental Status Exam (MMSE; Folstein, 2001) is the granddaddy of all brief cognitive impairment screening devices. It assesses orientation, registration, attention and calculation, construction, memory, and language, and takes between 10 and 15 minutes to administer. First published in the *Journal of Psychiatric Research* in 1975, it has gone on to become the most widely used bedside cognitive screening test for physicians (Crum, Anthony, Bassett, & Folstein, 1993). It has many critics, but many more users. It has been revised and now is available as a commercial product. It also has been translated into nearly 50 languages and is available in an electronic form. Mitchell (2009) concluded that it was more useful in ruling out dementia (specificity) than in identifying dementia or MCI (sensitivity). In a review of 22 screening tests, Velayudhan et al. (2014) report that although all of the tests have comparable specificity in identifying dementia, the MMSE had the least sensitivity, especially in identifying mild cognitive impairment. Therefore positive findings using the MMSE can be interpreted with greater confidence than negative findings.

**St. Louis Mental Status Examination**

After the MMSE was copyrighted and published in the late 1990s, the authors and publishers sought legal support of the copyright against the many individuals who were accustomed to using the free version available in the research literature. This in turn led to other clinical researchers developing alternate instruments such as the St. Louis University Mental Status Examination (SLUMS; Holsinger, Deveau, Boustan, and Williams, 2007). Additionally, the SLUMS was developed partly in response to validity concerns regarding the Folstein MMSE. Because of the simplicity of the items, the MMSE has ceiling effects, that is, individuals with higher levels of premorbid functioning or educational attainment may develop decrements in skill and still perform on some tests in the intact range.

One of the more frequent concerns brought to the attention of clinicians who treat older patients is that of mild cognitive impairment. Patients may report lapses in attention or memory that concern them, even though they may be relatively mild. The SLUMS contains items of greater difficulty than the MMSE and may therefore be more sensitive to subtle cognitive impairment. Feliciano et al. (2013) found that the SLUMS correlated significantly with standard neuropsychological tests of cognitive function. The SLUMS also correlated with the MMSE, and multiple regression models indicated that the SLUMS correlated with these tests even when the effects of demographic variables and the MMSE scores were statistically removed. Unfortunately, the SLUMS has been mainly developed and researched in populations of veterans and may not have equivalent utility in civilian populations, although the Feliciano study did use a nonveteran population.

Howland, Tatsuoka, Smyth, and Sajatovic (2016) report that the SLUMS may be less sensitive than the MMSE to cognitive changes over a 1-year period. Cruz-Oliver, Malmstrom, Roegner, Tumosa, and Grossberg (2014) used a retrospective chart review to demonstrate that when the cognitive impairment
had reversible causes that were adequately treated, scores on the SLUMS showed improvement. Cummings-Vaughn et al. (2014) report that the SLUMS, MoCA, and the Short Test of Mental Status (STMS; Kokmen, Smith, Petersen, Tangalos, & Ivnik, 1991) are equivalent in the detection of mild cognitive impairment and dementia.

NCSE Neurobehavioral Cognitive Status Examination
The Neurobehavioral Cognitive Status Exam (NCSE; Kiernan, Mueller, Langston, & Van Dyke, 1987), also known as Cognistat, is somewhat longer than most screening instruments and additionally requires a small amount of equipment (colored plastic squares, etc.). In addition to elderly subjects, the NCSE has been used with head-injured patients and stroke patients and has normative information for adolescents, adults, and elderly patients. It contains items to evaluate orientation, memory, calculation, language, and reasoning and takes about 25 minutes to administer. Kazmark (1997) reported generally adequate sensitivity and sensitivity when the criteria were the results of standardized neuropsychological tests. Eisenstein et al. (2002) provided data on scores from the NCSE in healthy elderly subjects, which allows judgments regarding cognitive functioning in community-dwelling individuals. The NCSE has been found to be equivalent to MoCA and Short Test of Mental Status (Cummings-Vaughn et al., 2014).

Clock Drawing
The Clock Drawing Test (Mainland & Shulman, 2013) is simple and requires no special equipment. It requires simply that the examiner ask the patient to draw a clock face and set the hands to a specific time. The task itself is simple, but it requires a combination of complex cognitive operations including construction and executive skills such as planning. Because those skills are affected in Alzheimer’s disease, the Clock Drawing Test may be useful in uncovering dementia (Amodeo et al., 2015; Shulman, 2000). However, it may not be sensitive to other forms of cognitive disorders such as may be seen in concussion. For example, Adamis, Meagher, O’Neill, and McCarthy (2015) found that the Clock Drawing Test was relatively insensitive to the cognitive impairment found in delirium.

Three Word Recall
The Three Word Recall Test (3WR; Chandler et al., 2004) asks the subject to recall three words after a brief delay. It has been incorporated into other procedures such as the MMSE, but it is possible to administer as a stand-alone procedure. Chandler et al. (2004) found that although performance on the Three Word Recall test decreases with normal aging, these aging effects are less than that found in disease processes. The discrimination requires age-related norms that are reported by those same investigators.

7-Minute Screen
The 7-Minute Screen (7MS) includes assessment of temporal orientation, construction, verbal fluency, and memory. Meulen et al. (2004) report the 7MS is sensitive to a variety of dementing conditions. Unfortunately, it is also sensitive to the cognitive impairment found in depression and may not be helpful in assessing those patients.

AB Cognitive Screen
The items of the AB Cognitive Screen (ABCS) (Solomon & Pendlebury, 1998) assess orientation, initial memory registration, clock drawing, delayed recall, and word fluency (Molloy, Standish, & Lewis, 2005). Standish, Molle, Cunje, and Lewis (2007) report that the ABCS can discriminate between dementia and mild cognitive impairment or normal cognition. The ABCS has increased sensitivity
over the MMSE in differentiating mild cognitive impairment from normal cognition. With minimal ceiling effect, it is also less influenced by age and education.

**Addenbrooke’s Cognitive Examination (ACE-R)**
Dudas, Berriors, and Hodges (2005) found that the Addenbrooke’s Cognitive Examination (ACE-R) can discriminate between dementia and affective disorder, giving it utility when the clinician wishes to perform an initial screen of possible cognitive impairment when depression is present and may present reversible cognitive impairment. Reyes et al. (2009) state that the ACE-R can uncover cognitive impairment in patients with Parkinson’s disease. Kaszás et al. (2012) report that the ACE-R is less accurate than the Mattis Dementia Rating Scale. However, the relative brevity of the ACE-R compared to the Mattis may bolster its value in a busy primary care practice.

**Alzheimer’s Disease 8**
Galvin et al.’s (2005) Alzheimer’s Disease 8 (AD8) contains items about the participant’s cognitive abilities in the areas of memory (consistent problems with memory, repetition, remembering appointments), temporal orientation, judgment (making decisions, handling finances), and function (reduced interest in activities, use of appliances). Research indicates it provides reasonable discrimination of demented from nondemented subjects. It is sensitive to early signs of a progressive dementia and it correlates both with objective measures of cognitive function and with clinicians’ rating of dementia symptoms (the Clinical Dementia Rating). Galvin et al. (2006) reports adequate reliability and reasonable external validity (agreement with standardized tests) in a sample of 255 patients and informants. Moreover, Galvin et al. (2007) report that the AD8 can also be used as a self-report instrument with good agreement with informant ratings and clinician ratings.

**Abbreviated Mental Test**
In reviewing the literature, Lees et al. (2014) found that the Abbreviated Mental Test (AMT; Jitapunkul, Pillay, & Ebrahim, 1991) was no less sensitive to cognitive impairment than were other, lengthier procedures. Pendlebury, Klaus, Mather, de Brito, and Wharton (2015) report that the AMT is as accurate as the MoCA in identifying possible cognitive impairment.

**Brief Alzheimer Screen**
The Brief Alzheimer Screen (BAS; Schmitt, Mendiondo, Kryscio, & Ashford, 2006) is just that—a brief instrument aimed at detecting the cognitive impairment associated with Alzheimer’s disease. It may have limited utility in the initial assessment of other causes of cognitive impairment. Its component procedures tap those areas known to be affected in Alzheimer’s, namely memory, orientation, verbal fluency, and mental concentration.

**Cognitive Abilities Screening Instrument**
The Cognitive Abilities Screening Instrument (CASI) has items that assess attention, orientation, mental concentration, verbal fluency, short- and long-term memory, abstraction judgment, and language (Teng et al., 1994). One benefit of the CASI for practitioners in a multicultural practice setting is that the CASI has been validated for use with Chinese-speaking subjects (Lin et al., 2002).

**Cognitive Assessment Screening Test**
The Cognitive Assessment Screening Test (CAST; Swearer et al., 2002) is a brief paper-and-pencil test that can be administered and scored by a nonspecialist. It has adequate test-retest reliability and has reasonable sensitivity and specificity when compared to other screening tests (Swearer et al., 2002).
It is self-administered and is three pages long. The first page has skill items such as simple arithmetic; the second page has somewhat more difficult items such as verbal abstraction questions; and the third page has items regarding the patient’s subjective experience and awareness of cognitive decline.

**DemTect**
The DemTect (Kalbe et al., 2004) has five procedures: a word list learning task, a number transcoding task, a word fluency task, digit span reverse, and delayed recall of the word list. It takes 8–10 minutes to administer and is relatively free of the effects of age and education (Kalbe et al., 2004).

**General Practitioner Assessment of Cognition**
The General Practitioner Assessment of Cognition (GPCOG; Brodaty, Kemp, & Low, 2004) combines assessment of cognitive performance (which takes about four minutes to administer) with a brief interview of a collateral source such as a relative (which takes about two minutes to administer). Brodaty et al. (2002) report adequate sensitivity and specificity on identifying dementia. Further research indicated lesser confounding influences of age and education (Brodaty et al., 2004). Based on its ease and brevity of administration as well as its sensitivity and specificity in identifying cognitive impairment, Brodaty et al. (2006) recommends its use in a general practice setting.

**Informant Questionnaire on Cognitive Decline in the Elderly**
The Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) is a set of 26 items administered to someone who knows the patient well. It has been translated into multiple languages and can be downloaded at http://crahw.anu.edu.au/risk-assessment-tools/informant-questionnaire-cognitive-decline-elderly (Jorm, 2004). It correlates well with other screening measures such as the MMSE and the Abbreviated Mental Test. The IQCODE has small correlations with age and education and only modest correlations with affective variables such as depression or anxiety. There is a short form available with 16 items.

**Short and Sweet Screening Instrument**
The Short and Sweet Screening Instrument (SASSI) involves three short tests of cognitive function derived from a larger battery (Belle et al. 2000). It can be administered in 10 minutes and is reasonably accurate in detecting general cognitive impairment. Unfortunately, it does not sample memory and may have limited utility in those instances where memory impairment is the chief complaint. Furthermore, there has not been an empirical evaluation of the effects of education or demographic variables.

**Six-Item Screener**
The Six-Item Screener (SIS; Callahan et al., 2002) is shorter even than most screening instruments. It consists of three orientation items and a three-word recall procedure. Because of its brevity and the fact that it can be administered without any equipment or scoring sheets, it has proven to be attractive in the emergency department (ED). It also lends itself easily to a general outpatient clinic or office.

The SIS shows greater agreement with the MMSE than does the Mini-Cog (Wilber, Lofgren, Mager, Blanda, & Gerson, 2005), which should not be surprising because the items of the SIS are drawn from the MMSE. The SIS does not include motor items nor does it require visual stimuli; therefore, it may be adapted for use over the telephone. The SIS shows good agreement with the MMSE in a community-dwelling population, even though it is much shorter (Callahan et al. 2002). The SIS is sensitive to changes in cognitive status due to cerebrovascular events (Unverzagt et al., 2011).
Although earlier studies showed promise, Wilber et al. (2008) found lower sensitivity of the SIS in a sample of elderly emergency department (ED) patients when the MMSE was used as the criterion for identifying cognitive impairment. Of course, with any short screening procedure, impairment in the cognitive domains that are not tapped by the instrument will go undetected. For that reason, Hirschman et al. (2011) recommend supplementing the SIS with the Clock Drawing Test in order to screen for executive impairment.

Six-Item Cognitive Impairment Test (6CIT)
The Six-Item Cognitive Impairment Test (6CIT; Jefferies & Gale, 2013) not to be confused with the SIS) is a very brief test of one memory item, two attention items, and three orientation items. Essentially, it is a shortened form of the 26-item Blessed Information-Memory-Concentration Scale. It has been found to be useful particularly in a busy outpatient practice or ED. However, its brevity also means that not all relevant areas of cognitive impairment are screened by it and the scoring can be cumbersome. Tuijl et al. (2012) report that the 6CIT is accurate in agreeing with the MMSE, takes less than five minutes to administer, and is not sensitive to the effects of education. It was found to have adequate accuracy in screening for cognitive impairment in acutely ill older patients (Goring et al., 2004).

Short Memory Questionnaire
The Short Memory Questionnaire (SMQ) is a brief, informant-based questionnaire. It asks the informant to answer questions regarding the presence of memory impairment in older patients. Koss, Patterson, Ownby, Stuckey, and Whitehouse (1993) report substantial correlations with objective measures of cognition as well good internal reliability and excellent negative predictive power.

Abbreviated Mental Test (AMT)
The Abbreviated Mental Test (AMT; Jitapunkul et al. 1991) is a 10-item instrument that assesses short- and long-term memory, attention, and orientation. Jackson, Naqvi, and Sheehan (2013) used meta-analytic techniques on nine studies and determined that the AMT had acceptable sensitivity and specificity. However, the emphasis on memory may limit the utility of the AMT when memory is intact but other cognitive impairment exists. Additionally, the specificity and sensitivity of the AMT may be less than that of the MMSE (Woodford & George, 2007).

Mini-Cog
Borson, Scanlan, Brush, Vitaliano, and Dokmak (2000) described the development of a brief instrument that combines a three-item verbal memory task with a clock drawing procedure. In a later population-based sample, the Mini-Cog was found to have sensitivity and specificity equivalent to the MMSE (Borson, 2003). As an added advantage, Borson et al. (2001) report that the Mini-Cog can be reliably administered by relatively naïve raters.

Rapid Dementia Screening Test
The Rapid Dementia Screening Test (RDST; Kalbe, Calabrese, Schwalen, & Kessler, 2003) has only two procedures—a verbal fluency task and a number-symbol transformation task. It is very brief and takes only three minutes to administer. Initial data indicates reasonable sensitivity to the presence of dementia.

Short Test of Mental Status
The Short Test of Mental Status (Kokmen et al., 1991) takes approximately five minutes to administer. It is similar to the MMSE in that it has items to assess orientation, attention, immediate recall,
3-minute delayed recall, and construction. Additionally, it has items that assess arithmetic, abstraction, and general information. Data regarding its validity are somewhat limited.

Montreal Cognitive Assessment
The Montreal Cognitive Assessment (MoCA) is the result of a series of research studies to develop a quick but accurate test of cognitive impairment to aid in the diagnosis of dementia, the detection of MCI, and the identification of nonimpaired patients (Nasreddine et al., 2005). It is available for free download after the clinician registers online. The website has downloads available in three alternate forms (to avoid practice effects in multiple administrations) as well as versions in 46 different languages. Additionally, the webpage (http://www.mocatest.org.) has a growing list of research citations, as well as opportunities for collaboration and online training culminating in certification. It assesses orientation, memory, attention, language functions, and construction. The list of references is organized by type of disorder or clinical question to be answered and is regularly updated. Velayudhan et al. (2014) report that in reviewing 22 different screening tests, all of them were generally able to distinguish between unimpaired and dementia subjects, with the MoCA being superior to all of the others in its capacity to identify mild cognitive impairment. Larner (2014) reviewed the literature on five different screening instruments, the MMSE, 6CIT, the Mini-Mental Parkinson, the Test Your Memory test, and ACE-R. He calculated the effect of sizes using Cohen’s $d$, a measure of the standardized difference between groups (in this case, between subjects with and without cognitive impairment). Larner reported that although five tests had equivalent effect sizes when comparing patient with dementia to intact individuals, the MoCA had the largest effect size when comparing subjects with MCI to intact subjects.

SUGGESTIONS FOR APPLYING THE CONCEPTS
As a result of their basic clinical training, the primary care psychologist already has the basic information needed to consider the possibility of cognitive impairment. However, information, like dairy products, has an expiration date. The neurosciences in particular have a rapid rate of change in the recent past. It is important for the psychologist to keep abreast of changes in information and procedures. This does not mean it is necessary to subscribe to a specialty professional journal. Most of that information is likely to be esoteric and not useful in a primary care setting. However, regional and local psychological associations often have education offerings related to these topics. Although it might not be the first choice when deciding among the offerings, registering for one of these on an occasional basis can help keep oneself fairly current.

Additionally, it can be helpful to reach out to the professional resources in the community. The primary care psychologist probably has specialists to whom suspected cases are referred for a more comprehensive evaluation of possible cognitive impairment. A telephone call or brief office consultation can provide valuable information. It can be especially important to develop a reiterative feedback pattern so that preliminary information from the specialist may aid in the decision to refer and the results of prior referrals can be used to judge the appropriateness of a future referral.

Having knowledge and skill in administration of the screening instrument is not sufficient. The practitioner should also have this procedure embedded in a system of knowledge and actions. The first stage is to have knowledge regarding when it would be useful to administer the screening procedure. The second step is to acquire the information, necessary materials, and behavioral skill necessary to use the screening instrument. The third step is to have a plan of action regarding what to do with the results: Refer for specialized assessment? Counsel the patient to wait and observe?
After reviewing the large body related to screening measure for dementia, Cordell et al. (2013) concluded that they could not offer a recommendation for a single best instrument. Detection of cognitive impairment is unlikely to be completed in a single visit or on the basis of a single set of information. Instead, it is a reiterative process using multiple sources of information. Observation by the physician during an interview, while helpful, is not sufficient, especially without objective cognitive performance measures. It is necessary to use some standardized screening measure to provide a baseline against which potential future changes can be compared or to inform a decision to refer.

There are two very important but sometimes overlooked sources of information. The first is the clinical staff who may spend extended time observing the patient or interacting with the patient either in person (e.g., in the waiting room) or via telephone. The clinical staff see the patient under more relaxed conditions than does the primary care physician or psychologist, and these staff see the patient under conditions in which memory or problem-solving skills may be required (e.g., making appointments, navigating insurance forms).

Finally, the use of screening procedures should be contained in a context sensitive to the emotional effect of a possible diagnosis. Brief counseling sessions before and after use of the screening procedure can help provide for acceptance of the results without emotional distress.

Everyone seems to be developing their own brief screening measures, resulting in a proliferation of available methods. However, for many of these instruments, there is little validation beyond the initial studies being done by investigators other than the developers. Perhaps more effort can be expended in researching the relative advantages and disadvantages of the different instruments in the different populations. It is unlikely that there will be one single instrument that is superior for all diagnostic populations and age groups in the primary care setting. This is true partly because by their nature, screening instruments are brief and therefore must leave out some cognitive skills while sampling from others. It is also partly true because with age-related changes in level of skill for different cognitive functions, what might constitute failure for a younger person might constitute acceptable performance for an older subject.

**A PROPOSED PROCESS IN SCREENING FOR COGNITIVE IMPAIRMENT**

Screening for cognitive impairment in a general outpatient practice can be thought of as a process with reiterative steps and decision points along the way. Table 14.1 gives a brief outline of the steps proposed.

The first step is to consider the reasons why the screening process might begin. The clinician should consider the context of the patient and the information that might lead to a decision to initiate the process. Such information can include observations of the clinician. The age of patient can be

<table>
<thead>
<tr>
<th>Table 14.1 Steps in Cognitive Screening Decisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Carefully consider the information related to the initial complaint.</td>
</tr>
<tr>
<td>2. Review the personal medical and social history of the patient.</td>
</tr>
<tr>
<td>3. Review the family history.</td>
</tr>
<tr>
<td>4. Interview the patient.</td>
</tr>
<tr>
<td>5. Obtain observations from clinical staff.</td>
</tr>
<tr>
<td>6. Choose a screening instrument.</td>
</tr>
<tr>
<td>7. Review results with the patient/family.</td>
</tr>
<tr>
<td>8. Make a referral if necessary.</td>
</tr>
</tbody>
</table>
considered first. The older the patient, the more likely it is that some form of cognitive decline has occurred. The clinician can consider the background of the patient, including the family history and medical history. Having a family history of individuals who have themselves developed cognitive impairment increases the risk that such impairment might occur in the patient. Other factors such as having a history of recurrent depressive episodes or a history of chronic diseases such as diabetes or hypertension may also serve to raise the suspicion of the clinician.

Second, observations of the clinician or of the office staff may raise suspicion. The staff might report that the patient seems to have trouble understanding the registration process of the office or how to satisfy insurance requirements. Additionally, the occurrence of multiple phone calls asking for the same information or making multiple appointments may indicate memory deficits. The patient might have missed appointments or come on the wrong day or time for an appointment. During a routine visit, a patient may seem to have trouble remembering particulars of prior visits or familiar prescribed medication regimens.

Next, the clinician may seek information from the patient or a family member. There may have been recent indications of memory problems or changes in the capacity of the patient to perform typical activities. There may also be reports of changes in personality. A formerly easygoing patient may become irritable. The patient may lose interest in formerly enjoyed activities or become less social. The patient or the family member may report instances of the patient becoming lost or confused when driving a familiar route. It is important to remember that some of these changes might also signal emotional conditions such as depression. Therefore, a careful evaluation of the emotional state of the patient is important here as well.

The clinical interview can serve two functions. The clinician may use the interview to obtain more information about the possible presence of cognitive changes. But if screening is to proceed, the interview can also be used to assess the patient’s willingness to receive the screening test and to assess the possible reaction of the patient to either set of results. The potential outcomes of the screening test should be discussed with the patient and the family, including the possible benefit identifying a potentially reversible cause for any cognitive changes.

Next, the clinician needs to decide on a screening instrument and a method of administration of the test. Will the clinician administer the test personally or will office clinical staff be used for that activity? In either case, a quiet, calm, private environment will be necessary. Time should be allotted for the patient to debrief after administration of the test. As this chapter has described, there are multiple screening tests available. The choice of one will depend upon multiple factors including the length of the test, the ease of administration, the presence of an appropriate normative base of research to guide decisions, and the prior use of the test in identifying whatever disorder is suspected in the patient. Table 14.2 is an outline of the variables involved in choosing a screening test.

The generalist clinician may feel unprepared to make such decisions alone, and here is where having a relationship with a specialist can be helpful. Not only can the specialist be a venue for referring patients for further evaluation, but the specialist can also be used as a consultant in choosing a screening instrument. The specialist, such as a clinical neuropsychologist, is more likely to be knowledgeable about which instruments may be helpful in a particular instance.

<table>
<thead>
<tr>
<th>Table 14.2 Questions to Ask When Choosing a Screening Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is it sensitive to the suspected area of cognitive impairment?</td>
</tr>
<tr>
<td>Has it been tested in the diagnostic group of interest in this case?</td>
</tr>
<tr>
<td>Are there appropriate age and education norms?</td>
</tr>
<tr>
<td>Can it be used reliably with minimal training?</td>
</tr>
</tbody>
</table>
Developing a relationship with a specialist is also integral to the next step in the screening process. The generalist clinician will need to set up a process for feedback on the appropriateness of the referrals that have been made. By having regular brief communication with the specialist, the generalist can determine if too high or too low a bar has been set in deciding to refer on. The generalist and the specialist together can discuss what resources are available in the community for the management of whatever disorder or condition is uncovered in the patient by the screening process. This partnership is the final step in the screening process and ultimately determines the utility of the entire process.

SUMMARY

As our population ages, and as we become more aware of the possibility of disease caused cognitive impairment, it is becoming increasingly important to screen for cognitive impairment in a larger number of patients. The psychologist can play an important role in this aspect of health care in primary care settings as well as other settings. Screening should be thought of not just as the administration of short tests of cognitive function but also as a set of processes involving the collection of multiple types of information from the patient and the family. The results of screening can be used to help guide the possible referral to specialists and to ultimately obtain optimal management and care of the cognitively compromised patient.

REFERENCES


of cognitive impairment during the Medicare annual wellness visit in a primary care setting. *Alzheimer's & Dementia*, 9(2), 141–150.


Cruz-Oliver, D. M., Malmstrom, T. K., Roegner, M., Tumosa, N., & Grossberg, G. T. (2014). Cognitive deficit reversal as shown by changes in the Veterans Affairs Saint Louis University Mental Status (SLUMS) Examination scores 7.5 years later. *Journal of the American Medical Directors Association*, 15(9), 687–e5.


Screening for Cognitive Impairment


Substance use disorders (SUDs) and their comorbid conditions such as liver failure, heart disease, and renal failure continue to be some of the most significant health problems within the United States. According to the results from the 2014 National Survey on Drug Use and Health, among those 12 and older, approximately 21.5 million, or 8.1%, of people in the United States had a substance use disorder (“Behavioral Health Trends in the United States: Results From the 2014 National Survey on Drug Use and Health,” 2015). Consequently, detecting the signs and symptoms of SUDs and identifying the early warning signs for SUDs in primary care are a critical part of effective approaches for prevention and early intervention for health outcomes.

The goal of this chapter is to establish the importance of screening and assessment of SUDs in primary care. It will address three types of SUD: (1) alcohol misuse; (2) illicit drug use such as heroin, cocaine, and marijuana; and (3) prescription drug use. In each area, we will review the epidemiological and population-based foundation for establishing these healthcare screenings. It will also identify which screening and assessment tools are reliable and valid for a primary care setting for adolescents and adult populations in primary care. And last, this chapter will address the barriers for incorporating these standardized tools and assessments into primary care. It will describe how screening tools are used in a clinical setting by defining roles and responsibilities for team members, identifying practice management and financial considerations, and describing relevant opportunities for quality improvement. In this chapter the terms “substance use disorders” (specifying which type of substance) and “substance misuse” will be used. These terms reflect evolving terminology in the field and the DSM-5 away from the use of “abuse,” which is seen as stigmatizing and an inaccurate description of behavior.

BACKGROUND AND ROLE OF SUBSTANCE ABUSE SCREENING AND ASSESSMENT IN PRIMARY CARE

The standard practice for evaluating patients’ risks for substance use disorders in primary care settings starts with screening. Screening is the process of measuring and detecting the signs and symptoms of a disorder to facilitate early identification and treatment. Screening is a population-based approach for prevention and early intervention in primary care settings. Consequently, screening for “at-risk” substance use behaviors in primary care is essential, yet often poorly implemented. While the evidence on the effectiveness of implementing tools for screening and assessment of substance use disorders is advancing, the application of these tools into standard clinical practice has lagged behind due to workflow and teamwork barriers in primary care settings (Talen, Baumer, & Mann, 2013).

Screening for behavioral health conditions within a population is a relatively new prevention initiative in health care. The US Preventative Services Task Force (USPSTF), which is now under the auspices of Agency of Healthcare Research and Quality (AHRQ)’s Prevention and Care Management Portfolio, makes healthcare recommendations based on the support of an Evidence-Based Practice.
Substance Abuse Screening and Assessment

These recommendations are based on systematic reviews of the evidence on specific topics in clinical prevention. These recommendations are meant to prevent complications, improve outcomes, and contain healthcare costs. The USPSTF recommends alcohol misuse screening for all adult patients and especially pregnant women, due to the increased risks to the fetus and pregnancy complications. USPSTF also recommends substance screening for patients who have tobacco use disorders, frequent trauma-related medical visits, and/or a family history of alcoholism, because these patients are at a greater risk for substance use disorders. None of the USPSTF systematic reviews found sufficient evidence to screen for substance use disorders in adolescent populations (Jonas, Garbutt, Brown et al., 2012).

The screening and detection of any disorder is only the first step and requires follow-up assessment and diagnosis. The purpose of screening is for managing the health of a targeted patient population. Screening is not a diagnostic litmus test, but an indicator that further assessment is needed. Clinical interviewing, professional judgment, and decision-making processes by a provider are not replaced with, but augmented by systematic screening and standardized psychological assessments. These tools supplement the provider–patient relationship, with the understanding that unique patient needs, style, and cultural and family backgrounds among other factors may influence the validity of a standardized tool. Choosing the appropriate tool among the many screening tools and assessment instruments is just one of the initial steps in implementation. The larger healthcare context plays a significant role in the effective process of screening and assessment (Talen et al., 2013).

Incorporating screening tools into clinical practice is a process that may begin with identifying the target population condition and choosing reliable standardized screening tools to detect “at-risk” patients. Then a follow-up diagnostic assessment is performed by a primary care provider (PCP) or behavioral health provider. Depending on the results of the patient’s assessments, they can be coached on health behavior change or referred to specialty care settings for treatment.

Most of the screening and assessment for substance use disorders is modeled after the Screening, Brief Intervention, Referral to Treatment Model (SBIRT) model (Agerwala & McCance-Katz, 2012; Babor et al., 2007; Harris, Louis-Jacques, & Knight, 2014; McCance-Katz & Satterfield, 2012). SBIRT is an evidence-based approach of early intervention in primary care for patients who have “at-risk” substance abuse behaviors ranging from misuse to dependence. The screening and brief intervention are the first line of clinical care. This is a combination strategy to screen for alcohol and substance use and for providers to immediately intervene with a brief intervention of the risks of misuse, offering patients appropriate psychopharmaceuticals (e.g., SSRIs, suboxone, naltrexone) and strategies and resources to change their behaviors. A follow-up session can be scheduled for patients who have a moderate or high level of substance dependence. Based on this diagnostic process, patients may receive short-term therapy, medication, and/or a referral to treatment. The SBIRT model has been widely studied and shown to be an effective model of care (Babor et al., 2007). A systematic review and meta-analysis found that brief alcohol intervention in primary care settings reduced alcohol consumption for both men and women at 6- and 12-month follow-ups (Agerwala & McCance-Katz, 2012; Babor et al., 2007; Harris et al., 2014).


Prevalence of Alcohol, Drugs, and Prescription Misuse and Dependence in Primary Care

In this section, the epidemiology of alcohol misuse, illicit drug use, and prescription drug use will be reviewed. The prevalence of substance misuse and disorders and the physical, social, and psychological
consequences of substance misuse will be identified. The impact of substance use on specific primary care populations, such as pregnant patients and adolescents, will also be described.

Alcohol
Alcohol misuse, which includes risky/hazardous drinking to alcohol dependence, is associated with numerous health and social problems (Strobbe, 2014). It is estimated to be the third leading cause of preventable mortality in the United States following tobacco use and obesity. These conditions span organ systems and include hypertension, gastritis, liver disease and cirrhosis, pancreatitis, certain types of cancer (e.g., breast and esophageal), cognitive impairment, anxiety, and depression (Corrao, Bagnardi, Zambon, & La Vecchia, 2004). In addition to disease processes, alcohol misuse has contributed to morbidity and mortality as a result of trauma, including falls, drowning, fires, motor vehicle crashes, homicide, and suicide (Cherpitel & Ye, 2008). Specifically in pregnancy, use of alcohol has been associated to a pattern of developmental abnormalities known as fetal alcohol syndrome, which occurs in about 0.2 to 1.5 per 1,000 live births in the United States (“Update: Trends in Fetal Alcohol Syndrome—United States, 1979–1993,” 1995).

Estimations of alcohol use in the general population have been difficult with studies ranging from 4% to 29% in the general population. Additionally, the definitions of alcohol misuse have changed over time and estimating the prevalence of alcohol misuse is challenging. Jonas, Garbutt, Amick et al. (2012) provide current definitions of the spectrum of alcohol misuse used by the USPSTF (2013), which is most often used in primary care settings.

Among those aged 12 and older, approximately 53% of US residents used alcohol in the past month (“Behavioral Health Trends,” 2015). Current estimates indicate that approximately 30% of the US population engages in alcohol misuse, with the majority of those being considered risky drinkers (Saitz, 2005). In the primary care setting, the American Academy of Family Physicians National Research Network demonstrated that 21% of patients revealed risky or hazardous drinking as assessed by the Alcohol Use Disorders Identification Test (AUDIT) and assessing the frequency of drinking questions. The proportion of patients with alcohol misuse seen in the primary care office is similar to the proportion of patients with diabetes and hypertension (Jonas, Garbutt, Amick et al., 2012; Jonas, Garbutt, Brown et al., 2012). Among adolescents, approximately 14% in the 8th grade and 41% in the 12th grade reported using alcohol at least once within the past 30 days; 7% and 23%, respectively, reported consuming at least five or more drinks on a single occasion (an episode of heavy use) within the previous two weeks (Cherpitel & Ye, 2008). Additionally, adolescents with at least one mental disorder are significantly more likely to have used alcohol and transition across the stages of alcohol misuse (Conway, Swendsen, Husky, He, & Merikangas, 2016).

Alcohol dependence is significantly more prevalent among men, whites, Native Americans, younger and unmarried adults, and those with lower incomes. In addition, significant disability has been associated with alcohol dependence. Unfortunately, only 24% of those with alcohol dependence were ever treated, which is slightly less than the treatment rate found 10 years earlier (Hasin, Stinson, Ogburn, & Grant, 2007). According to Hasin et al. (2007):

Strong associations between other substance use disorders and alcohol use disorders (odds ratios, 2.0–18.7) were lower but remained strong and significant (odds ratios, 1.8–7.5) when controlling for other comorbidities. Significant associations between mood, anxiety, and personality disorders and alcohol dependence (odds ratios, 2.1–4.8) were reduced in number and magnitude (odds ratios, 1.5–2.0) when controlling for other comorbidity.

(p. 830)

Based on these health risks, the USPSTF has made the following recommendations: for adults 18 and older, there is a B recommendation (meaning there is a high certainty that the net benefit is
moderate or there is moderate certainty that the net benefit is moderate to substantial) indicating that “primary care clinicians should screen for alcohol misuse and provide persons engaged in risky or hazardous drinking with brief behavioral counseling interventions to reduce alcohol misuse” (Moyer, 2013, p. 210). For the evaluation of adolescents, USPSTF has found insufficient evidence to recommend screening adolescents for alcohol misuse (Moyer, 2013).

Illicit Drug Misuse and Dependence
According to the 2014 National Survey on Drug Use and Health (NSDUH) conducted by the Substance Abuse and Mental Health Services Administration (SAMHSA), 10% (27 million) persons aged 12 and older reported using illicit drugs during the past 30 days (“Behavioral Health Trends,” 2015). This survey indicated higher use in every age range than in previous years. In the United States, marijuana has consistently been the most commonly used illicit drug, with approximately 8% of those aged 12 and older surveyed using in the past month. Approximately 4.2 million people, or under 2% of those 12 and older, met the criteria for marijuana use disorder. The marijuana use increase seems to account for the general increase of illicit drug use among all age ranges (“Behavioral Health Trends,” 2015). Cocaine is the second most commonly used nonprescription drug, and is used by 0.6% of the population (“Behavioral Health Trends,” 2015). Although a small minority of the population uses other illicit substances—hallucinogens (0.4%), inhalants (0.2%), heroin (0.2%)—great concern remains because the potential for misuse or dependence is quite high. Although it is difficult to tell yet if it is a trend, the use of heroin increased in the population and is a reason for concern (“Behavioral Health Trends,” 2015).

The time of peak use of illicit drugs is between the ages of 18 and 20, with percentages steadily decreasing as people become older. According to Polen, Whitlock, Wisdom, Nygren, and Bougatsos (2008):

Rates of illicit drug use vary significantly across racial/ethnic groups, with persons reporting more than one race having the highest rates (13.3%), followed by American Indians or Alaska Natives (12.3%), African Americans (8.7%), whites (8.1%), and Hispanics (7.2%). Men are more likely to engage in drug use than women (9.9% vs. 6.1%), but adolescent rates (age 12 to 17) of current illicit drug use are similar for boys and girls (10.6% for both).

(p. 1)

Although pregnant women are less likely to engage in illicit drug use when compared with nonpregnant women of the same age, numerous studies have found poor antenatal, neonatal, and childhood outcomes among women who used substance during pregnancy (Chasnoff et al., 1998).

Drug Misuse and Dependence
Again, according to the 2014 National Survey on Drug Use and Health conducted by the SAMHSA, nonmedical pain reliever use was the second most common type of illicit drug use. Approximately 4.3 million people, or 1.6% of the population aged 12 and older, use illicit drugs for pain. This percentage was, however, lower than percentages previously seen in 2002 to 2012 (“Behavioral Health Trends,” 2015). Of the general population, the same survey demonstrated that 1.9 million persons, or 0.7%, had a pain reliever use disorder. Approximately 1.9 million people, or 0.7% of the US population aged 12 and older, are nonmedical users of tranquilizers. Additionally, there are approximately 1.6 million persons, or 0.6% of the US population aged 12 and older, who are nonmedical users of stimulants, including illegally manufactured methamphetamines (“Behavioral Health Trends,” 2015).
SCREENING FOR SUBSTANCE MISUSE AND DEPENDENCE IN PRIMARY CARE

This section will address the variety of screening tools for alcohol and other addictive substances in primary care settings. Alcohol screening has received the most attention for developing and implementing evidence-based screeners. The screening tools for illicit and prescription drug misuse will also be reviewed. Given the growing public health epidemic of opioid and benzodiazepine addictions, these tools are not yet as prevalent in primary care. However, new and existing tools will need to be augmented, developed, evaluated, and revised.

Screening Tools in Primary Care

Other than tobacco, alcohol has been a primary focus for substance use disorder screening in primary care. While there is a 30% prevalence rate of alcohol misuse disorders in the adult population the United States, identifying and referring at-risk patients occurs less than 10% of the time in primary care, indicating that substance use is inadequately and inconsistently screened for in those settings (Babor et al., 2007; McCance-Katz & Satterfield, 2012). There are a substantial number of studies that support a variety of reliable and valid screening measures for alcohol misuse. In contrast, there are fewer studies on screening for illicit drug or prescription drug misuse. Consequently, there is insufficient evidence for brief, valid, and reliable screening tools for illicit drug use in primary care (Babor et al., 2007; Fiellin, Reid, & O’Connor, 2000; McCance-Katz & Satterfield, 2012).

There have been a number of screening instruments that have been developed to detect the level of risk in the spectrum of alcohol misuse (see Table 15.1). These include a range of questionnaires from a single-question screening to self-report questionnaires (e.g., AUDIT, AUDIT-C) to clinical interviewing questions (e.g., Cut Down, Annoyed, Guilty, and Eye-Opener (CAGE). There are also a number of tests designed specifically for pregnant women, such as the Tolerance, Annoyed, Cut-Down, and Eye-Opener (T-ACE) and Tolerance, Worried, Eye-Openers, Amnesia, Kut-Down (TWEAK).

Table 15.1 Alcohol and Substance Use Disorder Screening Tools

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Measurement</th>
<th>Administration and Scoring</th>
<th>Target Population</th>
<th>Psychometrics and Interpretation (sensitivities; specificities)</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Screening</td>
<td>How many drinks do you have in one week? Have you ever had more than seven (men) or five (women) at one time?</td>
<td>Clinical Question: &gt;7 drinks for men and &gt;5 drinks for women = positive screen</td>
<td>Adults (&gt;18–65)</td>
<td>High to moderate ranges (0.82–0.87; 0.61–0.79)</td>
<td><a href="http://drugabuse.gov">http://drugabuse.gov</a></td>
</tr>
<tr>
<td>Initial Screeners</td>
<td>AUDIT: Alcohol Use disorders Identification Test</td>
<td>Self-report questionnaire 10 Questions</td>
<td>Adults (&gt;18) Normed with diverse populations</td>
<td>Misuse: &gt;5;4 Risky: &gt;7;6 Dependent: &gt;12</td>
<td><a href="http://drugabuse.gov">http://drugabuse.gov</a></td>
</tr>
</tbody>
</table>

(Continued)
### Table 15.1 (Continued)

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Measurement</th>
<th>Administration and Scoring</th>
<th>Target Population</th>
<th>Psychometrics and Interpretation (sensitivities; specificities)</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>T-ACE; TWEAK</td>
<td>Clinical Interview: 4 questions</td>
<td>Women</td>
<td>Low to high ranges (.07–.89–.073–.085)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAST</td>
<td>Self-report 20, 10</td>
<td>Adults</td>
<td>Misuse &gt;6 (0.41-0.95-0.68-0.99)</td>
<td><a href="http://www.integration.samhsa.gov/clinical-practice/screening-tools">http://www.integration.samhsa.gov/clinical-practice/screening-tools</a></td>
<td></td>
</tr>
<tr>
<td>ASSIST: Alcohol, Smoking, Substance Involvement Test</td>
<td>Self-report 8 questions</td>
<td>Adults</td>
<td>Low risk&gt;3 Moderate&gt;4-26 Problem&gt;27 (0.90-0.58)</td>
<td><a href="http://www.integration.samhsa.gov/clinical-practice/screening-tools">http://www.integration.samhsa.gov/clinical-practice/screening-tools</a></td>
<td></td>
</tr>
<tr>
<td>Adolescent Screeners</td>
<td>CRAFFT</td>
<td>Self-report</td>
<td>Adolescents</td>
<td>High to moderate ranges 0.92 (0.88-0.96); 0.64 (0.59–0.69)</td>
<td><a href="http://www.ceasar.org">www.ceasar.org</a></td>
</tr>
<tr>
<td>Brief Opioid Questionnaire (BSQ)</td>
<td>Self-report</td>
<td>Adults</td>
<td>TBD</td>
<td><a href="http://www.painmed.org/2014posters/abstract-206/">http://www.painmed.org/2014posters/abstract-206/</a></td>
<td></td>
</tr>
</tbody>
</table>

Other tools include the Michigan Alcoholism Screening Test, the Rapid Alcohol Problems Screen, and the Alcohol-Related Problems Survey. Several systematic reviews compared the characteristics of these screening tools in primary care populations (Harris et al., 2014; Johnson, Lee, Vinson, & Seale, 2013; Spithoff & Kahan, 2015; Timko, Kong, Vittorio, & Cuacciare, 2016). Based on these reviews, the USPSTF recommends the full AUDIT instrument, the abbreviated AUDIT-C, and single-question screening (Johnson et al., 2013; Moyer, 2013). These tools have the best specificity and sensitivity for detecting the full spectrum of alcohol misuse in adults, young adults, and pregnant women (Jonas et al., 2013).
Single-Question Screening
The screening questions for risky alcohol use start with a single question about the number of drinks an individual has had over a period of time (Babor et al., 2007; McNeely et al., 2015). The single question for assessing alcohol misuse is “How many times in the past year have you had 5 [for men] or 4 [for women and all adults older than 65 years] or more drinks in one day?” During the validation for this single-question assessment, when the researcher associates in the study were asked, they were instructed to clarify that a standard drink referred to 12 ounces of beer, 5 ounces of wine, or 1.5 ounces of 80 proof spirits. The sensitivities range from 0.82 to 0.87 and specificities range from 0.61 to 0.79 for detecting misuse among adults seen in primary care settings (Smith, Schmidt, Allensworth-Davies, & Saitz, 2009). A second single question, “How many drinks do you have in one week?” screens for the frequency of use (>7 for men; >5 for women). This screening questions on binge and frequency of drinking are endorsed by the National Institute on Alcohol Abuse and Alcoholism (NIAAA; Harris et al., 2014; McNeely et al., 2015).

Alcohol Use Disorders Identification Test (AUDIT)
The AUDIT is the gold standard for detecting alcohol misuse, which has high levels of sensitivity and specificity in primary care populations (Aalto, Alho, Halme, & Seppa, 2009; Bischof et al., 2007; Boschloo et al., 2010; Burns, Gray, & Smith, 2010). The AUDIT was originally developed to detect heavy drinkers, including those who are in the early stages without significant alcohol-related harms. There is an extensive literature citing its effectiveness in detecting heavy drinking in medical settings (Aalto et al., 2009; Bischof et al., 2007; Boschloo et al., 2010). The AUDIT is available as a clinician or self-administered structured questionnaire with 10 questions (8 multiple choice) asking respondents to indicate frequency of behavior from “never” to “daily or almost daily” (4-point Likert scale), and 2 yes/no questions. The total AUDIT score (range 0–40) is the most widely used index with scores 0–7 indicating low-risk drinking; 8–15 indicating risky, hazardous drinking; 16–19 indicating high-risk, harmful drinking; and 20 or more indicating high-risk, dependent drinking.

This questionnaire was developed with the support of the World Health Organization (WHO) and it has been normed across culturally diverse populations worldwide. The first three questions ask patients about the amount of alcohol consumed, and the remaining seven questions address the consequences caused by alcohol (Rist, Glockner-Rist, & Demmel, 2009). Although the AUDIT consists of only 10 questions, it is sometimes too long for a fast-paced primary care setting. Therefore, several abbreviated versions have been developed. The AUDIT-C consists of the first three questions that identify amount of alcohol consumed (Aalto et al., 2009; Johnson et al., 2013). The AUDIT-QF includes the two first questions regarding quantity consumed and frequency of drinking (Aalto et al., 2006). The AUDIT-3 consists of only the third question from the original AUDIT regarding binge drinking. Encouraging results have been obtained from these abbreviated versions, indicating that they may be as effective at identifying alcohol misuse as the original AUDIT (Aalto et al., 2009; Bischof et al., 2007; Johnson et al., 2013).

The scoring profile identifies patients on a continuum ranging from risky drinking behaviors (scores >5 for men; >4 for women) to alcohol misuse (scores >7 for men; >6 for women) to dependency (scores >12). The AUDIT shows an optimal balance of sensitivity and specificity for detecting all forms of alcohol misuse when cutoff points of 4 or more (sensitivity, 84%–85%; specificity, 77%–84%) or 5 or more (sensitivity, 70%–92%; specificity, 73%–94%) are used (Gache et al., 2005). Use of higher cutoff points increases specificity but reduces its sensitivity. The sensitivity and specificity of AUDIT-C are best when cutoff points of 4 or more (74%–76% and 80%–83%, respectively) and three or more (74%–88% and 64%–83%, respectively) are used. The briefer AUDIT-C version has similar
high rates of sensitivity and moderate specificity compared to the full version. The AUDIT and these shorter versions have been shown to be valid for detecting heavy drinking also in a general population sample (Aalto et al., 2009).

In addition, AUDIT and AUDIT-C are effective for both males and females. The optimal cut-points for males were found to be ≥7 or 8 for the AUDIT and ≥6 for AUDIT-C. Among females the optimal cut-points were found to be ≥5 for AUDIT and ≥4 for the AUDIT-C (Aalto et al., 2009). Scores for detecting unhealthy alcohol use improves by combining AUDIT cutoffs of 6 for men and 4 for women with a 30-day binge drinking measure (Johnson et al., 2013). AUDIT scores of 15 for men and 13 for women detected alcohol dependence with 100% specificity but has compromised with low sensitivity (18%). Combining lower AUDIT cutoff scores and binge drinking measures may increase the detection of unhealthy alcohol use in primary care (Johnson et al., 2013).

**CAGE**
The CAGE, an acronym for Cut down, Annoyed, Guilt, and Eye-opener, is a standard clinical interviewing screening protocol in primary care and is integrated into a routine history taking. A positive response indicates further assessment. It targets alcohol misuse and dependence, but is less effective at detecting risky drinking behaviors (Fiellin et al., 2000). The screening questions have shown less accuracy in identifying older patients, African Americans, and Latinos with substance abuse symptoms. Although the CAGE questionnaire has frequently been used in primary care settings as a low-burden screening tool for alcohol disorders, it has comparatively poor sensitivity for identifying risky or hazardous drinking, particularly among older adults (14%–39%) and pregnant women (38%–49%). The CAGE questionnaire has commonly been used because it is perceived as a low-burden screening tool in the primary care setting. Unfortunately, it has poor sensitivity especially when evaluating for risky drinking among adults, older adults, and pregnant women (Jonas, Garbutt, Amick et al., 2012).

A number of hybrid screening tools have been used within primary care settings for special primary care groups (e.g., women, pregnant women, seniors). Combining a few questions from CAGE and AUDIT has created a variety of this type of screening tool. For example, TWEAK (Tolerance, Worried, Eye-opener, Amnesia, K/Cut Down) and TRACE (Tolerance, Annoyance, Cut-down, Eye-opener) were developed specifically for alcohol screening with women, while the Five Shot and Rapid Alcohol Problems Screen (RAPS) have been validated as screening questions for men and women across diverse ethnic groups (Bischof et al., 2007) (Burns et al., 2010). For pregnant patients, at-risk drinking sensitivity was highest for T-ACE (69–88%), TWEAK (71–91%), and AUDIT-C (95%), with high specificity (71–89%, 73–83%, and 85%, respectively). Sensitivity of AUDIT-C at score 3 was high identifying alcohol dependence (100%) or alcohol use disorder (96%) with pregnant patients, but had only moderate specificity (71% each; (Burns et al., 2010).

**Michigan Alcohol Screening Tool (MAST and bMAST)**
The Michigan Alcoholism Screening Test is a 25-item, self-report questionnaire that has been used primarily with a clinical population rather than a general primary care population. The majority of the items on the MAST refer to consequences of using alcohol in lifetime situations (e.g., “Have you ever been arrested for drunk driving or driving after drinking?”; “Have you ever lost friends or girlfriends/boyfriends because of drinking?”; “Do you feel you are a normal drinker?”; “Do friends or relatives think you are a normal drinker?”). The MAST remains a widely used clinical assessment instrument (Connor, Grier, Feeney, & Young, 2007) and is a robust predictor of alcohol dependence in medical (Magruder-Habib, Stevens, & Alling, 1993) and psychiatric settings (Magruder-Habib et al., 1993).
The Brief Michigan Alcoholism Screening Test (bMAST) has been applied across a range of clinical and research settings and as an alcohol-related problem severity. Construct validity studies support using the bMAST single (e.g., severity) or two-factor scores (e.g., perception of drinking and negative consequences) as an index of alcohol problems (e.g., moderate [0.50–0.68] correlations with AUDIT, which has been a useful as a tool for aiding providers in treatment decision-making [Connor et al., 2007]).

**Drug Assessment Screening Tool (DAST)**
The Drug Abuse Screening Test (DAST) is a 28-item, face-valid self-report measure of problematic substance use that is used for clinical screening and treatment/evaluation research. Responses to the DAST are yes/no answers, with each item scoring one point for a “yes” response, thus yielding a total score ranging from 0 to 28. A cutoff score of 6 is generally used to indicate problem use (Yudko, Lozhkina, & Fouts, 2007). It can be used by both nonprofessional and professional personnel and takes 5–10 minutes to administer. There are a variety of shorter versions including the DAST-10 and DAST-A for adolescents. Although the sensitivity of the DAST is usually quite high, it varies significantly according to the group being studied. The sensitivity and specificity of the DAST ranges from 80% to 94% for populations that are more aligned with a primary care group. For DAST-10, results of studies on its sensitivity ranged from 95% to 41% while specificity ranged from 68% to 99%. For the DAST-A, sensitivity at the cutoff score of 6 was reported at 78.6%, specificity was reported at 84.5% (DSM-IV), and positive predictive power was reported in a range of 79% to 82.3% (DSM-IV) (Martino, Grilo, & Fehon, 2000). In general, specificity of the DAST is increased when the cutoff score is high but the sensitivity decreases (Yudko et al., 2007).

**Alcohol, Smoking, and Substance Involvement Test (ASSIST)**
The ASSIST is a screening tool for alcohol, smoking, and other substances that has been recommended by the WHO (O’Grady, Gryczynski, Mitchell, Ondersma, & Schwartz, 2016; Rubio Valldolid et al., 2014; Silva, Lucchese, Vargas, Benicio, & Vera, 2016; Tiburcio Sainz et al., 2016; “The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, Reliability and Feasibility,” 2002). Its purpose is to identify the use of psychoactive substances and any consequences from their use. ASSIST contains eight questions and measures the level of an individual’s dependence on substances. There are seven questions regarding the use and problems related to smoking, alcohol, marijuana, cocaine, stimulants, inhalants, hypnotics/sedatives, opiates, and hallucinogens; there is only one question on drug injection. Scores lower than 3 (or 10 for alcohol) identify that the person is at low risk for substance use–related problems; mean scores of 4 (or 11 for alcohol) to 26 are indicative of harmful or problematic substance use; scores above 27 for any substance suggest that the person is highly dependent. It has high reliability scores (0.90–0.58) and validation studies (“The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST),” 2002; McNeely et al., 2015). The level of the score also helps guide health providers toward making recommends for the intensity of interventions (Silva et al., 2016).

**Adolescent Screeners**

**CRAFFT**
The CRAFFT (CAR, RELAX, ALONE, FORGET, FRIENDS, TROUBLE) is a screening tool for adolescents aged 14 and older. It consists of a series of six questions developed to screen adolescents for high-risk alcohol and other drug use disorders. It has been endorsed by the American Academy of Pediatrics (AAP; 2001) for use with adolescents. A recent study found that a physician-conducted CRAFFT
screening interview required an average of 74 seconds to complete, whereas a computer self-administered version took an average of 49 seconds. Adolescents preferred paper forms and computerized questionnaires over interviews with physicians or nurses (Harris et al., 2014; Harris et al., 2016).

There have been several reviews that examine substance use screening instruments commonly used with adolescents in medical settings, their comparative usefulness, and the effectiveness of SBIRT (Harris et al., 2014; Pilowsky & Wu, 2013). CRAFFT has been the most studied instrument for screening for alcohol/drug use and related problems among adolescents, and it has strong evidence to support its use in medical settings (Mitchell et al., 2014). CRAFFT has been validated against DSM-IV and DSM-5 diagnostic criteria. As a continuous measure, area under the curve (AUC) values were 0.93 for problem use or higher and 0.97 for DSM-5 SUD. The cutoff point of 2 was the optimal score for identifying adolescents with alcohol-related problems and dependency. Despite changes in the DSM substance use diagnostic criteria, CRAFFT demonstrates strong sensitivity and specificity at its established cut-point of 2 (Mitchell et al., 2014).

The sensitivity of the CRAFFT is similar to the longer AUDIT and Problem Oriented Screening Instrument for Teenagers (POSIT) tests, and much greater than CAGE, which is not recommended for use with adolescents (Knight, Sherritt, Harris, Gates, & Chang, 2003). Optimal cut-points associated with problem use or higher were 2 for AUDIT, 1 for POSIT, 1 for CAGE, and 1 for CRAFFT. The CRAFFT works equally well for alcohol and drugs, for boys and girls, for younger and older adolescents, and for youth of diverse races and ethnicities (Knight et al., 2003).

The CRAFFT and AUDIT tools currently have the strongest validity for adolescents, whereas the validity of other widely used tools such as DAST-10, the National Institute on Drug Abuse (NIDA)-modified ASSIST (Alcohol, Smoking and Substance Involvement Screening Test), and ultra-brief screens (AUDIT-C, single-item screens) do not have strong enough evidence to support their use with adolescents. Studies are still needed to identify effective strategies to promote universal adolescent screening and the use of valid screening tools in general medical settings (USPSTF, 2010). Practice guides from the AAP and NIAAA, however, do provide a structured algorithm for specific screening and the recommended type of intervention based on level of risk (Harris et al., 2014).

**Brief Risk Questionnaire (BRQ)**

Opioid misuse is a growing concern in population health and primary care (Brown et al., 2011). Opioid risk assessment has been needed to address the potential addiction in medical settings. The Brief Risk Questionnaire (BRQ) was created and compared with a structured clinical interview and two commonly used patient-completed risk assessment tools: The Opioid Risk Tool (ORT) and Screener and Opioid Assessment for Patients with Pain-Revised (SOAPP-R). Results found that the BRQ was able to predict future medication risk behaviors better than the other two patient-completed risk measures and almost as well overall as a structured clinical interview rating system (Jones, Lookatch, & Moore, 2015). This BRQ holds promise as a useful new tool in assessing opioid risk assessment (Lopez-Pelayo, Batalla, Balcells, Colom, & Gual, 2015).

**Summary of Screening for Substance Use**

There is fair evidence that standardized questionnaires such as AUDIT, DAST and Single-Question Screener are considered short enough to be potentially useful in the practice setting and have acceptable accuracy and reliability in screening for drug use/misuse. The CRAFFT has been adequately validated for screening adolescents for drug use/misuse. Several other instruments of various lengths (ASSIST, CAGE-AID, and BRQ) have been validated for screening adults (Rumpf, Hapke, Meyer, & John, 2002). However, these tools are lacking the necessary positive predictive value when used in a
general medical patient population. Consequently, there is insufficient evidence for clinical use of these instruments (ASSIST, CAGE-AID, BRQ) when applied in a busy primary care practice setting, and especially in screening with pregnant women for drug use (Lanier & Ko, 2008).

**DIAGNOSTIC ASSESSMENT MEASURES FOR SUBSTANCE USE DISORDERS**

After an adult has a positive screen for at-risk substance misuse or dependence, a more thorough psychological assessment is indicated. The assessment process can be lengthy and require a higher level of patient literacy and clinical time, as well as a provider with credentials to complete a diagnostic assessment. Psychology providers or other behavioral health consultants should choose measures that align with the purpose of the assessment (e.g., detection vs. severity vs. treatment intensity) and characteristics of the patient population (e.g., cultural, underserved, comorbid conditions). Assessment is an opportunity for providers to understand patients’ needs and develop interventions that are consistent with the level and intensity of care (Sweetman, Raistrick, Mdege, & Crosby, 2013). In a recent systematic review of substance use disorder assessment tools, the authors found only 6 assessment instruments out of 595 referenced in English databases with strong psychometric properties (Sweetman et al., 2013). This list below is a brief summary of these instruments along with a several other comprehensive assessments based on DSM criteria. Table 15.2 summarizes the properties of each assessment instrument.

<table>
<thead>
<tr>
<th>Table 15.2 Substance Abuse Diagnostic Assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Assessment Type</strong></td>
</tr>
<tr>
<td><strong>Diagnostic Tools</strong></td>
</tr>
<tr>
<td>Structured Clinical Interview for DSM-5 (SCID)</td>
</tr>
<tr>
<td>Psychiatric Research Interview for Substance and Mental Disorders (PRISM)</td>
</tr>
<tr>
<td>Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA)</td>
</tr>
<tr>
<td>Chemical Use, Abuse and Dependence Scale (CUAD)</td>
</tr>
<tr>
<td>Form 90</td>
</tr>
</tbody>
</table>

(Continued)
### Table 15.2 (Continued)

<table>
<thead>
<tr>
<th>Assessment Type</th>
<th>Measurement</th>
<th>Administration and Scoring</th>
<th>Target Population</th>
<th>Psychometrics and Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maudsley Addiction Profile (MAP)</td>
<td>Brief 26-item structured interview. Four domains: use; injecting; sexual behavior; and physical, mental, social functioning</td>
<td>Adult</td>
<td>Concurrent; no sensitivity or specificity reported</td>
<td></td>
</tr>
<tr>
<td>Measurements in the Addictions for Triage and Evaluation (MATE)</td>
<td>Clinical interview with prompts. Seven domains: use, mental, physical, dependence, physical signs, personality, and functioning</td>
<td>Adult</td>
<td>Construct validity and factor analysis; no sensitivity–specificity rates</td>
<td></td>
</tr>
<tr>
<td>Substance Abuse Outcomes Module (SAOM)</td>
<td>Computer analysis. Useful for treatment outcomes. Self-report 110 questions on activities, health, friends/family, use, medications, dependence, problems, occupation and housing</td>
<td>Adult</td>
<td>Concurrent validity; high sensitivity and specificity (90%; 94%)</td>
<td></td>
</tr>
<tr>
<td>Addiction Severity Index (ASI)</td>
<td>Interview and labs results: 163 questions; 50–60 minutes Seven subscales: physical, employment, alcohol, drug, offending, family/social, and mental health</td>
<td>Adults. Available in many languages.</td>
<td>Strong validity and reliability; high sensitivity and specificity (0.83–0.70)</td>
<td></td>
</tr>
<tr>
<td>Readiness to Change</td>
<td>Change Assessment Questionnaire (URICA) Self-report; readiness for change scales</td>
<td>Adults/adolescents</td>
<td>Positive predictive value (PPV) for treatment progress</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Adolescent Substance Abuse Goal Attainment: ASAGA Self-report</td>
<td>Adolescents</td>
<td>PPV for treatment goals and progress</td>
<td></td>
</tr>
</tbody>
</table>

---

### The Semi-structured Assessment for Drug Dependence and Alcoholism (SSADDA)

SSADDA is a computer-assisted psychiatric interview schedule developed to assess the physical, psychological, social, and psychiatric manifestations of substance use (Malison et al., 2011; Pierucci-Lagha et al., 2005). The SSADDA was created from the Semi-structured Assessment for the Genetics of Alcoholism (SSAGA; Bucholz et al. 1994, 1995) to include along with alcohol a study of the influence of genetics on cocaine and opioid dependence. A feature of the SSADDA is the ability to assess for psychiatric disorders independently of substance problems, which have historically confounded the study of the former.

Included in the SSADDA are questions about the onset and progression of symptoms for cocaine, opioid dependence, and other substances. The interview uses the abuse and dependence criteria based on DSM-IV-TR. The SSADDA is designed to be administered by trained (nonclinician) interviewers. The computerized administration allows for “skip outs” and is also meant to improve accuracy of
collected information and reduce the reliance on the interviewer’s clinical judgment. The data also directly uploads to a database.

Studies have investigated the reliability and validity of the SSADDA (Malison et al., 2011; Pierucci-Lagha et al., 2005; Pierucci-Lagha et al., 2007). In a study, Pierucci-Lagha et al. (2005) examined the reliability of individual DSM-IV criteria for lifetime substance dependence diagnoses and the impact of those criteria on diagnostic reliability. The SSADDA was found to have moderate to strong reliability on the identification of abuse and/or dependence of nicotine, alcohol, opioids, cannabis, stimulants, sedatives, and other drugs. In that study, the $\kappa$ coefficient was used to analyze that data. Results demonstrated that the SSADDA is better at assessing dependence as defined by the DSM-IV-TR rather than abuse symptoms alone. Further, the SSADDA has better interrater reliability for opioid dependence (excellent) versus stimulants or sedative dependence (fair). Combined dependence scores for dependence with abuse improved except for sedatives, which decreased, and other drugs, which stayed the same. SSADDA had the strong interrater reliability for opioid dependence and the poor reliability for cannabis dependence (Pierucci-Lagha et al., 2005; Pierucci-Lagha et al., 2007). A study of opioid dependence in a Thai population explored diagnostic reliability and validity of a Thai version of the SSADDA and demonstrated that the Thai SSADDA and the Thai Mini Neuropsychiatric Interview (MINI) had high concordance ($\kappa = 0.97$) in the diagnosis of opiate dependence (Malison et al., 2011).

The reliability and validity established for this tool relates to the DSM-IV and has yet to be fully explored for the DSM-5. As a result, the reports on diagnostic reliability and validity are based on the criteria used in the DSM-IV such as “abuse” and “dependence.” Further, the length of test administration is significant such that it is most appropriate for research studies like the study of genetic linkage rather than clinical diagnoses and treatment planning in primary care.

**Addiction Severity Index (ASI)**

This assessment has 163 items with seven subscales: physical health, employment, alcohol use, drug use, offending, family and social circumstances, and mental health status (Butler, Redondo, Fernandez, & Villapiano, 2009; Cacciola, Alterman, McLellan, Lin, & Lynch, 2007; Diaz Mesa et al., 2010). The substance use section includes items pertaining to 11 drugs or classes of drugs and the frequency that the individual has used these drugs in the past 30 days. This is a time-intensive assessment (up to 60 minutes to complete) and requires higher levels of health literacy. It has been widely reviewed, found to have strong validity studies and reliability indexes, and has been normed with Spanish populations. Sensitivity in medical settings is 0.83 and specificity is 0.70 (Denis et al., 2016; Denis, Cacciola, & Alterman, 2013; Sweetman et al., 2013).

**Chemical Use, Abuse, and Dependence Scale (CUAD)**

The CUAD scale involves a two-stage assessment process. The first stage is an extensive drug use history use for nine drugs. The interviewer asks 17 follow-up questions on the level of dependence and problems associated with each drug used. This is a useful format validating a DSM substance use disorder (McGovern & Morrison, 1992). However, it has only concurrent validity information and no other reported psychometric properties (McGovern & Morrison, 1992; Sweetman et al., 2013).

**Form 90**

This semi-structured interview has 121 items in five domains—healthcare utilization, medication, offending, vehicle accidents, and employment—along with an item assessing frequency of substance
use over 90 days. It is another time-intensive assessment with criterion validity and test-retest reliability but low to moderate levels of sensitivity; except for opioid users, specificity was very strong (e.g., no false positives) (Sweetman et al., 2013).

**Maudsley Addiction Profile (MAP)**

The MAP is a brief, structured interview in five sections: use, injecting, sexual behavior, physical and psychological health, and social functioning. There are 26 interview items and a list of the quantity, frequency, and means of drugs used. There are several self-report questions. This assessment has concurrent validity, test-retest reliability, and factor analysis but no reported sensitivity or specificity.

**Measurements in the Addictions for Triage and Evaluation (MATE)**

This semi-structured interview has six sections: substance use, mental, physical dependence, symptoms, personality, and functioning. It assesses for quality, frequency, and lifetime use of 12 drugs, and it also includes gambling. There is also a needs assessment of social function and self-report questions on depression, anxiety, and cravings (Schippers, Broekman, Buchholz, Koeter, & van den Brink, 2010; Sweetman et al., 2013).

**Substance Abuse Outcomes Module (SAOM)**

This self-report, 110-item questionnaire covers activities, health, friend/family, substance use, medications, dependence, and related occupation and housing problems. There is also a detailed section on the quantity and frequency of substance use. This is a tool designed for treatment outcomes and can be computer administered and scored (Smith et al., 2006; Sweetman et al., 2013).

**Psychiatric Research Interview for Substance and Mental Disorders (PRISM)**

The Psychiatric Research Interview for Substance and Mental Disorders (PRISM; Caton et al., 2005) is a semi-structured diagnostic interview tool. Originally designed for research purposes, the authors suggest that it may be useful in matching treatment modalities to patients based on presenting comorbidities. The strength of the PRISM is its ability to identify psychiatric disorders independent of substance use disorders and associated symptoms. The PRISM corresponds to DSM-IV-TR criteria for abuse and dependence. The PRISM specifically addresses DSM-IV-TR current and lifetime criteria for substance use and dependence for the substances, Axis I and Axis II disorders, listed in Table 15.3.

After a series of questions there is a time line that is gathered and recorded mostly to determine if psychiatric symptoms preceded substance use or misuse. The PRISM can be hand or computer administered. Although the administration of the PRISM is lengthy (120 minutes), it is separated into modules that can be used independently, based on the research needs. Training is required for the PRISM, and a self-study training manual accompanies the PRISM. Additionally, interviewers go through a 2-day training, which includes group didactic sessions and role-playing with a trained expert, followed by taped interviews to assess fidelity and quality assurance.
Test-retest reliability and construct validity have been found for the PRISM. In a test-retest study of 285 substance-using participants, good to excellent reliability (using the kappa coefficient to determine agreement: kappas such that a reliability coefficient below 0.40 is considered to be poor; 0.40–0.59 is fair; 0.60–0.74 is good; and 0.75–1.00 is excellent) was found for substance dependence as defined by the DSM-IV for alcohol ($k = 0.82$), cocaine (0.90), cannabis (0.73), heroin ($k = 0.94$), opiates ($k = 0.62$), and sedatives (0.74). The substance abuse criteria used in the study was slightly less reliable, generally ranging from poor to fair. Prevalence of heroin and stimulant abuse was too low to determine reliability. Reliability of dependence severity and age of onset of dependence both ranged from good to excellent (Gonzalez-Saiz et al., 2009; Hasin et al., 2007).

The construct validity of the PRISM-IV was supported by studies with both Spanish- and English-speaking patient populations. Results demonstrated that the PRISM-IV was good to excellent at identifying past or current substance abuse and dependence. In a 2004 study, a Spanish version of the PRISM-IV and the SCID were compared against expert psychiatrist diagnosis using the Longitudinal Expert All Data (LEAD) procedure. Results demonstrated that when using the LEAD procedure, the Spanish version of the PRISM-IV was generally similar to the Spanish version of the SCID and both were good or excellent at identifying past or current substance abuse and dependence. The Spanish version of the PRISM had significantly better concordance with the LEAD when identifying current cannabis and cocaine dependence; and past alcohol abuse and dependence when compared to the SCID and LEAD (Torrens, Fonseca, Mateu, & Farre, 2005).

In a 2005 study of 400 patients presenting to an emergency room, PRISM was used to differentiate between substance-induced psychosis and primary psychotic disorders. The authors noted as a potential limitation that white participants were more likely to be diagnosed with primary psychotic disorders and nonwhites with substance-induced psychosis. This may represent a bias inherent in the PRISM (Caton et al., 2005). Overall, the PRISM was developed with many purposes in mind; however, the lengthy administration time may limit the feasibility of its use in many community, primary care, and/or other clinical settings.

**Rhode Island Change Assessment Questionnaire (URICA)**

The URICA is a 32-item, self-report measure that includes four subscales measuring the stages of change: Precontemplation, Contemplation, Action, and Maintenance (Hasler, Klaghofer, & Buddeberg, 2003; Henderson, Saules, & Galen, 2004). The subscales can be combined to assess readiness to change throughout the treatment process. This questionnaire has had predictive utility of the
stages-of-change scales in a variety of patient populations, including tobacco-, alcohol-, and heroin-addicted polysubstance-use patients. The questionnaire has internal reliability, content, and criteria validity studies. The Maintenance scale was positively related to drug-free symptoms and length in treatment (Hasler et al., 2003; Henderson et al., 2004).

**Adolescent Substance Abuse Goal Commitment (ASAGC)**

Commitment to change is an important factor in treatment of substance use disorders. The Adolescent Substance Abuse Goal Commitment (ASAGC) questionnaire is a 16-item, 4-point Likert scale developed to assess an adolescent’s commitment to his/her stated treatment goal (Kaminer, Ohanessian, McKay, & Burke, 2016). An exploratory factor analysis was conducted on the ASAGC items, which resulted in two scales: Commitment to Recovery and Commitment to Harm Reduction. The ASAGC scales were found to demonstrate a high level of internal consistency (alpha coefficients ranged from .92 to .96 over time) (Kaminer et al., 2016). The ASAGC is a reliable and valid clinical research instrument that can be used for the assessment of adolescents’ commitment to their substance abuse treatment goals. Clinicians may take advantage of the clinical utility of the ASAGC including its ability to differentiate between commitment to abstinence versus commitment to harm reduction (Kaminer et al., 2016).

**GUIDELINES FOR COLLABORATIVE TEAM-BASED CARE AND CLINICAL DECISION-MAKING**

An effective primary care process for screening and assessment of substance misuse in primary care settings should follow the clinical algorithm from the evidence-based SBIRT model. A team-based model requires multiple steps and factors to work effectively in primary care settings. These processes include three core areas: (1) who is involved, (2) how the clinical processes flow, and (3) what system supports are needed to sustain these processes (Talen et al., 2013). First, a clinical team needs to meet and decide which team members will be involved with the screening, interventions, assessment and follow-up. The physician, behavioral health consultant, and nurse or medical assistant (MA) all have a role in this screening process such as administration and scoring of the screening instrument. Team members need to define the roles and responsibilities of each member, including the requirements for licensure with respect to administration, scoring, and feedback of the screening and/or assessment results. Teams should decide how the results from screeners or assessments are documented and communicated with providers. Treatment teams may have access to each other’s evaluation and assessment of patients’ functioning within their electronic medical record (EMR) system. In some systems, however, there may be confidentiality policies that limit access to medical and mental health assessments.

Second, the clinical workflow needs to be clearly outlined. The team needs to review clinical guidelines on who needs to be screened and when and how often the screening should occur. Annual or new patient exams should include a screening for substance misuse. In addition, there needs to be clear guidelines on how positive results from a screening tool are managed. For example, an MA may ask the patient to complete the AUDIT and DAST and then score the results, which are then entered into the EMR. The physician reviews the score before seeing the patient and identifies the level of risk for misuse of substances and other possible medication–drug interactions or comorbid health risks for the patient. If the patient has a low risk, the physician may provide a brief intervention using motivational interview techniques and education on healthy behaviors (e.g., driving and substance
use). A moderate level of dependence would prompt the behavioral health consultant to meet with the patient as a warm handoff, review the results with the patient, and provide follow-up diagnostic assessments and brief interventions. If the patient has significant substance dependence, physicians or the psychologist may provide referrals for follow-up with specialty care inpatient or outpatient services. The team should also have “shared care plans” where the medical and psychological treatment plans are coordinated between providers.

The last factor for clinical implementation focuses on the sustainability and quality improvement with screening and assessment processes. Billing for the screening tools and assessments needs to be included in the EMR system. In addition, providers need to document results in the same field, and interventions and plans should be treated as shared documents to ensure continuity and collaborative communication. Documentation in the EMR will also provide data for tracking patient care indicators, follow-up services, and quality measures. This provides a platform for teams to institute quality improvement plans that can be analyzed, reviewed, and improved.

Table 15.4 presents a summary of factors and processes for collaborative team-based care.

Table 15.4 Factors and Processes for Collaborative Team-Based Care

<table>
<thead>
<tr>
<th>Who</th>
<th>Team Tasks</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Physicians, medical assistants, behavioral health providers | • What are their roles/responsibilities in the screening process?  
• What is the necessary training required for screening tools/assessment?  
• How do team members communicate about screening results and follow-up?  
Huddles  
EMR  
• Who will decide how and when a follow-up assessment is needed? | • In a team meeting, PCP, MA, and behavioral health provider discuss who administers a tool, who scores it, and who provides feedback on results.  
• Team plans for how information will be shared during clinic session (flag, face-to-face).  
• Team discusses how follow-up plans will be managed. |

<table>
<thead>
<tr>
<th>How</th>
<th>Clinical Process</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Clinical workflow and protocol | • What are the evidence-based screening tools to identify your target population?  
• Which follow-up assessment tool will be used? Who determines this?  
• How will follow-up brief treatment sessions be handled? | • Target population: adults and adolescents.  
• Adult patients who are establishing care or annual exams are administered screeners: AUDIT and DAST 10.  
• Physicians will provide education and brief intervention for low-risk patients.  
• Positive screens in moderate range will have a warm handoff with psychology provider and scheduled a follow-up diagnostic session.  
• Patient will high levels of dependence will be referred to specialty programs. |
Table 15.4 (Continued)

<table>
<thead>
<tr>
<th>What</th>
<th>Support System Tasks</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Financial system, QI, and effectiveness measurements | • How are screening and assessment billed?  
• How is the data entered and reported in EMR system? What processes are in place for routinely collecting, analyzing, and using screening data to improve care, quality, and effectiveness? | • Billing for screening: 96110 with .59 modifier  
• Bill health and behavior codes (CPT 95801) when behavioral health provider meets with patients with a positive screening and assesses patient more fully.  
• Results entered into one field in EMR for QI data. |

STRENGTHS AND LIMITATIONS OF SCREENING AND ASSESSMENT FOR SUBSTANCE USE IN PRIMARY CARE SETTINGS

The strength of screening and assessment for substance misuse in primary care has been well documented. It truly is an important avenue for preventative and primary care interventions that are cost-effective for quality improvements in health outcomes and public health. Even though this process has demonstrated strong evidence for 20 years, the implementation of this protocol as a standard of care is lagging in primary care settings. There are several barriers including staff participation and engagement, time constraints of PCPs, limited number of behavioral health providers in primary care settings, and the short-term financial cost for staffing (Barbosa, Cowell, Bray, & Aldridge, 2015; Barbosa, Cowell, Landwehr, Dowd, & Bray, 2016). In addition, patients who suffer with substance misuse need significant resources that are often in short supply, especially in underserved or rural areas. These patients often struggle with health literacy and have had significant social determinants of health; yet, our system of care addresses these needs in crisis emergency care settings. Most healthcare plans have not yet made this screening an important part of their coverage, which creates another barrier to implementation (Horgan, Garnick, Merrick, & Hoyt, 2007). However, with healthcare reform and a new focus on integrated behavioral health in primary care and preventative services, there are innovative opportunities to advance this model of care in the future.

CONCLUSION

SUDs are important concerns for the primary care setting. The use of the AUDIT, DAST, and CRAFFT screeners are critical for early identification of patients who are at risk for substance use disorder. Primary care providers, psychologists, and other medical staff should work to identify a sustainable workflow within the primary care setting to identify those at risk for SUDs and establish referral and follow-up processes for a diagnostic evaluation and brief treatments and links to specialty care. Diagnostic assessments such as PRISM, ASI, or SAOM include comprehensive tools that are aligned with DSM criteria. Based on the results on severity and frequency of substance use, providers can identify appropriate follow-up care including referrals and resources for those who need inpatient or intensive outpatient treatment programs. With more integrated screening and assessment in primary care settings, primary care providers can successfully implement treatment options such as buprenorphine...
and naltrexone. With the growing interest in the provision of these medications in primary care for early intervention, it is important to anticipate the increased need for psychology providers to provide assessments, brief treatments, and monitoring of patients in these settings.

REFERENCES


In the present chapter, the important topic of the assessment of pain in primary care settings will be discussed. Indeed, complaints of pain account for almost 80% of all physician visits (Gatchel, 2005). As a result, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO; 2000) now makes it a requirement that physicians consider pain as a *fifth vital sign* (along with pulse, blood pressure, core temperature, and respiration) when routinely evaluating patients. A Numeric Rating Scale (NRS) was suggested for the purpose, ranging from 0 (No Pain) to 10 (Worst Possible Pain) for the patient’s current pain level.

Other rating systems will be reviewed in this chapter. Also, many of the psychosocial and personality characteristics often found in pain patients, and how they can be reliably assessed, will be included. This will be done under the general umbrella of the *biopsychosocial model of pain*, which is now the most heuristic approach to the comprehensive understanding of pain, as well as its effective assessment and treatment (e.g., Gatchel, McGeary, McGeary, & Lippe, 2014; Turk & Monarch, 2002). Examples of the most psychometrically sound assessment tools and what constructs they measure will be provided.

**THE MYTH OF THE “PAIN-PRONE PERSONALITY TYPE”**

At the outset, it should be highlighted that, in numerous publications, it has been consistently concluded that there is *no one consistent pain-prone personality syndrome* (e.g., see Gatchel & Dersh, 2002; Gatchel & Weisberg, 2000; Gatchel, Landers, Roberts, & Hulla, in press). That is not to say that a patient’s personality and psychosocial characteristics will often affect the nature of the acute and chronic pain, such as its severity and required treatment (e.g., Cano-García, Rodríguez-Franco, & López-Jiménez, 2013; Gatchel & Weisberg, 2000). The initial assumption that there may be a specific *pain-prone personality* was based on work by Sternbach (Sternbach, 1974), who originally proposed that three clinical profiles of the Minnesota Multiphasic Personality Inventory (MMPI) were useful for predicting the treatment outcomes of patients suffering from pain syndromes. However, equivocal results were subsequently reported in many other investigations (cf. Gatchel & Weisberg, 2000), and Sternbach’s profile was found to be less prevalent in patients with chronic spinal pain (Gatchel, Mayer, & Eddington, 2006). These latter authors actually elucidated an MMPI “disability profile,” which was comparable to the “floating profile” initially reported in the psychiatric literature (defined as elevations of four or more MMPI clinical scales [with *T*-scores greater than 65]). This floating profile is typically found in patients with borderline personality disorders. The only research with a pain population was that of Gatchel et al. (2006), who revealed that it was associated with high levels of psychopathology in a chronic spine pain population.

Rather than assuming the existence of a single “pain-prone personality,” it is best to consider each individual in terms of his/her specific personality/psychosocial characteristics. This is in keeping with
the biopsychosocial model of pain, to be discussed in the next section. Many of the characteristics that have been found useful in evaluating pain patients are listed in Table 16.1. It is important to take into account the possible presence of such characteristics in order to develop the best biopsychosocial assessment-treatment program for patients.

**THE BIOPSYCHOSOCIAL APPROACH TO PAIN: IMPLICATIONS FOR ASSESSMENT**

It has been repeatedly demonstrated that the biopsychosocial model is the most heuristic and comprehensive approach to the assessment and treatment of pain (e.g., Gatchel, 2005; Gatchel, Peng, Peters, Fuchs, & Turk, 2007; Turk & Monarch, 2002; Zale & Ditre, 2015). This model conceptualizes pain as a dynamic and complex interaction among biological/physical, psychological, and social factors that often results in, or maintains, the pain. As noted by Gatchel (2005):

> It cannot be broken down into distinct, independent psychosocial or physical components. Each individual also experiences pain uniquely. The complexity of pain is especially evident when it persists over time, as a range of psychological, social, and economic factors can interact with physical pathology to modulate a patient’s report of pain and subsequent disability. The model uses physiologic, biologic, cognitive, emotional, behavioral, and social factors, as well as their interplay, when explaining a patient’s report of pain.

(p. 23)

This also explains why there are frequent individual differences in the experience and expression of pain symptomatology. In the present chapter, the review will focus on one important component of this biopsychosocial approach: the psychosocial assessment. It should also be noted that this biopsychosocial model has stimulated the development of therapeutic- and cost-effective interdisciplinary pain management programs (e.g., Gatchel & Okifuji, 2006; Gatchel et al., 2014). Again, self-reported patient characteristics are important in tailoring such programs to the specific needs of patients.

**A Multidimensional Approach to Assessment: Important Clinical Constructs to Be Considered**

In this section of the chapter, a number of constructs (such as coping style) will be reviewed. A construct is an abstract conceptualization of some specific attribute; it often is not directly observable. However, constructs are important because they help one predict behavior, and may be related to

---

**Table 16.1: Psychosocial and Personality Characteristics Often Found in Patients With Pain**

- The MMPI “Floating Profile”
- Personality disorders (especially borderline personality disorder)
- Pain catastrophizing
- Pain-related fear avoidance
- Inadequate coping skills
- Low perceived control
- Comorbid mental and physical health disorders.
other constructs. In order for a construct to be useful, it must be proven to be valid. Gatchel and Mears (1982) earlier provided a simple definition of validity:

A test is valid if it measures what we say it measures. A valid IQ test should measure differences in intelligence. But how do we know if a test is valid? For intelligence, we might take IQ scores and correlate them with some other set of scores that is presumably related to intelligence (e.g., grades in school).

Because of the importance of constructs in psychology, and in general, the American Psychological Association convened a meeting of psychologists in 1954, and this committee was the first to introduce the concept of construct validity (Cronbach & Meehl, 1955). It is beyond the scope of the present chapter to discuss the issues of constructs, construct validity, and nomological networks (i.e., the overall pattern of relationship constructs and their operational definitions).

**Coping Style**

In terms of the construct of coping style, Taylor (2015) recently delineated it as the degree to which one deals with a stressful event, such as pain, in a particular way. Taylor (2015) went on to delineate a number of broad styles summarized below:

- **Approach coping style** (i.e., directly confronting the stressor, or being more vigilant of it);
- **Avoidance coping style** (i.e., avoiding or minimizing the threat of the stressor);
- **Problem-focused coping style** (i.e., doing something constructive about the stressor or threat, such as seeking help from others or taking direct action oneself);
- **Emotion-focused coping style** (i.e., learning about possible methods to regulate/manage emotions that are precipitated by a stressor, such as through counseling/reading information);
- **Emotional approach coping** (i.e., directly work on the best method to directly deal with a stressor, such as obtaining better stress management skills);
- **Proactive coping** (i.e., anticipating the potential occurrence of a stressor or threat, and then acting in advance to either prevent or reduce its impact).

As appropriately noted earlier by DeGood (2000), the term coping style is a construct often seen in the pain literature, and it can be inferred from a variety of sources. Some of these include standardized self-report questionnaires evaluating different types of coping (e.g., adaptive vs. dysfunctional coping, cognitive and behavioral coping styles), as well as the observation of the emotional state and physical functioning of an individual experiencing pain. Of course, construct validity will have to ultimately be determined. DeGood (2000) goes on to review many of the pain-coping style instruments that have been developed and commonly used. Many of such measures of this coping styles construct (such as fear avoidance and catastrophizing) will be reviewed next.

**Catastrophizing**

It has long been recognized that cognitive factors impact the experience of chronic pain, and influence its course. Turk and Rudy (1986), in an early review, delineated such factors into two broad categories:

a. Thoughts, self-statements or evaluations when in pain;
b. Beliefs, interpretations or attributions about . . . pain . . . medical conditions, and cognitive reactions or appraisals regarding the impact of pain. (p. 760)
Chief among the cognitive factors that has been the subject of empirical investigation is the construct of catastrophizing. Catastrophizing is defined as “an exaggerated negative mental set brought to bear during actual or anticipated painful experience” (Sullivan et al., 2001). Such a negative mental set has been more specifically “characterized by the tendency to magnify the threat value of the pain stimulus and to feel helpless in the context of pain, and by a relative inability to inhibit pain-related thoughts in anticipation of, during, or following a painful encounter” (Quartana, Campbell, & Edwards, 2009). Thus, catastrophizing has the potential to magnify the impact of pain, not only its subjective experience but also its effects on behavior and emotion.

A number of questionnaires that aim to assess the construct of catastrophizing have shown the extent of this impact. Wertli et al. (2014), in a recent systematic review of the literature on catastrophizing as a prognostic factor for outcome in low back pain patients, focused on the three most widely utilized of these scales: The Coping Strategies Questionnaire (Rosenstiel & Keefe, 1983); the Pain-Related Self-Statements Scale (Flor, Behle, & Birbaumer, 1993); and the Pain Catastrophizing Scale (Sullivan, 2009), the latter of which will be reviewed in more detail later in this chapter. Wertli and colleagues (2014) concluded that higher scores on catastrophizing scales are associated with poorer outcomes of both conservative and surgical treatments for low back pain, in a “dose-dependent fashion.” These results hold true for all levels of pain chronicity: acute, subacute, and chronic. Thus, assessment of catastrophizing is critical both in predicting treatment outcomes and in identifying a potential target for psychosocial intervention.

**Fear Avoidance**

In the case of an acute injury, pain serves as a warning signal: it indicates that tissue damage is occurring or is about to occur. In such a case, pain elicits automatic responses, such as muscle tension and sympathetic nervous system activation, which allow for rapid responses to either avoid a worsening of the injury or to escape from the potential source of damage. As pain becomes protracted, individuals learn that certain actions and situations are associated with increases in pain—associations that can be persistent and even counterproductive. For example, after injury to the low back, individuals learn to restrict certain activities, such as lifting or twisting, that increase their pain. Such restrictions, while reducing pain after an initial injury, can continue even after the injury has healed, leading to constrained behavior that weakens spinal muscles, reduces flexibility, and sets the stage for general “deconditioning” (see Mayer et al., 1987). It is now widely believed that fear avoidance (FA: fear of pain, leading to avoidance of movements and activities) is a central component of chronic pain syndromes (Vlaeyen & Linton, 2000; Vlaeyen & Linton, 2012).

Traditionally, two major questionnaires have been utilized to assess FA. The Fear Avoidance Beliefs Questionnaire (FABQ), which examines FA separately for work and physical activities, is by far the most extensively studied. The Tampa Scale of Kinesiophobia (Kori, Miller, & Todd, 1990; Miller, Kori, & Todd, 1991) also aims to assess pain-related fear of activity or reinjury in low back pain patients. Wertli et al. (2014) reviewed 78 studies examining FA beliefs utilizing either the FABQ or Tampa scale, including 23 randomized controlled trials. The authors concluded that FA beliefs were prognostic for continued pain and disability in a subacute pain population, but not a chronic pain population. Moreover, there is moderate evidence that higher FA scores, particularly on the FABQ Work scale, are associated with poorer return to work rates (cf., Heymans et al., 2007; Turner et al., 2006). The FABQ and Tampa scales were found to be equivalent in determining prognosis. While FA beliefs appear to be an important component of the chronic pain experience, at this point questionnaires that assess FA beliefs are most effective in assessing populations who have pain with a duration of 6 weeks to 3 months, and in predicting vocational disability. However, a new questionnaire has been developed...
by Gatchel and colleagues (Gatchel, Neblett, Kishino, & Ray, 2016; Neblett, Mayer, Hartzell, Williams, & Gatchel, 2015) which can be effectively used with chronic pain patients—The Fear Avoidance Components Scale (FACS). The FACS will be reviewed in the next major section of this chapter.

Comorbid Mental and Physical Health Disorders

The International Association for the Study of Pain (IASP Task Force on Taxonomy, 1994) defines pain as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage” (p. 210). As one might expect, the prevalence of mental health disorders is high among individuals whose lives have been impacted by pain, especially if it becomes chronic in nature. Not only is pain an unpleasant sensory experience, but it impacts one’s ability to engage effectively in work, recreational activities, and interpersonal relationships (Gatchel, 2005). Assessing the prevalence of comorbid mental health difficulties is a challenging undertaking because the effect of chronic pain can overlap with signs and symptoms of mental health disorders. For instance, individuals who are depressed may have difficulties with initiating or maintaining sleep, and individuals with chronic pain, who are not depressed, may report similar issues with sleep. In fact, pain and depression impact overlapping brain circuitry (Robinson et al., 2009).

In an attempt to study the relationship between mental health disorders and chronic pain, Polatin and colleagues (1993) administered the Structured Clinical Interview for DSM-III-R to 200 participants of a functional restoration, interdisciplinary pain program. When somatoform pain disorder was not considered, 77% of participants met criteria for a lifetime mental health diagnosis and 59% met current diagnostic criteria. High rates of depression, anxiety, and substance use disorders were found among this population. In general, an estimated 34%–54% of individuals with chronic pain meet criteria for major depressive disorder (Fishbain, Cutler, Rosomoff, & Rosomoff, 1997; Polatin et al., 1993) and 35% meet criteria for an anxiety disorder (McWilliams, Cox, & Enns, 2003). In one of the largest analyses to date, Gureje, Simon, and Von Korff (2001) examined World Health Organization data from 3,197 individuals receiving treatment at 15 different primary care sites across 14 countries. Individuals who reported persistent pain had higher rates of depressive and anxiety disorders than those who did not report persistent pain.

The relationship between pain and mental health disorders remains complex, as it is difficult to determine if depression and anxiety are solely a consequence of pain, a risk factor for chronic pain, or related to underlying factors (e.g., catastrophizing) that place individuals at risk for both pain and emotional disturbance. Furthermore, it is likely that pain intensity, chronicity, and the impact of pain on social functioning likely serve as moderating or mediating variables and different chronic pain populations are likely to have differential rates of psychopathology.

Personality Disorders

According to the DSM-5 (American Psychiatric Association, 2013), the general criteria for personality disorders includes: “An enduring pattern of inner experience and behavior that deviates markedly from the expectations of the individual’s culture” (p. 646). Furthermore, these difficulties impact at least two areas of functioning (i.e., thinking, affect, relationships, and/or control of impulses). As mentioned previously, clinical research does not support the concept of a pain-prone personality, but some evidence supports that the chronic pain population has higher rates of personality disorders than the general population (Fishbain, Goldberg, Meagher, Steele, & Rosomoff, 1986; Weisberg & Keefe, 1997). For instance, Polatin and colleagues (Polatin et al., 1993) reported that 51% of individuals with chronic pain in their sample met criteria for a personality disorder. By definition, individuals
with personality disorders are at risk for limited coping and low stress tolerance. After an injury or an illness, one might expect that individuals with a personality disorder will have difficulty adhering to self-care regimens that could assist in their recovery.

Much has been written about individuals who fall into the Cluster B category of personality disorders, and who are described as “dramatic, emotional, or erratic” (American Psychiatric Association, 2013, p. 646). Most practitioners are familiar with the diagnosis of borderline personality disorder, which has become an opprobrium and stigma. However, individuals who truly meet criteria for this disorder can challenge practitioners with intense and often unreasonable, angry, and manipulative-seeming behavior. However, when borderline personality disorder is conceptualized from a developmental perspective, seemingly contradictory and upsetting behavior from patients may be more understandable. Attachment theorists, such as Bowlby, Ainsworth, and Main, have explored the way an individual’s understanding of themselves and others are shaped by early attachment relationship with caregivers (Karen, 1998). Most individuals who meet criteria for borderline personality disorder would be characterized as having a disorganized attachment style (Fonagy, 2000). These individuals have developmental histories of being fearful of attachment figures that contribute to creating an approach-avoidance conflict. As such, cognitive structures to self-soothe, take the perspective of others, and maintain a sense of self-worth are not fully developed. Affective dysregulation can be seen not only as a result of impaired development, but also as a factor that hinders the growth of self-soothing behavior. As such, much of the manipulative-seeming behavior is a coping strategy, albeit a poor one, to reduce uncertainty and regulate emotions (Sharp & Fonagy, 2015). For a more complete understanding of this conceptualization, please refer to the work of Fonagy, Bateman, and others regarding the rationale for mentalization-based approaches for the treatment of borderline personality disorder.

METHODS OF ASSESSMENT

Approaching assessment of individuals with pain is best performed within a theoretical framework that guides the selection of instruments and provides a structure, or outline, to identify the most relevant factors to assess. As discussed, the biopsychosocial approach is an empirically derived, theoretical model that is ideal for providing the structure needed for effective assessment of individuals with pain (Gatchel, 2005; Gatchel et al., 2007; Turk & Monarch, 2002; Zale & Ditre, 2015). Under the “Bio” portion of this theoretical model, laboratory and imaging studies as well as self-report measures, or testing of physical limitations resulting from pain provide information to be used for treatment tailoring and monitoring of progress. Psychological factors, such as catastrophizing, depression, and anxiety, are both a consequence of pain and a barrier to recovery (Gatchel, 2005), and assessment of these areas, even if simply screening for depression and anxiety, can guide referrals and incorporation of other resources. Lastly, social factors can include gathering additional information from collateral resources as well as assessing the impact of pain on social functioning and satisfaction. Although each setting in which individuals with chronic pain may be seen may require slightly different batteries, developing a battery within the biopsychosocial approach provides the best structure to identifying the most relevant factors for effective treatment.

Examples of Self-Report Tests

Visual Analogue Scale
A common tool used to measure pain intensity is the Visual Analog Scale (VAS). This is similar to the NRS discussed at the outset of this chapter. It is a single-item scale that asks respondents to a mark the intensity of their pain on a 100 mm line, with one side indicating least possible pain and the other
indicating worst possible pain. Some researchers have suggested that a mark at a location greater than 5 mm indicates mild pain, 45 mm or more indicates moderate pain, and 75 mm or more indicates severe pain (Hawker, Mian, Kendzerska, & French, 2011); others believe that a mark at a location greater than 26 mm may indicate moderate pain, while a mark above 54 mm may indicate severe pain (Collins, Moore, & McQuay, 1997). This scale converges with other measures of pain including verbal rating scales (Ohnhaus & Adler, 1975). Additionally, the VAS has been adapted for use on electronic systems including an iPad, and scores on electronic versions have shown to be consistent with paper administration of the VAS (Bird et al., 2016; Jamison et al., 2002). Some studies have found that is not effective in detecting a change in pain over time and, thus, should not be used as an indicator of treatment efficacy (Fosnocht, Chapman, Swanson, & Donaldson, 2005).

Quantified Pain Drawings
As noted by Hartzell, Liegey-Dougall, Kishino and Gatchel (in press), pain drawings have been historically used to evaluate the location of pain, as well as the extent of a patient’s pain (Ingham & Portenoy, 1998). They are also excellent in assessing the location of pain (Buenaver & Edwards, 2007; Karoly & Jensen, 1987). In addition, they are typically brief to administer, and the scoring is relatively straightforward (Vucetic, Maattanen, & Svensson, 1995). Moreover, they can readily be used with children or the elderly, and do not require a great command of language (Vucetic et al., 1995).

The Quantified Pain Drawing is a standard Pain Drawing, which consists of the anterior and posterior views of the body. As noted by Gatchel (Gatchel, 2006), there are three commonly used scoring methods: the penalty-point method; the body-region method; and the grid-scoring method. The body-region method divides the anterior and posterior outlines in the body figure into 45 anatomical regions. Therefore, scoring can be accomplished on a region-specific basis. In addition, Margolis and colleagues (Margolis, Tait, & Krause, 1986) had developed a method by which a calculated weighted score can be derived to yield a percentage-of-the-body in pain. The test-retest reliability of this method is .85, and the rate of agreement between scorers was 88.27%.

MMPI-2-RF
As noted earlier, the original body of research utilizing the MMPI and MMPI-2 focused primarily on identifying a “pain-prone personality.” More recent research, utilizing a restructured form of this test—the MMPI-2-RF—has taken an actuarial approach to the assessment of chronic pain patients and candidates for both spine surgery and spinal cord stimulation. The MMPI-2-RF has several advantages over earlier forms of the MMPI family. It is 40% shorter and, most importantly, has improved psychometric properties. As noted by Ben-Porath and Tellegen (2008/2011), there was considerable overlap in the item content of the MMPI clinical scales, which clouded the interpretation of pain-related MMPI research. Ben-Porath (2012) described these overlapping items as assessing a “general” psychological factor, which he termed “demoralization.” The MMPI-2-RF removed these overlapping items and created a new Demoralization Scale (RCd). In so doing, the MMPI-2-RF also improved the discriminant validity of other clinical scales.

Research on the MMPI-2-RF has established mean clinical profiles for spine surgery and spinal cord stimulation candidates (Block, Ben-Porath, & Burchett, 2011), and for chronic pain patients (Tarescavage, Scheman, & Ben-Porath, 2015). Poorer surgical and treatment outcomes are significantly correlated with elevations on the RCd scale, and other restructured clinical scales, including Somatic Complaints (RC1), Low Positive Emotion (RC2), and Dysfunctional Negative Emotions (RC7; (Block, Ben-Porath, & Marek, 2013; Block, Ben-Porath, Marek, & Ohnmeiss, 2014; Block, Marek, Ben-Porath, & Kukal, 2015; Marek, Block, & Ben-Porath, 2015). Several MMPI-2-RF Specific Problem scales have also been shown to predict poorer treatment outcomes, including Malaise (MLS), Helplessness (HLP), Self-Doubt (SFD), Stress/Worry (STW), Anxiety (AXY), and Family Problems (FML), as has the
personality psychopathology scale (Negative Emotionality/Neuroticism-Revised [NEGE-r]). Given that the MMPI family of tests is historically the most widely utilized tool in the assessment of pain patients, and that the MMPI-2-RF has a substantial, growing research base, this restructured form of the MMPI provides an invaluable resource to assess patient characteristics, determine the likelihood of obtaining good treatment outcomes, and target areas for intervention. The use of the MMPI-2-RF is discussed in Chapter 23 of this book.

**Patient Health Questionnaire (PHQ)**
The PHQ is a self-administered questionnaire, developed to screen for mental health disorders (Spitzer, Kroenke, & Williams, 1999). The PHQ consists of five modules: depression, anxiety, somatoform disorders, alcohol, and eating disorders. All of PHQ modules have a dichotomous diagnosis component. The dichotomous diagnosis components determine a provisional diagnosis. Some of the modules (depression, anxiety, and somatization) have a total score that represents severity. The highest scores for depression (nine items) and anxiety (seven items) modules are 27 and 21, respectively, based on the 0–3 score range. The highest score for somatization (15 items) is 30, based on the 0–2 score range. In its original validation study, conducted on outpatients in a primary care setting, PHQ-derived diagnoses showed good agreement with the diagnoses of mental health professionals, with sensitivity of 75% and specificity of 90% (Spitzer et al., 1999). The PHQ demonstrates a good agreement with the older Structured Clinical Interview for DSM-IV-Axis I (SCID), with 90% overall accuracy, 84% sensitivity, and 97% specificity (Navinés et al., 2012). The PHQ-9 depression screener is discussed at length in Chapter 9 of this book.

**SF-36 Health Survey**
The Medical Outcomes Study 36-item Short Form questionnaire was designed to measure eight domains related to physical and mental health and their impact on functioning in a clinical and/or research setting (Ware & Sherbourne, 1992). These eight domains are: physical functioning; social functioning; role limitations due to physical problems; role limitations due to emotional problems; bodily pain; mental health; vitality; and general health perceptions. Studies have found that its validity and reliability vary by the sample on which the measure is tested but, overall, most studies find it to be reliable and valid (McPherson & Martin, 2013). For instance, it effectively distinguished between people with minor medical issues, people with major medical issues, people with only psychiatric issues, and people with both psychiatric and serious medical issues in a diverse sample of people seeking treatment for medical complaints (McHorney, Ware, & Rac-zek, 1993).

Subsequent to the SF-36’s initial development, a revised version—the SF-36v2—was developed. The SF-36v2 incorporates many improvements over the original version and has much more recent norms, which were the product of a large, nationwide renorming project conducted in 2009 (see Maruish, 2011). In addition, a 12-item abbreviated version of the SF-36v2—the SF-12v2—is also available (Maruish, 2012). Both surveys are discussed at length in Chapter 20 of this book.

**Pain Disability Questionnaire**
The Pain Disability Questionnaire (PDQ) was developed to assess pain and functioning across individuals with musculoskeletal disorders who have developed chronic pain (Anagnostis, Gatchel, & Mayer, 2004). Unlike many other measures of disability, the PDQ allows for the assessment of outcomes across a variety of pain conditions, and provides a range of scores from 0 to 150. The psychometric properties of the PDQ are strong with a test-retest reliability coefficients of .94 to .98, and a Cronbach’s coefficient alpha of .96 (reflecting a high degree of internal consistency). Convergent validity
Assessment of Pain in Primary Care

Some benefits of this assessment over other measures of pain include that this measure has excellent psychometric properties, as explained earlier; assesses a multitude of chronic painful musculoskeletal disorders as opposed to focusing on one area of pain, for instance low back pain; contains two factors that measure functional and psychosocial aspects of patients’ lives and thus can be additive in understanding the effect of these factors on pain; and is capable of capturing a change in pain and function over time (Gatchel, Mayer, & Theodore, 2006). Also, the PDQ has been translated to Brazilian Portuguese and retained its good psychometric properties (Giordano, Alexandre, Rodrigues, & Coluci, 2012). On the other hand, this measure is relatively new and has not been applied to research beyond chronic musculoskeletal pain treated in a multidisciplinary treatment setting and, thus, requires further testing in other setting such as primary care physiotherapy treatment facilities (Anagnostis, Gatchel, & Mayer, 2004).

Beck Depression Inventory-II
The assessment of symptoms of depression is a critical component of understanding the impact of pain, as well as its potentially aggravating and maintaining factors. Although many measures exist, the most commonly used is the Beck Depression Inventory-II (BDI-II). The BDI-II has demonstrated strong psychometric properties, with high internal consistency (i.e., Cronbach’s coefficient α = .91; (Beck, Steer, Ball, & Ranieri, 1996). Furthermore, when the BDI-II was evaluated within a primary care medical setting, high convergent validity was found when correlated with the Short-Form General Health Survey (SF-20), and a receiver operating curve analysis provided evidence of criterion validity (i.e., the BDI-II was able to predict the presence of a major depressive disorder) (Arnau, Meagher, Norris, & Bramson, 2001). A discussion of the BDI-II is include in Chapter 9 of this book.

Pain Catastrophizing Scale
Cognitive errors related to chronic pain can often be classified as catastrophizing. Specifically, individuals with chronic pain often have beliefs of unwarranted negativity given the current facts. Sullivan, Bishop and Pivik (1995) developed the Pain Catastrophizing Scale (PCS) to assess just such issues. The PCS consists of 13 items, and assesses three forms of catastrophizing: rumination, magnification, and helplessness. The PCS has also demonstrated strong psychometric properties (e.g., Cronbach’s coefficient alpha of .87) (Sullivan, 2009). Furthermore, construct validity has been established in several ways, including assessing the frequencies of catastrophizing thoughts during a cold pressor test (Sullivan et al., 1995).

Fear-Avoidance Components Scale (FACS)
As noted by Neblett and colleagues (Neblett et al., 2015), pain-related fear avoidance is a common problem that affects patients with painful medical conditions. This fear avoidance can lead to significant avoidance of activities of daily living and work that, in turn, can lead to a decrease in physical functioning, depression, and even disability. Neblett and colleagues developed the FACS for measuring this important construct. It consists of 20 items, each scored from 0 to 5, with a total possible score of 100. Moreover, the following severity levels have been suggested for clinical use: subclinical (0–20); mild (21–40); moderate (41–60); severe (61–80); and extreme (81–100). Examples of FACS items are: “I try to avoid activities and movements that make my pain worse”; “My painful medical condition puts me at risk for future injuries (or reinjuries) for the rest of my life.” The FACS has been found to be related to decreased physical performance tasks (e.g., lifting), as well as...
psychosocial variable (e.g., pain severity, perceived disability, and depression). The FACS has good psychometric properties, with high internal consistency (Cronbach’s $\alpha = .92$) and high test-retest reliability ($r = .94$).

**Pain Medication Questionnaire (PMQ)**
The PMQ detects risk of opioid medication misuse, and serves as one means of tracking adherence to opioid medication (Adams et al., 2004). The PMQ has strong psychometric properties, with a Cronbach’s coefficient alpha of .73, and test-retest reliability of .85. In addition, Holmes and colleagues (2006) found that individuals who scored as “high risk” were more likely to require early refills and to prematurely terminate treatment. The PMQ consists of 26 items and takes less than 10 minutes to complete (Holmes et al., 2006).

**Behavioral Observation**
Wilbert Fordyce was one of the first psychologists to apply psychological/behavioral principles to the treatment of individuals with chronic pain and viewed pain as a form of interpersonal communication (Fordyce, 1976). Familiar to all physicians is the individual with severe injuries who appears stoic, or an individual with a slight injury but who grimaces, moans, and walks in a guarded manner. Treatment for individuals with chronic pain was previously undertaken within a behavioral framework. As such, operant conditioning principles were utilized to aid in the diminution of less than adaptive behavior, such as postures related to guarded or bracing or low levels of activity. Self-reports of the sensory, cognitive, and affective components of pain were seen as relevant but inadequate targets of change from a strictly behavioral perspective. Rather, the reduction of pain behavior was a far more objective target within this framework. Currently, conceptualizing the assessment of pain from a multidimensional perspective is understood to provide the best results; however, behavioral observations provided an important supplement to self-report data.

Although most physicians have training in identifying problematic pain behavior (i.e., presentations that are inconsistent with physical findings or considered “nonorganic”; Main & Waddell, 1998), these methods have mostly, but not exclusively, been informal or based on a wealth of clinical experience. Fordyce (1976) began to apply more formal methods and discussed the importance of having patients track pain, activity level, and medication usage in a diary. He also developed interviews designed to elicit information about behaviors that could be modified through operant conditioning principles. In more recent years, other behavioral assessment protocols were developed. Keefe and Smith (2002) described the essential elements of structured behavioral observations:

> First, the patient is provided with a rationale and instructions for the observation session . . . Second, each protocol incorporates a series of tasks designed to elicit pain behavior . . . Third, each protocol includes some method for the recording and coding of pain behavior . . . Finally, each protocol yields quantifiable pain behavior.

(p. 119)

One of the first structured protocols was developed by Keefe and Block (1982) and measured pain behaviors such as guarding, bracing, and grimacing during a series of standardized task. Their protocol demonstrated high interrater reliability and established construct validity through high correlations with pain ratings (Keefe & Block, 1982). Other measures have been developed and have maintained fidelity to the principles outlined earlier (Keefe & Smith, 2002).
THE PATIENT-REPORTED OUTCOMES MEASUREMENT INFORMATION SYSTEM (PROMIS) FOR ASSESSING PATIENTS WITH PAIN

Over 10 years ago, the National Institutes of Health (NIH) began a series of meetings to evaluate and develop strategies to overcome obstacles to the advancement of scientific research (Cella et al., 2007). The goal of the PROMIS initiative was to develop measures, using state-of-the-art psychometric techniques, that could be used across clinical trials, as well as to improve patient quality of care (Broderick, DeWitt, Rothrock, Crane, & Forrest, 2013). Although paper-and-pencil versions of the PROMIS measures have been developed, the strength of the PROMIS measures involves the use of item response theory (IRT) and computerized adaptive testing (CAT). Simply put, the PROMIS measures ask targeted questions based on previous responses from the research participant or patient as well as normative data from a similar population to select follow-up items. This approach decreases patient burden and provides more precise assessment of the constructs of interest. Both clinician and researchers can gain free access to PROMIS measures via NIH’s assessment center, where patients or participants can complete these measures online (Broderick et al., 2013).

Scales Included on the PROMIS Platform

PROMIS domains have been developed for the adult population, as well as for children and their parents. This discussion will be limited to the adult population. Major domains focus on the following areas: (1) global health, (2) physical health, (3) mental health, and (4) social health. Of these domains, global health is a measure in and of itself. The global health measure assesses overall health and is intended to provide an overview based on reports of physical functioning, pain, fatigue, emotional distress, and social health. Within the general domain of physical health, measures exist that assess physical function, pain intensity, pain interference, fatigue, and sleep disturbance. Additional domains in this category include pain behavior, pain quality, sleep-related impairment, sexual function, gastrointestinal symptoms, and dyspnea (National Institutes of Health, 2016).

Within the general domain of mental health, measures include depression and anxiety. Furthermore, additional measures assess anger, cognitive function, alcohol use consequences and expectancies, smoking, substance abuse, psychosocial illness impact, and self-efficacy. Finally, within the social health domain, the primary measure focuses upon ability to participate in social roles and activities. Additional measures assess satisfaction with social roles and activities, social support, social isolation, and companionship (National Institutes of Health, 2016).

Comparisons Made to Updated Norms

The NIH PROMIS measures use T-scores, with 50 points representing the normative average and 10 points representing one standard deviation. PROMIS measures also provide comparisons to the US population for the majority of the domains. However, some domains developed normative data from patient populations as well (National Institutes of Health, 2016).

Feedback for Healthcare Professionals and Patients

As previously mentioned, as part of the NIH’s PROMIS initiative a free website was developed to allow for the administration and scoring of PROMIS domains once a researcher or clinician has
registered. Paper-and-pencil short forms are also available for these measures to address potential confidentiality and privacy concerns. Within the NIH’s scoring website assessment center, an individual graphic patient health report can be generated that compares the individual’s scores to that of the normative sample, whether the general US population or a clinical sample (NIH, 2016).

**ASSESSMENT IN PRIMARY CARE SETTINGS**

Conducting assessment of the relevant biopsychosocial factors needed to make appropriate referrals and guide treatment for individuals within primary care settings poses a unique set of challenges. First, individuals with pain may be understandably suspicious of measures that assess psychological factors associated with depression and anxiety. Not only is there a stigma associated with mental health diagnoses (Corrigan, 2004), but an individual with pain may fear that their physician believes their pain to be “psychological” in nature. These issues can be minimized when patients are explained that depression and anxiety are frequently natural consequences to pain and its impact on sleep, functioning, and quality of life. Education concerning the biopsychosocial approach is also useful at this stage. The biopsychosocial approach assumes that an injury or illness initiated the pain process and that factors such as stress can simply aggravate or maintain the pain (Turk & Gatchel, 2002). As can be seen, this approach moves beyond the inadequate and overly simplistic **physical** or **psychological** distinction, and in our experience, is understood intuitively by our patients.

Second, a growing number of providers are conducting assessments prior to the initiation of opioids as one means to impact positively the opioid pain medication epidemic (Chou et al., 2015). Individuals with pain may have a mistaken belief that opioids are “real” pain medication and all other non-opioids are not truly pain medication. Education into the limited efficacy of opioid medications for extended periods of time as well as the risk associated with opioids that go beyond concerns with opioid use disorder, such as endocrinopathy, increased risk of fractures, and cardiac events, can be helpful to address this concern in patients with pain. In our own practice, we have observed that a discussion of opioid induced hyperalgesia (OIH) has proven to be a useful intervention for many who are ambivalent about their opioid use.

Third, the time demands of a primary care practice are great. Ideally, education is provided by the physician, but this is not always practical. Delivery of information via written materials and conversations with nursing staff serve as alternative methods. In general, explaining the rationale for an assessment and how it can benefit the patient improves engagement in the assessment and treatment process (Smith, Finn, Swain, & Handler, 2010).

In general, within primary care settings, we recommend a stepwise approach detailed in the following section. All individuals within a primary care setting who present with pain should be administered brief screening measures of depression and anxiety. Clinically significant symptoms of depression, including suicidal ideation, and anxiety should trigger a more thorough evaluation. An individual with pain that has lasted longer than three months should also be evaluated fully even if depressive or anxious symptoms are not present. It is well established that chronic pain can begin to impact more areas of one’s life the longer it lasts and places one at risk for mental health disorders (Gatchel, 2005). Early identification of potentially problematic coping strategies is one way in which the impact of chronic pain on an individual’s life can be negated. Lastly, an individual who is being considered for chronic opioid therapy (COT) should also be evaluated in a comprehensive manner.
STEPWISE APPROACH TO ASSESSMENT: A CASE EXAMPLE

Primary care practices have continuous, and growing, demands on time and resources that can threaten or overwhelm even the most dedicated and organized practitioner. We recommend a stepwise approach to the assessment of individuals with pain. Thus, each patient, whether presenting with acute or chronic pain, should be initially assessed for depression and pain intensity. Individuals with depression, or who have experienced pain for longer than three months, should be considered to be sent to a pain management behavioral health specialist for a more thorough evaluation and treatment recommendations. In addition, physicians who are considering a prescription for opioids, even for short-term use, would benefit from the administration of the PMQ, or similar opioid risk screeners, prior to prescribing opioid pain medication. Below is a clinical example of this approach.

John is a 54-year-old white male who injured his back approximately four months ago. He indicated that he was working in the yard when he felt a sharp pain in his lower back and right leg. He sought treatment from his primary care physician within a month of his injury and no evidence of depression was found on the BDI-II, and his pain was rated as a 5 out of 10. Appropriate imaging and referrals were made at that time. Over the next few months, his leg pain resolved, but he continued to experience persistent low back pain. Findings from imaging revealed a bulging disc at L4-5, with no evidence of nerve impingement. A neurosurgical consultation recommended conservative care at this time. Aside from hypertension that is being managed medically, John has no other significant medical problems.

John subsequently returns for a follow-up visit. During this visit, he rated his pain as a 7 out of 10, and he complained that his sleep is poor. His ratings on the BDI-II have risen, but not to a degree that suggests depression. At this point, John should be considered for a referral to a behavioral health specialist with expertise in pain, such as a pain psychologist. Although there is not significant evidence of depression, his pain has lasted longer than three months. With the pain psychologist, a comprehensive evaluation occurs, including a psychosocial interview and the administration of the Pain Catastrophizing Scale, Fear Avoidance Scale, Pain Disability Questionnaire and MMPI-2-RF. Results reveal that John is experiencing mild depressive symptoms and difficulties with sleep. Furthermore, his pain and impaired sleeping have negatively affected his performance at work, and he is unable to engage in recreational activities, such as golf. The recommendation is made that he enter an interdisciplinary treatment program, which combines cognitive behavioral therapy, physical therapy, and group education. Upon his first visit, PROMIS measures related to pain, functioning, depression, and sleep are administered, and used to tailor a specific treatment program for him. These measures are also readministered at midtreatment in order to make adjustments to the treatment strategy if gains are not being made, and then again at the end of treatment in order to document treatment effectiveness or guide additional follow-up. If significant clinical improvement is not evidenced, then additional “booster” treatment sessions should be recommended, focusing on the specific biopsychosocial factor(s) that still need improvement.

SUMMARY AND CONCLUSIONS

The reliable and valid assessment of pain in primary care settings is especially important because complaints of pain account for 80% of all physician visits. This high prevalence now makes it a requirement that physicians consider pain as a fifth vital sign (along with pulse, blood pressure, core
temperature, and respiration) when routinely evaluating patients. Initially, a simple Numeric Rating Scale was suggested for the purpose, ranging from 0 (No Pain) to 10 (Worst Possible Pain), relating to the patient’s current pain level.

It should be kept in mind, though, that a biopsychosocial model of pain is now accepted as the most heuristic approach to the most comprehensive assessment and treatment of pain. As a consequence, a number of other interacting variables need to be taken into account besides a simple numerical rating of pain. Because it has been consistently reviewed that there is no one consistent pain-prone personality syndrome, several different psychosocial constructs have been introduced in order to develop a more complete understanding of a patient’s self-reported pain that, in turn, will help develop more effective treatments tailored to the needs of each specific patient. Many of the most important of these psychosocial constructs were reviewed (such as coping style, personality disorders, and comorbid disorders) that would be essential for a clinically relevant multidimensional approach to the assessment of pain. Under the general Coping Style umbrella, important measures such as catastrophizing and fear avoidance were presented, along with the most psychometrically sound instruments used to assess them. Moreover, a summary of many other often-used instruments/methods were reviewed.

It should also be noted that because pain is now recognized as a major national problem, the National Institutes of Health developed the PROMIS system to encourage all clinicians/researchers to utilize a common nomenclature and metric so that the comparison across clinic and research sites will become more consistent. Before its introduction, there were large disparities across sites in measures used. Consequently, equivocal findings could not be fully understood because it was like trying to “compare apples to oranges” in the measures presented for outcome-reporting purposes. The new PROMIS has provided a major advance in evidence-based pain assessment/treatment outcome reporting. A major clinical research task for the future is to determine how effective PROMIS is in the primary care setting with acute pain patients.

From a patient-centered and practical perspective, a stepwise approach to assessment decreases burden on both patient and practitioner. Individuals who present with pain should be screened for depressive symptoms, as this is highly prevalent among individuals with pain (Polatin, Kinney, Gatchel, Lillo, & Mayer, 1993). Furthermore, patients who are being considered for an opioid prescription would benefit from the administration of the PMQ or a similar measure to assess the risk of opioid misuse. Individuals whose pain has lasted for longer than three months, or have screened positive for depressive symptoms, would benefit from more comprehensive evaluations by behavioral health specialists with expertise in pain management. Lastly, individuals with chronic pain who are undergoing a course of treatment would benefit from the administration of the NIH PROMIS measures at several points over the course of their treatment to aid in the tailoring of interventions.

REFERENCES


Exposure to traumatic events, including but not limited to interpersonal violence, child abuse, motor vehicle and other accidents, and natural disasters is an unfortunately common occurrence. The resulting impact of trauma can have deleterious effects on one’s functioning and overall health as both physical health complications and mental illness comorbidities are common. Trauma contributes significantly to the cost of health care provided to victims, as well as the cost of loss of productivity by those facing the symptoms that are experienced subsequent to exposure to traumatic events. The onset of posttraumatic stress disorder (PTSD), the resulting mental health disorder frequently occurring after an adverse or catastrophic life event, is a manifestation of symptoms that can be chronic and debilitating. Overall, PTSD costs an estimated $3 billion a year in lost productivity in the United States—similar to that of major depression (Kessler, 2000). The importance of recognizing a person’s exposure to a traumatic event, assessing the impact of such events on functioning, and developing effective, evidence-based treatments to address such concerns cannot be overstated. This is especially the case in a primary care setting as it is likely the first place a person will seek help to address the resulting impact of trauma. Fortunately, in recent decades there has been significant progress in establishing psychometrically sound measures designed to recognize trauma exposure and PTSD, as well as related syndromes such as acute stress disorder (ASD; Foa & Yadin, 2011).

This chapter will provide a description of trauma, the clinical classification of associated symptoms, and a profile of trauma presentation in primary care. The significant link between exposure to traumatic events and a person’s physical health will also be presented. Broad considerations for assessing trauma in a primary care setting including several practical steps for effective screening will be discussed. Primarily, an in-depth description of some of the most useful screening measures, including their psychometric properties and strategies for administration in a real-world setting and patient population, will be presented. The chapter will conclude with several cultural considerations for screening and assessing of trauma, with the inclusion of a case study that underscores the importance of the strategies and recommendations presented.

POSTTRAUMATIC STRESS DISORDER (PTSD): CLINICAL MANIFESTATIONS AND DIAGNOSTIC CRITERIA

While the clinical manifestation and concerns of those presenting to treatment following a trauma has remained relatively constant, the diagnostic criteria and coding manuals utilized to categorize such criteria has changed significantly over time. Clinical observation and increased research literature in the area of traumatic stress has changed the conceptualization of trauma and PTSD (Weathers, Marx, Friedman, & Schnurr, 2014). With regard to the Diagnostic and Statistical Manual of Mental
Disorders, 5th edition (DSM-5; American Psychiatric Association, 2013), the diagnostic criteria were revised significantly. Changes include: (1) development of a new Trauma- and Stress-or-Related Disorders chapter; (2) removal of Criterion A2 (the person’s response involved intense fear, helplessness, or horror); (3) transition from single avoidance and numbing cluster to two separate clusters labeled avoidance and negative alterations in cognition and mood; (4) addition of three new symptoms (persistent and distorted blame or self or others, persistent negative emotional state, and reckless or destructive behavior) and reconceptualization of others; (5) addition of dissociative subtype; (6) development of separate criteria for preschool children; and (7) elaboration of the concept of indirect exposure to trauma and emphasis on the importance of connecting symptoms with the traumatic event (Weathers et al., 2014). Obviously, such changes will necessitate revision of standardized measures for assessing trauma exposures and PTSD symptoms and it is expected that numerous studies designed to demonstrate consensus of the new definition of PTSD—beyond what is described in this chapter—will be forthcoming (Weathers et al., 2014).

Stein et al. (2014) provided an overview of the development of the DSM-5 and ICD-11 (Cloitre, Garvert, Brewin, Bryant, & Maercker, 2013) criteria for PTSD. In it, they delve into the history of symptom criteria and how it has changed over time. Currently, it is likely that the mostly widely utilized diagnostic criteria utilized by primary clinicians in practice—including in the documentation of a clinical diagnosis—is based on the ICD-10 and the recently developed DSM-5. That said, it is also likely that many of the mental health clinicians currently practicing in an integrated primary care setting were trained in the DSM-IV-TR (American Psychiatric Association, 2000). It is important to recognize that diagnostic criteria are not uniform across versions of coding manuals. Clinicians should become familiar with the subtle differences that exist and develop a high degree of comfort with the diagnostic coding system utilized by their respective clinical settings.

Stein et al. (2014) detailed the between-system differences in indicators of clinical severity. For example, between the ICD-10, ICD-11, DSM-IV-TR, and DSM-5, criteria for posttraumatic stress disorder in the ICD-10 are the least strict, while criteria for the DSM-IV-TR are the most strict. Further, it is important to note that the use of any one diagnostic system may overlook many individuals. The measures described in this chapter are likely to be versatile enough to recognize symptoms based on any diagnostic system utilized.

It is vital not only to manage symptoms of PTSD, but also to develop approaches to manage sub-threshold levels of PTSD (i.e., presenting symptoms that do not meet full diagnostic criteria (Corso et al., 2009). Many patients will present to primary care with a history of trauma exposure and may not meet criteria for PTSD. Current research indicates four trajectories of symptomology after a trauma. Acute PTSD symptoms are manifested by high anxiety and distress immediately after the event. These patients will recover over time. Chronic distress is characterized as a constant high level of symptoms indefinitely. Those who do not initially experience symptoms of PTSD but later develop problematic signs of distress could be considered to have delayed onset PTSD (Hanley, deRoon-Cassini, & Brasel, 2013).

Asymptomatic patients never go on to develop characteristic symptoms of PTSD (Hanley et al., 2013). However, when a patient does not meet full criteria for PTSD, it should not be taken as an indication that follow-up assessment and treatment is not warranted. Positive screens only indicate that assessment or further evaluation is needed, while negative screens do not necessarily indicate the patient does not have symptoms that warrant follow-up (SAMHSA, 2014b). The likelihood of a patient under-reporting symptoms due to stigma, fear of a report of abuse being made to social services, fear of retribution from a perpetrator, along with several cultural considerations, can be common. At the same time, the degree to which onset of the full criteria of PTSD and its impact on
physical health can be prevented is something an integrated behavioral health clinician, in conjunction with the primary care provider, should take into account. Therefore, some type of follow-up (e.g., screening for symptoms at a later time) with any patient that has experienced trauma is typically necessitated as research has demonstrated evidence that early intervention can reduce the occurrence of the development of PTSD (Hanley et al., 2013). The measures described in this chapter help a clinician recognize the difference between trauma exposure and full symptom criteria of PTSD, both of which warrant follow-up.

**TRAUMA PRESENTATIONS IN PRIMARY CARE**

PTSD is common among primary care patients and is closely tied with functional impairment, physical health concerns and high rates of medical utilization, and mental health comorbidities (Possemato, 2011). In addition, primary care is often the first point of contact for patients experiencing PTSD because most people with common mental health disorder do not seek treatment from specialty mental health providers (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Ouimette, Wade, Prins, & Schohn, 2008). Primary care tends to be the initial point of contact for patients with PTSD, although these patients rarely identify themselves as having experienced a traumatic event in the past and often do not express current symptoms of concern related to the trauma (Cameron & Gusman, 2003).

Rates of PTSD in the civilian population and those presenting to non–Veterans Affairs (VA) primary care clinics are somewhat less clear, as so much of the research on trauma, especially for validating the measures described later in this chapter, has involved military personnel presenting to the VA. Previous research has found a prevalence rate of 8% in the general population (Kessler et al., 1995; Meredith et al., 2009) and up to 17% in primary care patients (Meredith et al., 2009). Ouimette et al. (2008) reported rates of PTSD in primary care ranging from 11% to 12% in community-based primary care settings, which is higher than the prevalence of PTSD in the general population (Kessler et al., 1995). Further, prevalence of PTSD and associated impairments in physical functioning and increased bodily pain may be as high as 25% among civilian primary care patients (Gillock, Zayfert, Hegel, & Ferguson, 2005).

According to the US Department of Veterans Affairs National Center for PTSD, experiencing trauma is not rare. About 60% of men and 50% of women experience at least one trauma in their lives, and approximately 7% of the population in the United States will develop PTSD in their lifetime. In a 12-month period, approximately eight million adults meet criteria for PTSD. About 10% of women and 4% of men develop PTSD in their lifetime. Whereas women are more likely to experience sexual assault and child sexual abuse, men are more likely to experience accidents, physical assault, combat, or disaster, or to witness death or injury (National Center for PTSD, 2016a).

Previous data from the current authors’ clinic of service indicated rates of positive screens for PTSD (as measured by the PC-PTSD) to be 12% (Auxier & Runyan, 2013). In a meaningful description of managing a military sample with PTSD in primary care, Corso and colleagues (2009) noted that avoidant coping strategies with PTSD and fear of seeking treatment decreased the likelihood that active duty members will seek treatment for mental health symptoms like those of PTSD. This is also likely the case in a civilian, nonmilitary population as well.

As could be stated for individuals with varying mental health disorders, individuals with PTSD often have difficulties seeking and accessing appropriate mental health services (Possemato, 2011). The various barriers impacting access to appropriate healthcare services, especially those seeking
mental health services, will be further detailed later in this chapter. Fortunately, as detailed by several contributing authors in this book, primary care offers the opportunity to develop unique and innovative integrated behavioral health models that lend themselves to addressing such barriers, thus dramatically improving access to treatment and reducing health disparities. Primary care patients with mental health problems often do not receive adequate mental health treatment and are more likely to seek care from their primary care medical provider than a mental health provider. Further, integrated primary care programs with a team of behavioral health clinicians can improve access, reduce stigma, and increase the convenience of receiving treatment to address all of their healthcare needs in one setting (Possemato, 2011). System factors play an important role in PTSD management in primary care and it is suggested that interventions are needed in primary care practices to make mental health services more integrated, in addition to enhancing community linkages between primary care to specialty mental health in the most severe of cases (Meredith et al., 2009). For example, behavioral health interventions in primary care are likely to take place in the context of a patient’s medical visit, which allows the opportunity for both the medical and behavioral health clinicians to explain the nature of their collaborative approach and “normalize” behavioral health as a part of the healthcare service delivery model. The primary care provider can introduce the mental health provider to the patient, which can greatly help increase the patient’s engagement (Blount, 2003). A model for integrating mental health into primary care can be an effective way to address PTSD for those who are unable or unwilling to participate in specialty mental health services, as well as treat those experiencing subthreshold levels of PTSD (Corso et al., 2009).

Patients and providers are more accepting of brief interventions in the fast-paced workflows that exist in primary care (Gunn & Blount, 2009; McDaniel & LeRoux, 2007). This lends itself to prevention efforts at three levels: (1) primary prevention (e.g., providing supportive parenting interventions on dealing with fussy babies to pregnant women to prevent toxic stress for that child); (2) secondary prevention (e.g., recognizing exposure to a trauma prior to onset of symptoms of ASD or PTSD and providing coping strategies); and (3) tertiary prevention (e.g., implementing brief yet comprehensive assessment strategies for PTSD’s impact on physical health concerns and developing a collaborative treatment approach with the patient and medical providers to address the physical and mental health comorbidities).

Behavioral health clinicians working as part of the primary care team have a substantial opportunity to provide effective brief interventions (Possemato, 2011). This is especially true when assessment and interventions are focused on improving one’s functioning and developing tools to manage reactions to stress (e.g., recommendations for sleep hygiene, deep breathing exercises), as opposed to attempting full remediation of deep-seated symptoms, characterological deficits, and interpersonal dysfunction.

With an effective population-based approach, the integrated primary care delivery model can provide the setting in which PTSD symptoms are most likely to be identified (Corso et al., 2009). Within the civilian population patients generally have more frequent contact with medical providers than mental health providers, underscoring the importance of addressing trauma exposure and PTSD symptoms in primary care. In order to do this, an effective assessment and intervention approach must be available in an integrated primary care model (Corso et al., 2009). Ouimette and colleagues (2008) outline the importance of disorder-specific screens in primary care in their validation study of the PC-PTSD for use in primary care. They stated that effective treatments for the disorder, as well as screens with good operating characteristics (i.e., good sensitivity and specificity) are available, thus making the case for screening for PTSD in primary care (Ouimette et al., 2008).
ACUTE STRESS DISORDER (ASD) VERSUS POSTTRAUMATIC STRESS DISORDER (PTSD)

ASD is worth noting in this chapter given the previously noted changes to the DSM-5 and the advent of Trauma and Stressor-Related Disorders as a single category in the latest version of the DSM. The diagnostic criteria for both ASD and PTSD are nearly identical, although there are two fundamental differences: (1) a diagnosis of ASD can only be given within the first month following exposure to trauma; and (2) ASD has a greater emphasis on dissociative symptoms than PTSD. According to the DSM-5, ASD is distinguished from PTSD in that the symptom pattern must occur within one month of the traumatic event and resolve within that one-month period. If symptoms persist and continue to meet criteria for PTSD beyond one month, a diagnosis of PTSD is warranted. Furthermore, unlike PTSD, an ASD diagnosis requires that a person experience three symptoms of dissociation (e.g., numbing, reduced awareness, depersonalization, derealization, or amnesia; Gibson, 2016). To these authors’ knowledge, currently there are no scientific studies validating the use of the measures described in this chapter for the identification of ASD.

IMPACT OF TRAUMA ON PHYSICAL HEALTH: SUMMARY OF THE ACE’S STUDY

The effect of PTSD on patients, families, and society is profound as it is often associated with functional impairment and behaviors that represent health risks, such as smoking, overeating, inactivity, and non-adherence to treatment (Freedy & Brock, 2010). Ouimette and colleagues (2008) summarized the nature of PTSD’s relation to physical health concerns and healthcare utilization rates. Additional studies have found higher rates of diabetes, cardiovascular disease, autoimmune diseases, and hypertension in patients also experiencing symptoms of PTSD (Andersen, Wade, Possemato, & Ouimette, 2010; Boyko et al., 2010).

It is often the case that patients with PTSD are hesitant to talk about psychological problems yet often present with vague and persistent physical complaints. Anecdotally, this is not always an intentional minimization of psychological distress with increased willingness to discuss physical complaints. Rather, the likelihood of discussing physical concerns may be due to culture or lack of understanding of stress and its manifestations. It is common for both patients and medical providers to lack awareness of the interplay between psychological distress and physical complaints, and that typical mental health presentations such as depression and anxiety might actually be a presentation better explained by underlying PTSD (Freedy & Brock, 2010). Once again, this underscores the importance of having well-trained behavioral health clinicians adept in recognizing trauma symptoms as these might otherwise be out of the awareness of patients and medical providers.

Perhaps the most significant contribution to the trauma field as it relates to the link between trauma exposure and physical health outcomes has been the Adverse Childhood Events (ACE) Study (Felitti et al., 1998) and the subsequent study on the convergence of evidence on the epidemiological and neurobiological effects of childhood trauma (Anda et al., 2006). Prior to this research, the relationship of health risk behavior and disease in adulthood to exposure to emotional, physical, or sexual abuse, and household dysfunction had not previously been described. The ACE score (ACES) research team’s initial efforts included an examination of a questionnaire about adverse childhood experiences in a sample of 9,508 patients from a large health maintenance organization (HMO). They studied seven categories of adverse experiences: psychological, physical, or sexual abuse; violence
against mother; or living with household members who were substance abusers, mentally ill or suicidal, or ever imprisoned. The number of categories endorsed was compared to measures of adult risk behavior, health status, and disease. More than half of the respondents reported exposure to an adverse childhood event in at least one category, and one-fourth reported at least two categories of childhood exposures. They found a graded relationship between the number of categories of childhood exposure and each of the physical health outcomes analyzed at a statistically significant level. This included significantly increased risk for alcoholism, drug abuse, depression, and suicide attempt. Further, they demonstrated a graded relationship between adverse childhood events and the presence of diseases in adulthood including ischemic heart disease, cancer, chronic lung disease, skeletal fractures, and liver disease. The seven categories of adverse experiences were strongly interrelated. Further, participants with multiple categories of traumatic exposures were likely to have multiple health risk factors later in life (Felitti et al., 1998).

Anda et al. (2006) later used the original ACE study as an epidemiological “case example” of the convergence between epidemiologic and neurobiological evidence of the effects of childhood trauma. They looked at the number of ACEs as a measure of cumulative childhood stress and hypothesized a “dose-response” relationship of the ACE score to 18 selected outcome health-related behaviors or problem sources and to the total number of comorbid outcomes. They found the risk of every outcome in the affective, somatic, substance abuse, memory, sexual, and aggression-related domains increased significantly in a graded fashion as the ACE score increased. Further, the mean number of comorbid outcomes tripled across the range of the ACE score. They concluded that the cumulative exposure of the developing brain to the stress response resulted in impairment in multiple brain structures and functions (Anda et al., 2006).

**IMPLEMENTATION: DOS AND DON’TS FOR TRAUMA ASSESSMENT IN PRIMARY CARE**

Given the nature of integrated primary care, rapid assessment is critical and can be a difficult and daunting task. As such, there are several considerations that should be made in the implementation of trauma assessment in primary care. Suggestions for effective implementation are provided.

**Develop a Trauma-Informed Care Approach**

The Substance Abuse and Mental Health Services Administration (SAMHSA) has provided a comprehensive, practical guide for the provision of behavioral health services, including a detailed trauma-informed care approach for clinical practices (SAMHSA, 2014a). It is a useful guide for developing trauma awareness, understanding the impact of trauma, and recommendations for screening and assessment of trauma exposure and PTSD. SAMHSA has also released a Treatment Improvement Protocol (TIP) of trauma-informed care specific to the behavioral health sector (TIP 57; SAMHSA, 2014b). The following recommendations are drawn largely from these documents. As noted by these contributors, screening to identify those with histories of trauma and subsequent symptoms is a prevention strategy (SAMHSA, 2014b).

**Take a Population-Based, Stepped Approach**

Exposure to trauma is common, and it has been reported that as many as half of the respondents to clinical surveys report a history of trauma. Many of the mental health comorbidities are more difficult
to treat without the recognition of a trauma history (SAMHSA, 2014b). Not addressing traumatic stress symptoms can lead to poor engagement in treatment, worse outcomes, premature termination from treatment, and greater risk of psychological symptoms. Freedy and Brock (2010) recommended a systems-based approach for PTSD, similar to scheduled HbA1C for diabetes. Further, as a stepped approach, the Primary Care-PTSD Screen (PC-PTSD) could be utilized as part of an initial screen and if the patient scores positive on this brief screen, the 20-item PTSD Checklist for DSM-5 (PCL-5) could be used as follow-up. It is also possible to consider training medical support staff in the administration of the initial screen. Ideally, a trained behavioral health clinician embedded in the clinic would then provide the follow-up assessment and brief intervention (Freedy & Brock, 2010). Ouimette et al. (2008) also suggested that combining the results of two screens might be the most effective practice to identify PTSD in primary care (e.g., PC-PTSD and General Health Questionnaire [GHQ]). It should be noted that within their study, the PC-PTSD performed slightly better than the GHQ in identifying PTSD.

**Address Barriers**

Many significant barriers exist for primary care patients with PTSD (Possemato, 2011). It can be helpful to think of barriers to seeking care and to providing effective care within a framework consisting of two major components. Logistical barriers (e.g., restraints related to time, scheduling, finances) and perceived barriers (e.g., the perception of value for attending or providing treatment to address trauma concerns) occur for both patients and clinicians.

**A Trauma Is Not a Trauma: Patient Variables**

As noted briefly earlier, patients may often be reluctant to discuss psychological distress with their care providers. Many patients avoid talking about their problems because of the stigma of mental illness, while others may know little about PTSD and are unaware that previous traumatic events—even those that occurred years ago—can have a profound effect on them now (Freedy & Brock, 2010). Bliese et al. (2008) also suggested that concerns related to stigma during the screening process can lead to under-reporting of symptoms. The contribution from Freedy and Brock (2010) further spells out some of the barriers that patients bring to care that can inhibit the screening and assessment process. These may include fear of retribution (from perpetrators of interpersonal violence); embarrassment, shame, and guilt (which can be reinforced by not screening or discussing such concerns); low self-esteem; learned helplessness; and limited insight. With regard to the final barrier previously noted, it is often the case that patients are simply not aware a trauma is a trauma, as they may have been repeatedly victimized so often they do not know their circumstances are indeed outside the norm. This underscores the importance of providing an explanation—including specific examples—of what any given screening measure is targeting, in addition to (as noted later) some psychoeducation and direct feedback to the patient on what the results of a screen signify.

**“I Don’t Want to Go There”: Clinician Variables**

Primary care providers may not have sufficient knowledge about PTSD recognition, and subsequent discomfort in asking about trauma histories and their ability to offer care may be limited when dependent on referrals to specialty mental health (Meredith et al., 2009). It has been suggested that integrating care, or linking systems of care (i.e., referrals to specialty mental health) can lead to better outcomes (Meredith et al., 2009). As noted earlier, the need to refer to specialty mental health can be alleviated when integrated behavioral health clinicians are adept in assessing for trauma exposure and PTSD in a collaborative capacity with primary care providers.

Time constraints are a legitimate concern given the amount of recommended routine screenings to be completed by primary care providers, as well as the limitations in care that are inherent
in a fee-for-service payment system (i.e., productivity often comes first, even at the potential cost of quality). However, it has been strongly suggested that PTSD often has such a profound impact on the patient’s well-being and overall health that it is not advised to forgo screening (Freedy & Brock, 2010). The benefit of integrated behavioral health providers notwithstanding, the variable of time constraints can impact the patient, medical provider, and behavioral health provider. This has been recognized by clinical practitioners as one of the most common barriers to PTSD screening (Hanley et al., 2013). Therefore, the use of the most efficient screening method in the acute care setting for identifying high-risk patients in the interest of connecting them with comprehensive care is essential (Hanley et al., 2013).

In addition to time, there may be an education deficit in medical providers that may preclude their ability to recognize symptoms of trauma victimization (Freedy & Brock, 2010). This may also include a lack of understanding of the clear link between trauma exposure and physical health concerns.

As noted earlier, physician discomfort, or the “I don’t want to go there” factor, can also be a barrier for physical health providers. Medical providers may be reticent to discuss the nature of trauma, as details of such events can be uncomfortable and even distressing to the medical provider. Whether this is an intentional avoidance on behalf of the medical provider to promote their own self-care or is out of their awareness, the integrated behavioral health clinician can play a vital role. This can include not only responding to the patient’s needs, or taking the case over once it is recognized that trauma is the etiology of the presenting concern, but also providing education to providers about the nature of trauma, helpful responses to patients experiencing trauma, comments medical providers can make to enhance treatment adherence, and strategies to convey empathy.

**Provide Education**

A basic learning model can be used in the initial screening and assessment phase in order to provide a simple and accessible method for patients to understand the nature of their experiences, which can be done at the initial step of a PTSD symptom screen (Corso et al., 2009). Psychoeducation and support from the outset of service should begin with an explanation of screening and assessment, including the screening process and why specific questions are important (SAMHSA, 2014b). It can be important to explain to the patient what they should expect with the screening and assessment process. This should include explaining to the patient that the screening process will help identify problematic symptoms that will be the focus of treatment in an effort to improve their overall functioning—even if discussing a history of trauma exposure can be difficult. Patients should also be given the opportunity to delay a response or refuse to answer a question. Further, it should be explained to the patient that difficult emotions that may arise during the screening process are normal (SAMHSA, 2014b). Finally, providing immediate feedback about the results of the screening can be an opportunity to help reduce vulnerability, as well as enhance strengths and develop coping strategies. This should be done in a way that avoids complicated theoretical jargon, answers questions in a direct and compassionate manner, and helps the client understand the clinician is working as an advocate on their behalf (SAMHSA, 2014). Freedy and Brock (2010) also recommended providing immediate feedback to the patient, including an explanation that PTSD is treatable. They also suggested answering questions directly and truthfully, while maintaining a high degree of empathy (Freedy & Brock, 2010).

A plan for follow-up treatment should be developed following a positive screen, in addition to whatever brief interventions are immediately provided (Freedy & Brock, 2010). It is also important to keep in mind that treatment in a primary care setting is not about “fixing” the patient’s problems, or fully remediating the effects of one’s trauma history. Rather, the emphasis should be on enhanced *functioning* and the development of strategies to manage the symptoms the patient is experiencing.
No Need to Go “Deep”: Screening Does Not Require Details of the Event(s)

Behavioral health clinicians may feel compelled to delve into a patient’s trauma following a positive screen. However, it is crucial that clinicians resist the urge to immediately learn about details about traumatic events. For the screening process, it is not necessary to probe deeply about details. For diagnostic purposes it is sufficient to focus on determining whether trauma occurred, general information about the trauma (e.g., type of trauma and when it occurred), and whether any recent events have exacerbated symptoms (Spoont et al., 2015). Although some patients may be inclined to provide details about trauma, it is the clinician’s duty to maintain the patient’s safety and avoid retraumatization, particularly when it is unknown whether the patient possesses skills necessary to discuss traumatic events.

Clinicians must be cautious about not conveying that a patient’s experience with trauma does not warrant clinical attention. According to TIP 57 (SAMHSA, 2014b), tone of voice and message are important. It may be helpful to mention to the patient that their safety is the most important at the time without arousing intense feelings. Further, it should be reiterated that it is indicated to first establish resources and discuss coping skills before proceeding to in-depth conversation about trauma exposure.

TOOLS FOR SCREENING AND ASSESSMENT OF TRAUMA IN PRIMARY CARE

A positive screen should indicate to the clinician that follow-up is necessary. A clinical assessment allows the clinician to develop a better understanding of the patient’s past and current experience, psychosocial and cultural history, assets, and values (SAMHSA, 2014b). In other words, a positive screen for symptoms of PTSD may be considered necessary, yet not fully sufficient for developing a comprehensive conceptualization of a patient’s presentation. A full-length diagnostic assessment may allow the clinician to recognize a positive screen was false and that there is no significant cause for concern. The reverse is also possible, in which a patient may minimize the full extent of symptoms on a questionnaire, yet may be more forthcoming about the true extent of symptoms. A patient may gain insight as to the connection of a trauma exposure with the presenting physical concern, therefore allowing the clinician to fine-tune their understanding of the patient’s underlying reason for their physical complaint (SAMHSA, 2014b).

At the same time, depending on varying clinic operations related to scheduling, workflow and productivity requirements, detailed diagnostic interviews and assessments can be impractical in primary care. Brief screening tools that are relatively easy for a patient to complete can help identify patients with the primary feature of PTSD (Prins et al., 2003). Early recognition is critical so that proper interventions can prevent problematic or more severe symptoms in the future. Several screening questionnaires exist to promptly recognize patients at high risk (Hanley et al., 2013).

Inclusion Criteria for the Measures Discussed in This Chapter

When a comprehensive screening process is not adapted, trauma-related symptoms are not detected and providers may direct services toward symptoms and disorders misaligned with the patient’s actual reason for presentation (SAMHSA, 2014b). PTSD assessment instruments that are psychometrically
sound can be utilized to measure various population samples impacted by trauma (Foa & Yadin, 2011). Further, Ouimette et al. (2008) note that screens should be easy to obtain, administer, complete, and score. It is also ideal for the measures to be validated in a primary care sample with multiple and different traumas, in addition to having strong psychometrics (Ouimette et al., 2008). The measures described subsequently in this chapter meet each of these standards.

Weathers and colleagues (2014) describe the process of revising the National Center for PTSD assessment instruments for DSM-5, including the PTSD Checklist, the Clinician-Administered PTSD Scale, and the Life Events Checklist. These instruments are among the most widely used and extensively validated in the field, each of which are included in this chapter. As Stein et al. (2014) note, it is possible that the use of any one diagnostic system may overlook many individuals. This consideration notwithstanding, the following measures were also selected as a strategy to recognize the construct of PTSD as a whole. Difference in cutoff scores, specificity, and sensitivity will be further discussed.

More generally, selections of the instruments described in this chapter were based on the following considerations: (1) the measures are available via public domain and free for use—as cost should be a significant variable for primary care providers, especially those working in a Federally Qualified Health Center (FQHC), as is the case for the current authors; (2) the measures are brief and can be administered in a time-efficient manner; (3) the measures are sensitive to various cultural considerations and amenable to recognizing PTSD in a diverse population; (4) the measures are available in Spanish; and 5) the measures are psychometrically sound. All measures discussed in this chapter are available from the National Center for PTSD at http://www.ptsd.va.gov. Despite their availability, these assessment tools should only be distributed to qualified mental health professionals and researchers.

It is important to consider that the following questionnaires should not be used for diagnostic purposes (Hanley et al., 2013) although the extended Clinician-Administered PTSD Scale (CAPS) may be an exception.

**Primary Care-PTSD (PC-PTSD)**

The Primary Care-PTSD Screen (see Appendix A) is a four-item screening measure that was designed for use in primary care and other medical settings and is currently used with all veterans using VA health care (National Center for PTSD, 2016b). The screen includes a prompt referring to a traumatic event and asks four subsequent questions, in a yes/no fashion, as to whether the respondent has experienced specific symptoms coinciding with criteria of PTSD. It is a simple and effective tool that can be easily utilized to recognize symptoms of PTSD in a primary care population (Davis, Whitworth, & Rickett, 2009) and can typically be administered in less than a minute (Hanley et al., 2013). It was the first PTSD screening measure developed with a primary care sample and is considered to be the single best screening test in primary care due to its brevity and diagnostic efficiency (Hanley et al., 2013). The PC-PTSD has been found to be a useful screening instrument for PTSD within a civilian population and to be as efficient as the General Health Questionnaire (GHQ-12) at predicting PTSD (Ouimette et al., 2008).

This tool assesses the underlying characteristics specific to PTSD: re-experiencing, numbing, avoidance, and hyperarousal (Prins et al., 2003). It was designed to be quick to administer and easy to read (Bliese et al., 2008). It is designed for patients with an eighth-grade reading level and has been validated in a Department of Veterans Affairs (VA) population (Prins et al., 2003).

A cutoff score of 3 “yes” answers has been suggested as the strategy to heighten the balance of sensitivity and specificity. With a cutoff score of 3, the PC-PTSD has a specificity of 87% and sensitivity of 78%, when compared with the Clinician-Administered PTSD Scale (CAPS; Prins et al., 2003).
A later study by Hanley and colleagues (2013) found that the PC-PTSD retained an even higher degree of specificity (93.4%) with slightly diminished sensitivity (72.4%) when compared with the PTSD Checklist-Civilian (PCL-C). It should be noted that in that study the PC-PTSD was not administered separately from the PCL-C; rather, PC-PTSD responses were extracted from the PCL-C, which could have skewed the responses. Although PC-PTSD items are included within the PCL-C, the PC-PTSD should be considered an independent measure apart from the PCL-C and not a subscale. Whereas the PCL-C is based on Likert responses, the PC-PTSD consists of dichotomous responses (“yes,” “no”). Furthermore, only the total score of the PCL-C should be interpreted. In an examination of the psychometric properties of the PC-PTSD, Prins et al. (2003) found the measure to have strong diagnostic efficiency in primary care. They also assessed a cutoff score of two “yes” responses yielded a sensitivity of 91% and specificity of 72%. Another examination of cutoff values showed that the PC-PTSD has reasonable sensitivity and specificity with either two or three “yes” responses (Bliese et al., 2008). They found a cutoff of three “yes” responses provided high degree of specificity (near 90%) while also maintaining a sensitivity value above 70%, and suggested a cutoff of three as preferable when simultaneous screening for other mental health concerns takes place.

A limitation of the PC-PTSD is that it is a binary screen (consisting of “yes/no” responses) and does not assess severity of PTSD. Rather, it only examines whether patients are beginning to show signs of the acute phase of PTSD (Hanley et al., 2013). Exposure to multiple lifetime traumas is common, and previous traumas may greatly influence a patient’s presentation for a separate trauma that precipitated the current presentation (Foa & Yadin, 2011). Overall, however, it was suggested that the PC-PTSD be used in a targeted approach to screening as a means to recognize PTSD, even in the context of several psychological and physical morbidities presenting for care (Freedy et al., 2010).

**PTSD Checklist (PCL-5)**

The PTSD Checklist (PCL) was originally designed as a 17-item DSM-correspondent self-report measure of PTSD and has become one of the most widely used and extensively validated PTSD measures (Weathers et al., 2014). There have been three versions of the PCL. The first two, the military (PCL-M) and civilian (PCL-C), were developed to assess those respective populations and refer generally to a stressful military experience or stressful experience in the past, while the third, the specific (PCL-S), is intended for the respondent to refer to a specific event (Weathers et al., 2014). It is intended primarily as a measure of PTSD symptom severity.

**Changes From Previous PCL for DSM-IV**

As noted by Weathers et al. (2014), the revision of the PTSD criteria for DSM-5 necessitated a parallel revision of existing DSM-correspondent measures of PTSD. They noted the several changes from previous versions of the PCL for the most updated version, the PCL-5 (see Appendix B), which include the following: (1) items were added to assess the three new symptoms, making it a 20-item measure, and other items were reworded to reflect other changes in wording of criteria in the DSM; (2) numerical anchors for the response scale were changed from 1–5 to 0–4. This makes no difference in psychometric analyses and is more intuitive, with a lowest possible score now at 0 rather than 17; (c) only one version of the symptom items was created, with no versions corresponding to the PCL-M or PCL-S (Weathers et al., 2014). Although there is only one version of the PCL-5 items, there are three formats that vary with respect to the assessment of Criterion A of the DSM: (1) without a Criterion A component; (2) with a Criterion A component; and (3) with the Life Events Checklist for DSM-5 (LEC-5) and extended Criterion A component (National Center for PTSD, 2016c).
Severity scores can be calculated for each symptom cluster, by summing item scores within a given cluster (re-experiencing, avoidance, negative alterations in cognitions and mood, hyperarousal); or for the entire disorder, by summing all 20 items. Rating scale descriptors are the same: “Not at all,” “A little bit,” “Moderately,” “Quite a bit,” and “Extremely.” The change in the rating scale, combined with the increase from 17 to 20 items means that PCL-5 scores are not compatible with PCL for DSM-IV scores and cannot be used interchangeably (National Center for PTSD, 2016c).

Administration and Scoring
The following recommendations on administration, scoring, and interpretation are provided by the National Center for PTSD (PTSD, 2016c). The PCL-5 is a self-report measure that can be completed by patients (e.g., given to the patient by a medical assistant while waiting for the behavioral health clinician). It takes approximately 5–10 minutes to complete. The PCL-5 can be administered with or without a brief Criterion A (brief instructions and items only), or with the revised LEC-5 and extended Criterion A assessment. It is important to note that when the PCL-5 is administered without Criterion A, trauma exposure should be assessed by another measure.

Interpretation
The PCL-5 should be interpreted by a qualified mental/behavioral health clinician. Several scoring options are available for the PCL-5 as indicated by the National Center for PTSD (2016c). A total symptom severity score (range: 0–80) can be obtained by summing the scores for each of the 20 items. Summing item scores within given clusters that correspond to DSM-5 criteria can provide a symptom cluster severity score: Criterion B (items 1–5), Criterion C (items 6–7), Criterion D (items 8–14), and Criterion E (items 15–20). Additionally, a provisional PTSD diagnosis can be made by interpreting each item rated 2 (“Moderately”) or higher as a symptom endorsed, then using the DSM-5 diagnostic rule requiring at least one Criterion B item (items 1–5), one Criterion C item (items 6–7), two Criterion D items (items 8–14), and two Criterion E items (items 15–20). According to the National Center for PTSD, preliminary validation work is sufficient to make a suggestion for initial cut-point of 33 for possible PTSD, although this may be subject to change as additional psychometric work becomes available.

Characteristics of a respondent’s setting should be considered when using PCL severity scores to make a provisional diagnosis, as recommended cutoff scores vary depending on the setting (Bliese et al., 2008; Wilkins, Lang, & Norman, 2011). A lower cutoff should be considered when screening or when it is desirable to maximize detection of possible cases, while a higher cutoff should be considered when attempting to make a provisional diagnosis or to minimize false positives (National Center for PTSD, 2016c). McDonald and Calhoun (2010) provided a comprehensive overview on selecting a cutoff score and the best uses of the PCL for various settings. Bliese et al. (2008) concluded when the PCL is used in primary care settings, cutoff values lower than 44 allow for administration and diagnostic efficiency. Further, they noted that lower cutoff scores may be useful in settings such as primary care, given that a screen is often administered to non-treatment-seeking patients and therefore, low PCL cutoffs with high sensitivity are ideal (Bliese et al., 2008).

Identifying a reasonable cutoff score based on population and treatment setting is a critical consideration. This should involve determining appropriate sensitivity, or the proportion of those with the disorder who are correctly identified by the measure, as well as specificity, which is the proportion of subjects without the disorder who are correctly identified (McDonald & Calhoun, 2010). Further, two other important determinants of diagnostic accuracy are positive predictive power (PPP), the proportion of those screening positive who actually have the disorder (i.e., the ratio of true positive among all positive screens) and negative predictive power (NPP), which is the proportion of those
screening negative who actually do not have the disorder (i.e., the number of false positives; McDonald & Calhoun, 2010). In their comprehensive review of several studies assessing the PCL, McDonald and Calhoun identified the PCL to have a PPP of 85% and an NPP of 15%. It is important to keep in mind that, as those investigators noted, a patient will present with symptoms on a continuum, and a cutoff score indicating a positive screen will only speak to the likelihood of true PTSD symptoms. Again, this underscores the importance of a multimethod assessment, and follow-up for patients with both a positive and negative screen is recommended. That is to say, further assessment (e.g., with use of CAPS) to rule in symptoms of PTSD is warranted, while those with a negative screen may still require clinical attention.

Wilkins et al. (2011) underscored that a universal cutoff score will not exist and a cutoff score should be determined based on the population being assessed. They suggested a higher cutoff score for populations with a high base rate of PTSD, such as veterans, while lower cutoff scores may be more useful with populations of lower base rates when it is crucial that symptoms of PTSD are not missed and when there is also potential for under-reporting (Wilkins et al., 2011).

Conybeare, Behar, Solomon, Newman, and Borkovec (2012) provided an assessment of the reliability, validity, and factor structure of the PCL in a nonclinical sample. They found the measure to have good internal consistency and retest reliability, in addition to strong patterns of convergent and discriminant validity (Conybeare et al., 2012). Importantly, they found the PCL-C did not perform well in discriminating trauma symptoms from other mental health symptoms. This is an important consideration and the current authors recommend screening for trauma exposure and PTSD take place in the context of a broader screening process involving an assessment of many mental health concerns, as well as substance abuse concerns. For a full review of various studies with recommended cutoff scores on the PCL ranging from 30 to 55, see McDonald and Calhoun (2010). Also, they note that when a screen is followed up by a more in-depth assessment, a cut score allowing for high sensitivity is preferred in order to not miss any patient presenting with symptoms of clinical concern.

When thoughtfully administered, the PCL can be extremely useful as a screening tool and can also be used to track changes of PTSD over time (McDonald & Calhoun, 2010; National Center for PTSD, 2016). This can be useful, but the onus falls on primary care clinics to follow up on positive screens, provide accurate diagnosis, and make available treatment options that are effective for primary care patients (Lang & Stein, 2005). This inherently should involve tracking a patient’s change in symptoms over time as a means to gauge meaningful outcomes.

In another review of each version of the PCL, Wilkins, Lang, and Norman (2011) found all three versions of the measure to be well validated. They demonstrated the PCL shows good temporal stability, internal consistency, test-retest reliability, and convergent validity. They also found the PCL correlated moderately highly with measures of anxiety, depression, poor quality of life, and other constructs of negative affect (Wilkins et al., 2011). These findings notwithstanding, a well-rounded assessment of various mental health concerns, even during a brief encounter with a patient, should be considered. Given that the PCL is also not intended as a diagnostic instrument, the PCL should be used in conjunction with other assessment tools (Wilkins et al., 2011).

Limitations

It should be noted that most of the studies on the PCL have reviewed the measure in the context of DSM-IV-TR criteria being utilized (McDonald & Calhoun, 2010; Wilkins et al., 2011). However, Weathers and colleagues (2014) predict the PCL-5 will likely maintain many of the favorable characteristics of the previous PCL and the changes noted earlier should be pertinent improvements. Regardless of the desirable characteristics, limitations of the measure still exist. Most significantly, it is not fully amenable to determining that symptoms are attributable to one or more Criterion A events,
and it does not fully assess functional impairment (Weathers et al., 2014). Additionally, neither the duration of the disturbance (Criterion E) nor the clinical significance of the symptoms (Criterion F) is assessed (McDonald & Calhoun, 2010). Finally, as previously noted, the PCL should not be used as a stand-alone instrument in providing a diagnosis of PTSD.

In addition, while the PCL is clearly one of the most widely used screening measures in primary care settings, having a patient complete the 20-item measure in addition to other measures assessing for comorbid symptoms can be time-consuming and difficult to complete in a short amount of time, as is the case when a screen or consultation is taking place in the context of the patient’s medical visit. In order to address this, research on brief versions of the PCL have been conducted. Lang & Stein (2005) developed four short forms of the PCL designed to capture a majority of the variance in the measures. The performance of these short forms was evaluated with a sample of primary care patients and it was found that the two-item and six-item versions had adequate psychometric properties (Lang & Stein, 2005). More recently, Han et al. (2016) conducted a study on the validity, reliability, and efficiency of cutoff values of the six-item PCL, termed the PCL-6, for underserved, largely minority patients in primary care settings of FQHCs. Scores ranged from 6 to 30, and reliability and efficiency statistics for a cutoff between 12 and 26 demonstrated a strong monotonic relationship between the PCL-6 scores and the probability of meeting CAPS diagnostic criteria. They concluded a single PCL-6 cutoff of 17 or higher might be suitable for identifying those with the greatest need of care (Han et al., 2016). Importantly, this study contained six items from the PCL-C, the 17-item version. It is hoped that studies on the development of brief measures derived from the updated PLC-5, the 20-item measure, will be forthcoming.

Finally, as an important note on the limitations of the PCL, Holliday, Smith, North, and Surís (2015) recognized that the PCL may not provide a full appraisal of symptoms as it does not identify symptom frequency from intensity as aspects of PTSD severity. Therefore, Holliday et al. modified the PCL to create the PCL-I/F as a means to identify both frequency and intensity of symptoms. Within a sample of male combat veterans they found statistically significant correlations between the PCL-I/F and the CAPS, suggesting initial validation of this measure. The creation of the PCL-I/F also stemmed from the 17-item version of the PCL.

**Summary and Overlap of the PC-PTSD and PCL**

Spoont et al. (2015) provided a systematic review of the utility of several self-report screening instruments for PTSD among primary care and high-risk populations. They concluded that both of these measures were the best performing instruments. Bliese et al. (2008) concluded the accuracy of the measures was virtually identical. These scales show reasonable performance characteristics for use in primary care clinics or in community settings with high-risk populations (Spoont et al., 2015). As previously noted, a stepped approach to screening and assessment is worth considering, and using the PC-PTSD with the PCL-5 as a follow-up measure could be a reasonable and fairly comprehensive approach to recognizing symptoms of concern in any primary care setting.

**Life Events Checklist (LEC-5) and Clinician-Administered PTSD Scale (CAPS-5)**

Several self-report measures and personality inventories are available for assessing PTSD. However, a structured interview is recommended to evaluate all of the diagnostic criteria, assess symptoms of comorbid disorders, and determine differential diagnosis (Foá & Yadin, 2011). This also allows for an ongoing demonstration of empathy and rapport building between the clinician and patient. As
previously noted, in addition to identifying a specific event, or series of events for symptom severity, often referred to as the index event or precipitating event, it is also important to assess the collective exposure to adverse events across a patient’s life span (Foa & Yadin, 2011). Given logistical restraints of time and workflow for screening and consultation during a medical visit, this is likely best accomplished at a full-length follow-up session.

The Life Events Checklist for DSM-5 (LEC-5; Weathers, Blake, Schnurr, Kaloupek, Marx et al., 2013; see Appendix C) is a 17-item trauma exposure measure that screens for exposure to potentially traumatic events in a respondent’s lifetime. The first 16 items pertain to events associated with risk for PTSD or distress (e.g., natural disaster, physical assault, sexual assault, captivity) while the final item pertains to any other stressful event not captured by the first 16 items. The LEC is the most widely used among the trauma exposure measures (Elhai, Gray, Kashdan, & Franklin, 2005). It was originally developed concurrently with the Clinician-Administered PTSD Scale for DSM-IV (CAPS; Blake et al., 1995) at the National Center for PTSD. Its intended purpose is to screen for history of an index traumatic event (worst or most salient) for the assessment of PTSD with a semi-structured interview, specifically the CAPS.

The CAPS is the gold standard assessment measure for PTSD, having served as the primary diagnostic measure in more than 200 empirical studies on PTSD (Weathers, Keane, & Davidson, 2001). The original CAPS (Blake et al., 1990) assessed for DSM-III-R symptoms of PTSD. It was then determined that two parallel versions that served distinct purposes were necessary. The CAPS-Diagnostic version (CAPS-DX) assessed PTSD symptom severity and diagnostic status in the past month or for the worst month since the trauma. The CAPS-Symptom Status version (CAPS-SX) assessed PTSD symptom severity over the past week. The CAPS-SX was designed for repeated use over brief time intervals in pharmacological research (e.g., Martenyi & Soldatenkova, 2006; Petty et al., 2001). Publication of the DSM-IV prompted a revision of the CAPS. Additional items were included for consistency with the DSM-IV criteria and the two CAPS were merged into a single CAPS scale that could function as a diagnostic measure or as a symptom severity scale (see Weathers et al., 2001 for review). Publication of the DSM-5 has once again prompted another revision and the development of the CAPS-5 (Weathers et al., 2013).

The CAPS-5 is a 30-item structured interview that assesses the 20 DSM-5 PTSD symptoms. Additionally, questions target onset and duration of symptoms, subjective distress, functional impairment, improvement since previous CAPS administration, overall PTSD severity, and endorsement of depersonalization versus derealization (National Center for PTSD, 2016d). The CAPS-5 can be used for diagnostic purposes (i.e., current/past month or lifetime diagnosis of PTSD) and to assess PTSD symptoms over the past week.

The LEC-5 is available in three formats: standard self-report, extended self-report, and interview. The 17-item standard self-report (see Appendix C) is used to screen for lifetime exposure to a potentially traumatic event(s) (i.e., did the traumatic event occur). Respondents are asked to report whether they experienced, witnessed, or learned about the potentially traumatic events listed in the form. The measure can be interpreted at the item level (i.e., specific events experienced), or a total score can be obtained by summing events that were endorsed. The extended self-report includes an additional set of questions designed to establish the worst event a respondent has experienced should the respondent report more than one potentially traumatic event. Lastly, the interview is a follow-up clinician-administered interview that is designed to obtain a history of child abuse and to establish whether Criterion A is met.

The CAPS-5, which includes the LEC-5, is the most current version and is administered by a clinician or a trained paraprofessional. The CAPS-5 scoring system allows for a total PTSD severity score and PTSD cluster severity scores. The total PTSD severity score is obtained by first summing frequency
and intensity ratings of each item (5-point scales ranging from 0 to 4) into single-item severity scores. Single-item severity scores for DSM-5 PTSD symptoms are then summed to calculate a total severity score. Symptom cluster scores are calculated by summing individual item severity scores for symptoms corresponding to DSM-5 symptom clusters: intrusion (Criterion B), avoidance (Criterion C), negative alterations in cognitions and mood (Criterion D), alterations in arousal and reactivity (Criterion E), duration (Criterion F), and clinically significant distress and functional impairment (Criterion G).

The CAPS has also been used as a diagnostic measure. Scoring rules have been developed to determine what frequency-intensity combination is necessary to establish the presence of a PTSD symptom. Several scoring rules that determine what frequency-intensity combination are required to determine whether a symptom is present have been developed for the CAPS (Weathers, Ruscio, & Keane, 1999). Selection of scoring rules must be done with caution, as implications for diagnosis are significant. Blanchard and colleagues (1995) found that the percentage of motor vehicle accident victims diagnosed with PTSD varied from 27% to 44% depending on whether a conservative or liberal scoring rule, respectively, was employed. The original CAPS did not identify an optimal cutoff score but did suggest a scoring rule for diagnostic purposes. According to the developers, a symptom is to be considered present if the corresponding item has a frequency rating of 1 (once a month) or higher and an intensity rating of 2 (moderate) or higher (scoring rule F1/I2). Scoring rules consisting of optimal cutoff scores for total PTSD severity have also been proposed (e.g., Orr, 1997). Weathers et al. (1999) argue that the decision to use one scoring rule over another should be based on the objective of assessment (i.e., screening vs. confirmation of a diagnosis vs. differential diagnosis). For instance, for screening purposes a more liberal scoring rule would be warranted to minimize risk of false negatives, whereas a more conservative scoring rule would be optimal for differential diagnosis to minimize risk of both false positives and false negatives (Weathers et al., 1999).

The LEC and CAPS have both been validated on civilian and combat veteran populations (Blake et al., 1990, 1995; Gray, Litz, Hsu, & Lombardo, 2004). The CAPS has also been adopted across traumatic events, including motor vehicle accidents (e.g., Blanchard et al., 1995), terrorist attacks (Palmieri, Weathers, Difede, & King, 2007), domestic violence (Griffin, Uhlmansiek, Resick, & Mechanic, 2004), and urban violence (Pupo et al., 2011). It has also been used with individuals with severe mental illness (Mueser et al., 2001), including schizophrenia (Gearon, Bellack, & Tenhula, 2004). The CAPS has been translated into several languages, including Spanish (Benuto, Olmo-Terrasa, & Reyes-Rabanillo, 2011), Chinese (Wu & Chan, 2004), Farsi and Russian (Renner, Salem, & Ottomeyer, 2006), German (Schnyder & Moergeli, 2002), Swedish (Paunovic & Ost, 2005), Bosnian (Charney & Keane, 2007), Cambodian (Hinton et al., 2006), Dutch (Hovens et al., 1994), and Albanian (Turner, Bowie, Dunn, Shapo, & Yule, 2003).

Psychometric properties for the LEC-5 are not currently available due to minimal revisions from the original version of the LEC. In fact, studies on psychometric properties on the LEC are minimal. Gray and colleagues (2004) conducted the most comprehensive study and reported that the instrument possesses generally adequate psychometric properties. The LEC compared favorably to other existing measures, including the Traumatic Life Events Questionnaire (TLEQ; r = .55), the Modified PTSD Symptom Scale (MPSS; r = .44), Mississippi Scale for Combat-Related PTSD (r = .33), the PTSD Checklist (PCL; r = .48), and the PTSD Checklist-Military Version (PCL-M; r = .43). Test-retest reliability was moderate, with 16 of the items with kappas over .50 (range .52–.80). One item’s kappa was .37, although the authors suggest that this item (“caused serious injury/death of another”) is seldom endorsed. Full-scale kappas were not as elevated as item-level kappas (range .36–.79). The authors argue that full-scale temporal stability is not relevant, as this instrument pertains to lifetime exposure to potentially traumatic events. Stability at the item level is more important because individual events
must be reported consistently. If respondents endorsed different items at different administrations, the measure would not be stable (Gray et al., 2004).

Currently, psychometric properties of the CAPS-5 are unavailable but are underway (Weathers et al., 2014). However, given minor changes the CAPS underwent when it was updated for use with the DSM-5, psychometric properties of previous CAPS will be reported (see Weathers et al., 2001 for review). A more recent review suggests that the CAPS possesses good psychometrics (Bardoshi et al., 2016). Weathers and colleagues (1999) provided psychometric properties for the symptom clusters and the full scale. Internal consistency for three symptom clusters ranged from .64 to .73 for frequency, .66 to .76 for intensity, and .69 to .78 for severity. For the full scale, alphas were .85 for frequency, .86 for intensity, and .87 for severity. Additional studies have demonstrated good test-retest reliability (Blake et al., 1990, 1995; Gearon et al., 2004), interrater reliability (Gearon et al., 2004; Mueser et al., 2001), and internal consistency (Benuto et al., 2011; Blake et al., 1990, 1995; Charney & Keane, 2007; Mueser et al., 2001). The CAPS is correlated with other trauma measures, including the PCL, the Patient Symptom Checklist, the Mississippi Scale for Combat-Related PTSD (M-PTSD), the Keane PTSD scale of MMPI, and the Combat Exposure Scale (Benuto et al., 2011; Blake et al., 1990; Weathers et al., 1999).

Research on the use of the LEC-5 in primary care is lacking. However, the simplicity and flexibility of the LEC-5 makes it suitable for use in primary care. Furthermore, although the LEC-5 is embedded in the CAPS-5, it has been validated as a stand-alone trauma measure (Gray et al., 2004), making it ideal for universal screenings of exposure to potentially traumatic events. For each item, respondents are asked to check: Happened to me, Witnessed it, Learned about it, Part of my job, Not sure, Doesn’t apply. In this way, it elicits information that would otherwise be overlooked. Endorsement of any of these events would trigger administration of the extended self-report designed to establish the worst event, if more than one. The clinician-administered interview would follow to establish whether Criterion A is met. The CAPS-5 would then be administered, only if Criterion A is met, as an index traumatic event would have been identified.

A major barrier to the use of the CAPS-5 in primary care is the time required to assess PTSD symptom severity or a possible diagnosis of PTSD should Criterion A be met. Depending on the model of integration of behavioral health services in primary care, a clinician might only have 15–20 minutes to meet with a patient. The CAPS-5 typically takes 45–60 minutes to administer. Given the length of time required to administer the full measure, administration of the LEC-5 stand-alone measure might be the most appropriate for screening. Although researchers have indicated that the CAPS can be used as a screener (Weathers et al., 1999), it would not be suitable or efficient for use in a brief encounter during a medical visit. Instead, when a screen is positive for exposure to a potential traumatic event, the CAPS would be administered during a follow-up encounter with a behavioral health clinician (e.g., diagnostic interview, therapy). The CAPS has demonstrated strong diagnostic utility with sensitivity that ranges from 82% to 91% and specificity that ranges from 71% to 91% (Weathers et al., 1999).

In primary care, it is also important to consider who will administer the measure. The LEC-5 standard and extended versions could be administered by medical support staff, as these are self-report measures. However, administration of the interview version and the CAPS would require a behavioral health clinician or trained paraprofessional to ensure that the interview is conducted in a clinically valid and reliable way. Furthermore, given that patients are required to recall aspects of their trauma and sequelae, a clinician with sound clinical abilities would be qualified to address any issues that come up (e.g., distress, dissociation).

Guidelines for decision-making with the assistance of the LEC-5 and CAPS-5 are lacking. However, clinical judgment along with findings from the LEC-5 and CAPS-5 can assist in determining the
need for behavioral health intervention. For instance, a positive history of exposure to a potential traumatic event on the LEC-5 stand-alone measure would warrant further attention from a behavioral health clinician to determine whether the patient meets Criterion A of the PTSD diagnosis in the DSM-5. The LEC interview version can be used to determine whether the respondent meets Criterion A, after which a more thorough assessment with a behavioral health clinician is warranted to assess for additional symptoms of PTSD and severity. The CAPS-5 can facilitate the diagnostic process and further inform next steps (i.e., continued treatment by the primary care provider vs. referral to behavioral health clinician vs. referral to specialty mental health).

Cutoffs on levels of severity of PTSD symptoms (i.e., mild, moderate, severe) are not available. However, research on scoring rules has suggested that individuals who warrant a diagnosis of PTSD when a conservative scoring rule is adopted present greater functional impairment and psychological distress than individuals who warrant a diagnosis of PTSD when a lenient scoring rule is adopted (Blanchard et al., 1995). Therefore, the adoption of certain scoring rules can inform severity as defined by functional impairment and psychological distress. For instance, patients who meet criteria for PTSD based on a conservative scoring rule might be more likely to benefit from a referral to specialty mental health. Patients who do not meet criteria for a PTSD diagnosis when a conservative scoring rule is adopted but do meet criteria when more liberal scoring is applied are likely to benefit from ongoing behavioral health services in primary care.

Ideally, trauma measures used in primary care should be flexible enough that they can serve as screeners or as outcome measures. Neither the LEC-5 nor the CAPS-5 is flexible enough as a stand-alone measure to serve these functions concurrently in primary care. While the LEC-5 is an adequate screener for exposure to potential traumatic events, the CAPS has been shown to be an effective outcome measure (e.g., Martenyi & Soldatenkova, 2006; Petty et al., 2001). One of the functions of the CAPS is to monitor changes in frequency and intensity (Blake et al., 1995). Multiple studies have used it as an outcome measure 8–12 weeks postintervention (e.g., Martenyi & Soldatenkova, 2006; Petty et al., 2001). Ratings have also been made over a one-week period (Nagy, Morgan, Southwick, & Charney, 1993). In fact, the CAPS-5 includes a version that is designed to assess symptoms in the past week, making it suitable to track improvement in symptom endorsement. Formal guidelines on the use of the CAPS as an outcome measure are not available to the best of our knowledge.

SYMPTOM OVERLAP AND DISEASE-SPECIFIC CONSIDERATIONS

Treatment of chronic medical concerns may be complicated by symptoms of general hyperarousal, decreased adherence to medical treatment, and engagement with problematic health behaviors (Anda et al., 2006; Felitti et al., 1998; Spoont et al., 2015). Resick and Miller (2009) provided a helpful overview of the reclassification of PTSD, making the case for a new category of diagnostic classification. They argued, in part, that PTSD is related to both the internationalization and externalization of pathology, and as a result does not clearly belong with disorders in either spectrum.

The following is a summary of comorbid mental health concerns that should be considered by integrated behavioral health clinicians, in conjunction with medical providers, as the likelihood of co-occurring disorders is high (SAMHSA, 2014b; Sonis, 2013; Spoont et al., 2015). According to Freedy et al. (2010), for patients with a prior trauma exposure, the mental health symptoms that can be associated with trauma, such as depression, anxiety, alcohol abuse, and somatic complaints, should all be screened and assessed concurrently. As previously noted, the PCL correlated highly with measures
of depression and generalized anxiety, indicating it may effectively detect negative affect, although this may not always be helpful in differentiating diagnosis when only a PTSD screen is administered (Wilkins et al., 2011). For an extensive list of resources on addressing comorbidities see the National Center for PTSD resource page (http://www.ptsd.va.gov/professional/co-occurring/index.asp).

**Anxiety Disorders**

It is essential to consider symptoms of anxiety when assessing patients with a history of trauma exposure, as the degree of symptom overlap is obviously high. Links between PTSD, specific phobias, generalized anxiety, obsessive-compulsive disorder, and panic disorder have been well documented, which is not surprising given PTSD’s previous placement in the DSM-IV-TR (APA, 2000) in the Anxiety Disorders category (Weathers et al., 2014).

Pietrzak et al. (2011) examined lifetime Axis I psychiatric comorbidity of PTSD in a nationally representative sample of US adults. Respondents with both PTSD and partial PTSD most commonly reported the unexpected death of someone close, serious illness or injury to someone close, and sexual assault as their worst stressful experiences. PTSD and partial PTSD were also associated with elevated lifetime rates of mood, anxiety, substance use disorders, and suicide attempts. Respondents with partial PTSD generally had intermediate odds of comorbid Axis I disorders and psychosocial impairment relative to trauma controls and full PTSD (Pietrzak et al., 2011). These findings also underscore the importance of comprehensive screening for various disorders in primary care, as well as follow-up assessments to aid in determining a differential diagnosis.

Milanak, Gros, Magruder, Brawman-Mintzer, and Frueh (2013) described generalized anxiety disorder (GAD) as a highly prevalent and distressing condition for individuals in both the community and primary care settings. They investigated the prevalence, comorbidity, physical and mental health impairment, and healthcare utilization of veteran participants with GAD, and compared symptoms of GAD and PTSD. They found that a large number of participants (12%) met diagnostic criteria for GAD and reported significantly worse emotional health, pain, and general health. In addition, these participants reported increased mental healthcare utilization and antidepressant medications. Importantly, GAD was found in 40% of participants with PTSD. Symptom severity and impairment were greatest among those with comorbid PTSD than those with GAD alone (Milanak et al., 2013).

Gros et al. (2011) found a high prevalence and features of panic disorder in a VA primary care setting when investigating prevalence, comorbidity, physical and mental health impairment, and healthcare utilization. Their participants completed diagnostic interviews and self-report questionnaires, and a chart review was also completed. Their results included a high degree of association with panic disorder and all of the other variables noted earlier. Further, panic disorder was highly comorbid with PTSD, with similar symptoms across all measures. These findings highlight the importance of improved recognition, assessment, and specialized treatments for panic disorder in VA Medical Centers and other healthcare settings (Gros et al., 2011).

**Borderline Personality Disorder**

McLean and Gallop (2003) examined the differences among women with a history of early onset versus late onset sexual abuse to determine whether they were more likely to meet diagnostic criteria for both borderline personality disorder (BPD) and complex PTSD. The diagnosis of both BPD and PTSD were significantly higher in women reporting early onset abuse than in those with late onset abuse, which may not be surprising given the characteristics of BPD must be present early in life to meet
criteria for the diagnosis. McLean and Gallop concluded that those in the sample with a sexual abuse history without BPD may be subsumed under the diagnosis of PTSD. Sexual abuse and paternal incest were significant predictors of both diagnoses (McLean & Gallop, 2003).

Pagura et al. (2010) noted that BPD and PTSD have important relationships with trauma, and the overlap between these disorders is well established. In a study examining the comorbidity of PTSD and BPD, they found that among individuals with identified BPD, 30.2% were also diagnosed with PTSD, whereas 24.2% of individuals with PTSD were also diagnosed with BPD. The individuals with comorbid PTSD-BPD were found to have reported poorer quality of life, increased comorbidity with other mental health disorders, higher odds of a lifetime suicide attempt, and an enhanced exposure to repeated childhood trauma than individuals with either trauma alone. Pagura et al. concluded that these disorders have a high degree of lifetime comorbidity, although the symptoms are not entirely overlapping, and they emphasized the clinical utility of diagnosing both conditions. Anecdotally, it is important to keep in mind that medical providers are not likely to have extensive training in the assessment of either BPD or PTSD. Whether to include either or both diagnoses in a patient’s medical chart should be discussed collaboratively among providers, as well as with the patient. This is especially true in primary care where patients are more likely to request access to their medical record as compared to other settings.

Depression

Freedy et al. (2010) found a substantial overlap of symptoms at a rate of 76.5% in their assessment of brief PTSD and depression screenings in primary care. Campbell et al. (2007) also reported on the high prevalence of depressive symptoms and PTSD comorbidity. Compared to those with depression alone, they found patients with depression and PTSD experienced higher degrees of pathology and numerous factors that complicate treatment (Campbell et al., 2007). They found rates of PTSD in patients previously identified with depression to be as high as 36%. They concluded comorbid PTSD among patients with depression is associated with illness burden, poorer prognosis, and delayed response to treatment (Campbell et al., 2007).

O’Donnell, Creamer, and Pattison (2004) raised the question of whether PTSD and depression are separate disorders in the aftermath of trauma or part of a singular construct. They concluded that PTSD and depression often occur together and reflect a shared vulnerability of distress. Their findings supported the existence of depression as a separate construct in the acute but not chronic aftermath of trauma as PTSD alone and comorbid PTSD and depression are indistinguishable (O’Donnell et al., 2004). Other studies have found a relationship between exposure to traumatic events and risk of comorbidity. In a study investigating the relationship among trauma history, immigration-related factors and mental health status, Kaltman et al. (2010) found that exposure to four or more traumatic events was associated with an increase in the probability of comorbidity. These results again emphasize the importance of comprehensive screening and assessment of several mental health disorders and that an ongoing review of symptom severity throughout the course of treatment is necessary in identifying the primary symptoms of concern.

Research on biological correlates of PTSD suggests that depression drives the relationship between PTSD and hypothalamic-pituitary-adrenal (HPA) abnormalities. Pinna, Johnston, and Delahanty (2014) found that depression contributed to elevations in cortisol responses in a sample of victims of intimate partner violence with PTSD. However, there have been contradicting findings suggesting that PTSD is associated with blunted waking cortisol response (see Morris, Compass, & Garber, 2012 for a review). Morris et al. (2012) report that daily cortisol output is lower for individuals with PTSD and comorbid PTSD and major depressive disorder (MDD) than no trauma controls. Pinna et al.
(2014) argued that several trauma-related factors, including how recent the trauma was and ongoing stressors (e.g., living in shelter), likely account for differences. Nonetheless, these findings highlight the extent to which PTSD and depression are related.

**Insomnia and Other Sleep Disorders**

Sleep problems, especially chronic insomnia and nightmares, are often some of the most disturbing aspects of PTSD (Gehrman, 2016). While these sleep problems are considered symptoms of PTSD, depending on the course of treatment and timing of exposure to the traumatic event, there is evidence to suggest that sleep disorders can become an independent and separate construct and disorder separate from the PTSD over time (Gehrman, 2016). Separately, and in addition to an assessment of PTSD symptoms, sleep-focused assessment and treatment should occur in the context of patients presenting to care following a traumatic event. For a review of the pharmacologic and cognitive behavioral treatment options available, see the National Center for PTSD website (http://www.ptsd.va.gov).

**Somatoform Disorders**

As previously noted, PTSD is associated with numerous adverse health consequences, including higher rates of diabetes, cardiovascular disease, autoimmune diseases, and hypertension, among several other health risk factors (Anda et al., 2006; Andersen et al., 2010; Boyko et al., 2010; Felitti et al., 1998). While PTSD is characterized by significant psychological distress, it is not uncommon for patients with a trauma exposure to initially present with somatic symptoms rather than more typical mental health complaints (Spoont et al., 2015). For those patients who do not respond to treatment for complaints typically identified as physical symptoms of concern (e.g., pain and sleep disorders), consideration for the presence of trauma exposure and PTSD should occur (Spoont et al., 2015). Increased rates of somatization and healthcare utilization by individuals with PTSD has been documented as an important consideration for accounting for high prevalence rates of PTSD in primary care (Escalona, Achilles, Waitzkin, & Yager, 2004).

**Substance Use Disorders**

Substance abuse is characterized by a maladaptive pattern of use leading to clinically significant impairment or distress (McCauley, Killeen, Gros, Brady, & Back, 2012). The comorbidities that exist between PTSD and substance use disorders (SUD) have been more widely researched and documented than the comorbidities between PTSD and other mental health disorders. Both disorders are prevalent and frequently occur separately, yet comorbid PTSD-SUD is associated with increased complexity of symptoms and higher cost of care when compared to treating either disorder alone (McCauley et al., 2012). This includes increased chronic physical health problems, poorer social functioning, higher rates of suicide attempts, more legal problems, increased risk of violence, lower treatment adherence, and less improvement during the overall course of treatment (McCauley et al., 2012). A treatment setting offering a multidisciplinary team-based approach to care that is able to respond to the physical health and mental health concerns in a parallel fashion is ideal.

Posttraumatic stress disorder has been found to have a high prevalence of comorbidity with patients also experiencing SUDs (Van Dam, Ehring, Vedel, & Emmelkamp. 2010). While prevalent, PTSD is often underdetected in SUD in general mental health treatment settings, underscoring the
need to consider use of a PTSD screening measure when a patient presents with primarily substance abuse concerns (Tiet, Schutte, & Leyva, 2013). Tiet et al. confirmed use of both the PCL and PC-PTSD to help identify PTSD in those presenting with SUDs in both SUD and general mental health treatment settings.

Jacobsen, Southwick, and Kosten (2001) also noted that alcohol and other SUDs are extremely common among patients with PTSD, and results of their research indicated the disorders are functionally related to one another. It is often the case that PTSD precedes substance abuse or dependence, as substances are often used to change the way the distressing emotional symptoms are experienced (i.e., substance use as an unhealthy means of coping; Jacobsen et al., 2001). Particularly as it relates to the hyperarousal and the physiological stress response system being especially active following a trauma, a patient may often seek sedatives, hypnotics, and alcohol in an effort to reduce bodily distress experienced. However, with the development of dependence, physiologic arousal resulting from substance withdrawal may also exacerbate symptoms of PTSD, leading to a return to the substance abuse and a cycle in which the SUD symptoms worsen the PTSD symptoms and vice versa (Jacobsen et al., 2001).

In a study designed to assess the diagnostic efficiency of the PC-PTSD as a screening measure for PTSD in a civilian patient sample with known substance abuse concern, van Dam and colleagues (2010) compared the PC-PTSD to an eight-item version of the PC-PTSD and another screening measure. Results showed a high sensitivity (.86) and moderate specificity (.57) for the PC-PTSD when using a cutoff score of 2. The diagnostic efficiency of the PC-PTSD was found to be equivalent to that of the two other screening measures, suggesting it is a useful screening instrument for PTSD within a civilian SUD population. This high comorbidity has important clinical implications related to symptom severity, worse outcomes, and difficulty in determining what to prioritize for treatment (van Dam et al., 2010). Therefore, developing effective screening of symptoms at the outset of treatment becomes critical.

Psychosis

The existing literature suggests there is a significant relationship between PTSD and psychosis. Trauma and PTSD are highly prevalent in individuals with severe mental illness (SMI) relative to the general population, yet trauma and PTSD remain overlooked in this population (Grubaugh, Zinzow, Paul, Egede, & Frueh, 2011). Very little is known about the manifestation and consequences of PTSD in more complicated patient populations, such as those with schizophrenia and major depressive disorder (see Grubaugh et al., 2011 for review). Seedat et al. (2003) conducted a review on the association between PTSD and psychosis and reported high rates of PTSD in patients with SMI. The authors also reported that psychotic symptoms may be common among patients with chronic PTSD. Higher rates of psychotic symptoms (e.g., auditory hallucinations) have been reported in certain racial/ethnic groups (David, Kutcher, Jackson, & Mellman, 1999; Wilcox, Briones, & Suess, 1991). Despite increased risk of trauma and PTSD among individuals with SMI, few systematic guidelines exist for treatment of comorbid PTSD and SMI (Seedat et al., 2003).

The scarcity of research on treatment of these comorbidities notwithstanding, some findings are promising. De Bont, van Minnen, and de Jongh (2013) conducted a study investigating the efficacy of prolonged exposure (PE) and eye movement desensitization and reprocessing (EMDR) therapy in patients with PTSD and a psychotic disorder. The researchers found that both PE and EMDR were associated with a reduction in PTSD symptom severity. At follow-up, 70% of patients no longer met criteria for PTSD. Moreover, patients did not experience any worsening of psychotic symptoms, general psychopathology, or social functioning. Findings from this study suggest that standard treatments for PTSD may be effective for individuals with comorbid SMI (De Bont et al., 2013).
Grubaugh et al. (2011) recommend a need to better understand the relationship between trauma and severe forms of mental illness and develop evidence-based treatments. They provide a strong review of extant literature and considerations for this population, as well as future directions for research and practice (Grubaugh et al., 2011).

CULTURAL CONSIDERATIONS

The nation’s demographics are changing, and it is anticipated that by 2044 racially/ethnically diverse populations will constitute the majority of the population in the United States (Colby & Ortman, 2014). Moreover, by 2060 one in five people in the United States will be foreign-born. Racially/ethnically diverse individuals are among those predominantly served in primary care. As such, several considerations must be made in the assessment of trauma and PTSD in this setting to ensure that a culturally sensitive approach is taken.

Language/Translation

Language is an important factor to consider when assessing PTSD among racial/ethnically diverse populations, as it has significant implications in the diagnosis and treatment of PTSD. In fact, language can account for disparate rates of diagnosis among certain racial/ethnic groups. For bilingual individuals whose native language is not English, assessment interviews conducted in English are associated with greater likelihood of being assigned a diagnosis or rated as pathological (e.g., Cheung & Snowden, 1990). Additionally, people whose native language is not English are more comfortable disclosing about emotionally laden content in their native language (Altarriba, 2003).

Given the role of language on adequate assessment and diagnosis of PTSD, there is growing interest in the development and validation of culturally appropriate measures in languages other than English. Although several instruments have been translated into non-English languages, the reporting on the translation process is often inadequate (Novy, Stanley, Averill, & Daza, 2001). Likewise, reliable and valid non-English-language assessment instruments are limited.

Psychometric Properties With Racially/Ethnically Diverse Populations

There is a vast literature on psychometric properties of PTSD assessment measures. Research on the psychometric properties of these assessment measures using culturally diverse populations is scarce (Benuto et al., 2011). Moreover, few assessment measures have been specifically developed for use with ethnically diverse populations. Among the measures adapted and translated for use with ethnically diverse populations, few have undergone extensive investigation to determine their validity and reliability with these populations.

Psychometric properties of PTSD assessment instruments developed and normed on the European American population do not consistently hold across all ethnically diverse populations. While psychometric properties for African Americans are similar to those of European whites (e.g., Ghee, Johnson, & Burlew, 2010), psychometric properties differ between Latinos and European Americans (e.g., Bourque & Shen, 2005). Bourque and Shen (2005) investigated the psychometric characteristics of the Spanish version of the Civilian Mississippi Scale (CMS). Low reliability of the measure was attributed to difficulty understanding items with reversed wording, thus highlighting the importance language in the development of instruments and in the general assessment of PTSD.
The Harvard Trauma Questionnaire (HTQ; Mollica et al., 1992) is an adaptation of the Indochinese version of the Hopkins Symptom Checklist-25 (HSCL-25; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974). The HTQ assesses for a series of events that refugee populations are likely to have experienced, including “torture,” “lack of food or water,” “rape,” and “murder of family or friend.” The HTQ was specifically developed for use with Indochinese patients. It combined refugee and culture-specific symptoms with core PTSD symptoms. The HTQ has since been adapted and validated for use with refugees from multiple ethnocultural backgrounds (e.g., Fouchier et al., 2012; Kleijn, Hovens, & Rodenburg, 2001; Oruc et al., 2008).

Instruments incorporating local idioms of distress have incremental validity and are able to predict functional impairment above and beyond translations of commonly used self-report measures (Jayawickreme, Jayawickreme, Atanasov, Goonasekera, & Foa, 2012). It has been suggested that the use of measures developed on Western (i.e., European American) populations likely contribute to disparate findings on PTSD prevalence, as these instruments have not demonstrated adequately validity and reliability in refugee populations (Jayawickreme et al., 2012).

**Additional Cultural Considerations**

Evidence demonstrating that PTSD can occur across any ethnocultural group following exposure to trauma suggests that PTSD is not a culture-bound syndrome that only affects individuals from Western cultures (Hinton & Lewis-Fernandez, 2011). These findings point to the likelihood that there is a universal core construct of PTSD. However, there is also evidence of ethnocultural-specific responses to traumatic events. Pole et al. (2008) provide a review of PTSD among ethnically/racially diverse populations. For instance, Latinos have been found to report greater severity of avoidance and hyperarousal symptoms (Ortega & Rosenheck, 2000), somatic complaints (Ruef, Litz, & Schlenger, 2000), and auditory hallucinations (David et al., 1999; Wilcox et al., 1991). Among African Americans, rates of trait dissociation have been found to be higher relative to their white counterparts (Zatzick, Marmar, Weiss, & Metzler, 1995), whereas for Asian Americans, somatization is a predominant feature of distress (Lee, Lei, & Sue, 2001; Mattson, 1993). Additionally, differences in PTSD symptom severity between ethnically diverse populations and European Americans have been reported (e.g., Ortega & Rosenheck, 2000).

Furthermore, accurate assessment of PTSD requires that idioms of distress that vary from core PTSD symptoms be identified (Lopez & Guarnaccia, 2000; Marsella, Friedman, & Spain, 1992). This is particularly important when working with individuals who have recently migrated (refugee or voluntary immigrant) or have a low level of acculturation. Common idioms of distress relevant to PTSD include “susto” and “nervios” among Latinos (Lewis-Fernandez et al., 2010), “thinking a lot” among Cambodians (Hinton, Kredlow, Pich, Bui, & Hofmann, 2013), and spirit possession among patients from Uganda (van Duijl, Nijenhuis, Komproe, Goernaat, & de Jong, 2010).

Additional factors related to endorsement of PTSD symptoms include response style and adherence to cultural norms. For instance, researchers have suggested that Latinos and Asian Americans are more likely to under-report distress and traumatic events consistent with social desirability (Hopwood, Flato, Ambwani, Garland, & Morey, 2009; Zhang, Snowden, & Sue, 1998). African Americans, on the other hand, may exhibit apprehension and healthy paranoia when discussing traumatic events (Wyatt, 1992).

Furthermore, certain sociocultural factors are associated with increased risk of PTSD, including exposure to acculturative stress (Ellis, MacDonald, Lincoln, & Cabral, 2008; Perilla, Norris, & Lavizzo, 2002). Likewise, chronic exposure to racial/ethnic discrimination and race-related stress increases vulnerability to PTSD among African Americans (Pole, Best, Metzler, & Marmar, 2005), Latinos (Fortuna, Porche, & Alegria, 2008), and Asian Americans (Loo, 1994). Trauma-associated racism and
“ethnoviolence” (i.e., violence directed at an ethnic group), whether experienced personally or vicariously, can be as detrimental as other types of trauma (e.g., rape, natural disasters, war) given the likelihood of psychological sequelae at the individual and group level (Helms, Nicolas, & Green, 2010).

**Immigration and Refugee Populations**

Distinctions must be made between immigrants and refugees, as of each of these groups migrate to the United States for different reasons and each is exposed to different traumatic events. Whereas immigrants voluntarily migrate to another country, refugees engage in an involuntary migration. Portes and Rumbaut (2006) make the distinction between being “pulled in” by perceived opportunities versus being “pushed out” by political/civil unrest or exposure to disaster.

Refugees have been found to be at greater risk for poor mental health, including PTSD, than voluntary immigrants (Cervantes, Salgado de Snyder, & Padilla, 1989; Hollander, Bruce, Burstrom, & Ekblad, 2011; Silove, Steel, McGorry, & Mohan, 1998). Pre-migration and post-migration factors (e.g., stress related to seeking asylum) might contribute to increased risk of PTSD and poor mental health among refugees (Porter & Haslam, 2005).

In light of refugees’ increased risk for PTSD and overall poor mental health, adequate assessments for PTSD are crucial in order to provide timely services. Several challenges in the assessment of PTSD among refugee populations have been identified, including the iatrogenic effect of eliciting memories of traumatic events when asking refugees to complete checklists and questionnaires (Mollica et al., 1992). Additionally, the development of instruments sensitive to a wide range of traumatic events is challenging due to diversity in geopolitical backgrounds (Mollica et al., 1992). Identification of culturally specific expressions of psychological across cultures is crucial for diagnostic and intervention purposes (Marsella et al., 1992; Miller et al., 2006). Moreover, the development of culturally meaningful mental health assessment is necessary to reduce the risk of false negatives and false positives.

**CASE STUDY**

The integrated primary care setting is one that serves a diverse population, racially and ethnically as well as socioeconomically. The integrated primary care setting in which the authors of this chapter work is an FQHC that serves a predominantly underserved and racially/ethnically diverse population. In working with such a diverse population, it is critical to engage in evidence-based practices that incorporate a culturally appropriate approach. Following is a clinical case vignette that highlights the clinical utility of selected measures while also accounting for cultural factors that may contribute to presentation and treatment. The clinical case vignette is loosely based on a patient with whom one of the authors worked closely. Several elements have been modified to maintain confidentiality and for didactic purposes. Following the vignette, an approach to assessment of PTSD in primary care is described.

Carmen is a 51-year-old Spanish-speaking Mexican woman who is unemployed and receives disability benefits. She immigrated to the United States 10 years ago and is in a relationship with a man she met upon her arrival to the United States. She is a mother of three adult children who reside in Mexico. Carmen presented to the clinic for a routine follow-up visit for management of diabetes mellitus. Her primary care provider (PCP) requested a behavioral health consult with the patient during the medical visit, as Carmen reported feeling depressed to her PCP and was tearful throughout the visit. Carmen’s PCP also expressed concern about the possibility that
Carmen’s presentation might be best accounted for by delusional disorder, as Carmen reported a story that the PCP believed was unlikely and indicative of possible psychological disturbance. Carmen reported that 6 weeks ago she received a phone call from her daughter in Mexico notifying her that her son and his family had been kidnapped and were being held for ransom. Carmen recalled the feeling of horror she experienced as she heard that her son, his wife, and their two daughters had been taken violently from their home. She expressed feeling “impotencia, mucha desesperacion” (helplessness and desperation) because she was unable to provide money to meet the demands of the kidnappers. She reported that for the past week she has experienced depressed mood, loss of interest in previously pleasurable activities, decreased energy, changes in appetite, crying spells, and feelings of guilt as she is unable to be with her family. She also reported difficulty with sleep onset and maintenance, as she has increased worry and nightmares about the many horrific things that could happen to her son and his family.

Upon hearing about such a horrific event, a behavioral health provider (BHP) might be inclined to agree with the PCP and assume that Carmen’s presentation is best accounted for by delusional disorder. Although this is definitely a plausible argument and should be part of an assessment, caution must be taken as an inclination toward this hypothesis might lead to an assessment focused on the confirmation of delusional disorder. Of course, mental health clinicians receive extensive training to ensure that they are objective and are able to engage in comprehensive assessments that provide equal weight to all explanations. Nonetheless, there are several steps that can be taken to safeguard against such a natural proclivity, that is, the confirmation bias.

Awareness of the sociopolitical factors that many Latino patients face in their native countries can be informative and contribute to a more culturally sensitive assessment. It is important to consider that violence has risen dramatically in Mexico and other parts of Latin America, including Honduras, El Salvador, and Venezuela, among other countries (Molzahn, Rios, & Shirk, 2012). Much of the violence that plagues Mexico is related to organized crime (e.g., drug trafficking). Violence disproportionately affects vulnerable populations, including women and children. Although homicide has been the most publicized example of crime south of the border, criminal activities like kidnapping and extortion have also risen dramatically and have a more direct effect on the general population. The BHP who met with Carmen was aware of the sociopolitical factors leading to migration and that continue to affect immigrant Latino patients, as many continue to have family in their native countries. An appreciation for the violence that affects the Mexican population provided the BHP with a crucial context for Carmen’s narrative. Upon considering Mexico’s sociopolitical climate and further assessment, the BHP concluded that it was likely that Carmen’s presentation was not consistent with delusional disorder. The BHP then proceeded to screen for PTSD.

The BHP recognized that although Carmen had not directly experienced a traumatic event, she met the DSM-5 Criterion A for a diagnosis of PTSD. Carmen learned about a traumatic event that threatened violent death that occurred to her family. Furthermore, this event was ongoing, as there was no indication of when her family would be released and whether they were being subjected to ongoing violence. Carmen noted at one point that she was concerned about the possibility of sexual violence directed toward her daughter-in-law and granddaughters. Upon recognizing that Carmen met Criterion A, the BHP proceeded to screen for PTSD using the PC-PTSD, particularly because Carmen had reported an onset of symptoms following the traumatic event. The PC-PTSD was selected because there is significant evidence supporting its use in primary care and it has been translated and validated for use with Spanish-speaking populations.

The BHP proceeded to the first item of the PC-PTSD, which corresponds to Criterion B (intrusion) of the PTSD diagnosis in the DSM-5. Carmen reiterated that she has nightmares about the
kidnapping and often finds herself imagining how her family was kidnapped and what they are being exposed to as she thinks about them. In response to the second item (Criterion C: avoidance), Carmen reported that she sometimes avoids answering the phone because she worries that she will receive news that her family is dead. The third item pertains to Criterion E (alterations in arousal and reactivity), which Carmen responded in the affirmative. She specified that she startles easily and is constantly looking over her back when she is out because she worries that she might become the victim of a violent crime. Prior to the screen, Carmen had also mentioned that she has not been able to sleep well since the kidnapping because she worries about her family, particularly at nighttime. Sleep disturbance is another symptom that comprises Criterion E. The BHP concluded the screen by asking Carmen the final item (Criterion D: negative alterations in cognitions and emotions). She reported that she sometimes feels “entumecida” (numb), as if she cannot feel anything, although at other times she feels overwhelmed with rage and self-blame because she should be able to do more to ensure that her family does not continue to suffer. As recommended by the authors of the PC-PTSD, Carmen’s screen was deemed “positive” as she responded “yes” to all items (three of four required for positive screen).

A positive PC-PTSD would ideally be followed by a structured interview, such as the CAPS-5, to confirm a diagnosis of PTSD. However, administering a structured interview during a 20-minute consult is not optimal in primary care. The PCL can be used to make a provisional diagnosis of PTSD when in the absence of a structured interview. The PCL has been revised for consistency with the DSM-5 PTSD criteria (PCL-5); however, it has yet to be translated and validated on a Spanish-speaking population. As such, the BHP decided to proceed with the administration of the Spanish translation of the PCL.

In order to make a provisional diagnosis, each item rated a 3 (“moderately”) or higher was treated as a symptom endorsed. The DSM-IV diagnostic rule was then used to determine whether Carmen met criteria for the diagnosis: one symptom from Criterion B (items 1–5), three symptoms from Criterion C (items 6–12), and two symptoms from Criterion D (items 13–17). Based on these scoring rules, Carmen met Criterion B (disturbing memories, nightmares, distress when reminded of event, physical reactions when reminded of event), Criterion C (loss of interest, feeling distant from others, feeling emotionally numb, foreshortened future), and Criterion D (trouble falling and staying asleep, irritability, difficulty concentrating, hypervigilance). Furthermore, Carmen had a total score of 56, suggesting that a PTSD diagnosis is likely. The BHP proceeded to schedule Carmen for a follow-up therapy appointment, as she was likely to benefit from ongoing support and strategies to continue coping with current situation. Additionally, the BHP followed up with Carmen’s PCP and shared her diagnostic impressions of Carmen.

Unfortunately, the case vignette described here is not uncommon in primary care settings that serve immigrant and refugee populations. In the past few years the United States has witnessed an influx of Central American immigrants fleeing from civil unrest, gang and drug violence, and extreme poverty seeking asylum. Countries outside the Western hemisphere (e.g., Syria) are experiencing an exodus due to sociopolitical strife (“Cumulative Summary of Refugee Admissions,” 2015). In light of the growing numbers of immigrants and refugees who have experienced traumatic events in their countries or during their migration journey, evidence-based PTSD assessment is critical. Behavioral health clinicians should strive to use measures that have been validated on culturally diverse populations and demonstrated sound psychometric properties. Moreover, an adequate PTSD assessment should account for cultural factors likely to contribute to presentation of symptoms.
SUMMARY

The resulting impact of trauma can have damaging effects on one’s functioning and overall health, as both physical health complications and mental illness comorbidities are common. The importance of recognizing a person’s exposure to a traumatic event, assessing the impact of such events on functioning, and developing effective, evidence-based treatments to address such concerns cannot be overstated. This is especially the case in a primary care setting as it is likely the first place a person will seek help to address the resulting impact of trauma. PTSD is common among primary care patients and is closely tied with functional impairment, physical health concerns, and high rates of medical utilization, and mental health comorbidities (Possemato, 2011).

Fortunately, in recent decades there has been significant progress in establishing psychometrically sound measures designed to recognize trauma exposure and PTSD, as well as related syndromes such as ASD (Foa & Yadin, 2011). While patients and providers are more accepting of brief interventions in the fast-paced workflows that exist in primary care (Gunn & Blount, 2009; McDaniel & LeRoux, 2007), the need to maximize the clinician’s time in order to best serve the patients in need remains critical. Based on the clear connection between adverse childhood events and exposure to trauma on health risk behaviors and poorer physical health outcomes, treating the respective disease states separately is indeed a recipe for inferior care.

Developing a trauma-informed care workflow, taking a population-management approach, developing a stepped process for screening and assessment, and addressing barriers to effective care are all important considerations for implementing a screening and assessment workflow for addressing PTSD in primary care. While the measures in this chapter are by no means an exhaustive list of screening questionnaires available to assess PTSD in primary care, careful consideration has been made in selecting the measures described here as they relate to efficiency, availability, and the strength of psychometric properties and research backing them.
APPENDIX A

Primary Care-PTSD Screen (PC-PTSD)

DESCRIPTION

The PC-PTSD is a four-item screen that was designed for use in primary care and other medical settings and is currently used to screen for PTSD in veterans at the VA. The screen includes an introductory sentence to cue respondents to traumatic events. The authors suggest that in most circumstances the results of the PC-PTSD should be considered “positive” if a patient answers “yes” to any three items. Those screening positive should then be assessed with a structured interview for PTSD. The screen does not include a list of potentially traumatic events.

SCALE

Instructions: In your life, have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month, you:

1. Have had nightmares about it or thought about it when you did not want to?
   YES/NO
2. Tried hard not to think about it or went out of your way to avoid situations that reminded you of it?
   YES/NO
3. Were constantly on guard, watchful, or easily startled?
   YES/NO
4. Felt numb or detached from others, activities, or your surroundings?
   YES/NO

Source: Prins et al. (2003). Scale available from the National Center for PTSD at http://www.ptsd.va.gov
APPENDIX B

PTSD Checklist for DSM-5 (PCL-5)

**Instructions:** Below is a list of problems that people sometimes have in response to a very stressful experience. Please read each problem carefully and then circle one of the numbers to the right to indicate how much you have been bothered by that problem in the past month.

<table>
<thead>
<tr>
<th>In the past month, how much were you bothered by:</th>
<th>Not at all</th>
<th>A little bit</th>
<th>Moderately</th>
<th>Quite a bit</th>
<th>Extremely</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Repeated, disturbing, and unwanted memories of the stressful experience?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Repeated, disturbing dreams of the stressful experience?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Suddenly feeling or acting as if the stressful experience were actually happening again (as if you were actually back there reliving it)?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Feeling very upset when something reminded you of the stressful experience?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Having strong physical reactions when something reminded you of the stressful experience (for example, heart pounding, trouble breathing, sweating)?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Avoiding memories, thoughts, or feelings related to the stressful experience?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Avoiding external reminders of the stressful experience (for example, people, places, conversations, activities, objects, or situations)?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Trouble remembering important parts of the stressful experience?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Having strong negative beliefs about yourself, other people, or the world (for example, having thoughts such as: I am bad, there is something seriously wrong with me, no one can be trusted, the world is completely dangerous)?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Blaming yourself or someone else for the stressful experience or what happened after it?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Having strong negative feelings such as fear, horror, anger, guilt, or shame?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Loss of interest in activities that you used to enjoy?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Feeling distant or cut off from other people?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Trouble experiencing positive feelings (for example, being unable to feel happiness or have loving feelings for people close to you)?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Irritable behavior, angry outbursts, or acting aggressively?</td>
<td>0 1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Not at all</td>
<td>A little bit</td>
<td>Moderately</td>
<td>Quite a bit</td>
<td>Extremely</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>16. Taking too many risks or doing things that could cause you harm?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>17. Being “superalert” or watchful or on guard?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>18. Feeling jumpy or easily startled?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>19. Having difficulty concentrating?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>20. Trouble falling or staying asleep?</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

APPENDIX C

Life Events Checklist—Version 5 (LEC-5)

Instructions: Listed below are a number of difficult or stressful things that sometimes happen to people. For each event check one or more of the boxes to the right to indicate that: (1) it happened to you personally; (b) you witnessed it happen to someone else; (3) you learned about it happening to a close family member or close friend; (4) you were exposed to it as part of your job (for example, paramedic, police, military, or other first responder); (5) you’re not sure if it fits; or (6) it doesn’t apply to you.

Be sure to consider your entire life (growing up as well as adulthood) as you go through the list of events.

<table>
<thead>
<tr>
<th>Event</th>
<th>Happened to me</th>
<th>Witnessed it</th>
<th>Learned about it</th>
<th>Part of my job</th>
<th>Not sure</th>
<th>Doesn’t apply</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Natural disaster (for example, flood, hurricane, tornado, earthquake)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Fire or explosion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Transportation accident (for example, car accident, boat accident, train wreck, plane crash)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Serious accident at work, home, or during recreational activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Exposure to toxic substance (for example, dangerous chemicals, radiation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Physical assault (for example, being attacked, hit, slapped, kicked, beaten up)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Assault with a weapon (for example, being shot, stabbed, threatened with a knife, gun, bomb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Sexual assault (rape, attempted rape, made to perform any type of sexual act through force or threat of harm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Other unwanted or uncomfortable sexual experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Combat or exposure to a war-zone (in the military or as a civilian)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Captivity (for example, being kidnapped, abducted, held hostage, prisoner of war)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Life-threatening illness or injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Severe human suffering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Sudden violent death (for example, homicide, suicide)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Sudden accidental death</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Serious injury, harm, or death you caused to someone else</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Any other very stressful event or experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
REFERENCES


PTSD in Integrated Primary Care


Sleep is increasingly recognized as a significant “vital sign” in the primary care setting. The amount and quality of time spent asleep is an important marker in an individual’s mental and physical health. It is estimated that 50–70 million American adults struggle with a sleep or wakefulness disorder (Institute of Medicine [IOM], 2006). Poor sleep can predict or present as a comorbid symptom of psychiatric or medical disease, and it is often difficult to treat insomnia as a secondary symptom. While the importance of sleep is not a necessarily novel concept, the methods by which one assesses and treats sleep disorders are rapidly evolving based on empirical evidence.

Typically, those practicing within the specialized field of sleep medicine have exclusively conducted the assessment and treatment of sleep. Sleep specialists can help identify sleep disorders such as obstructive sleep apnea, restless leg syndrome, circadian rhythm disorders, REM behavior disorders and insomnia, and can offer treatment options such as medication or specific appliances to help with breathing. More recently, behavioral sleep medicine clinicians, primarily psychologists, have entered the field with a host of evidence-based treatments aimed at improving the quality of sleep for those individuals presenting with sleep difficulties.

As a field, sleep medicine and behavioral sleep medicine are well fortified with empirically supported treatments for those patients who find their way to such a sleep specialist. However, this group of patients is relatively small compared to the general population of those suffering with chronic sleep problems. Primary care settings, on the other hand, are seen as the front line of detection of sleep disorders, with 49% of patients presenting with sleep complaints and 10%–15% of patients having chronic insomnia (Goodie & Hunter, 2014). At the same time, recognition of sleep assessment within this setting is limited by a lack of physicians who are trained to assess for sleep and a lack of specialists who are trained in treatment approaches located in these medical settings (Benca, 2005; Mindell et al., 2011; Papp, Penrod, & Strohl, 2002).

This chapter will focus on the role of behavioral sleep medicine assessment and its role within the context of primary care. It will outline the importance of direct inquiry about sleep within the primary care setting and present practical tools for assessing sleep disorders and related evidence-based treatment options. While the discussion will address the assessment of most sleep disorders, the primary focus will be on insomnia, given that this is the primary sleep complaint within primary care and the sleep disorder in which behavioral sleep medicine providers can have the biggest impact. Also provided is a discussion of cognitive behavioral therapy for insomnia, because this is the recommended first line of treatment and is easily administered in primary care at different levels of intensity.
BEHAVIORAL SLEEP MEDICINE

Behavioral sleep medicine (BSM) targets sleep issues and disorders from a behavioral perspective (Pigeon, Crabtree, & Scherer, 2007; Roth, Drake, & Roehrs, 2013). Historically, behavioral sleep medicine has been defined as

the branch of clinical sleep medicine and health psychology that: (1) focuses on the identification of the psychological (e.g., cognitive and/or behavioral) factors that contribute to the development and/or maintenance of sleep disorder and (2) specializes in developing and providing empirically validated cognitive, behavioral and/or other non-pharmacologic interventions for the entire spectrum of sleep disorders.

(Stepanski & Perlis, 2000, p. 343)

The term “behavioral sleep medicine” became part of the clinical lexicon at the beginning of the twenty-first century (Stepanski, 2003). However, the empirical roots of the field trace back to the 1930s with the work of Jacobson on progressive muscle relaxation (Stepanski, 2003). Although the best-known treatment in behavioral sleep medicine is cognitive behavioral therapy for insomnia (CBT-I), the field also consists of treatments for bedtime refusal, enuresis, parasomnias, CPAP compliance, and general behavioral medicine targets (i.e., smoking cessation, weight loss) as they are often related to sleep-related issues (Pigeon et al., 2007; Stepanski, 2003).

The growing recognition of the importance of behavioral mechanisms in sleep, in tandem with the pervasiveness of sleep disorders, has ignited the field and facilitated a rapid growth over recent years (Perlis, Smith, Cacialli, Nowakowski, & Orff, 2003). However, behavioral treatments for sleep disorders continue to be overlooked in medical settings (Mitchell, Gehman, Perlis, & Umscheid, 2012). For example, pharmaceuticals, such as benzodiazepines, are frequently prescribed as an initial treatment for chronic insomnia, despite multiple reports that identify cognitive behavioral therapy as the front-line treatment (Morgenthaler et al., 2006; Qaseem, Kansagara, Forciea, Cooke, & Denberg, 2016). Furthermore, as noted by Pigeon and his colleagues (2007), BSM specialists are not only specialized in evidence-based treatments for sleep but are also trained in the assessment and treatment of psychiatric disorders and quality of life issues that often co-occur with sleep and general medical disorders.

Despite such clinical indications, the uptake and dissemination of BSM approaches in medical settings is limited. Several barriers prevent both patients and clinicians from engaging in this intervention. Impediments to widespread use include lack of patient or physician awareness about BSM, a paucity of qualified clinicians, appointment time constraints, and limited insurance reimbursement. In addition, integration of behavioral health services into primary care, wherein most patients with sleep disorders are seen, faces structural and financial barriers that currently limit widespread application of these evidence-based approaches.

Barriers to Integration

An emerging body of literature has sought to better understand what factors contribute to the detection and treatment of sleep disorders within the primary care setting. One important barrier to consider is the infrequency with which sleep problems are addressed. One key finding has been that if a patient does not mention having difficulties with sleep, physicians are less likely to proactively inquire about the nature of their patients’ sleep. Williams and colleagues (2015) found that the odds of making a referral for obstructive sleep apnea (OSA) were 10 times more likely if the patient...
asked about OSA. Even when patients do present with reported sleep complaints, it is often the “quick fix” of medication that is requested due to ease and gaps in provider’s knowledge (Moloney, Konrad, & Zimmer, 2011). Furthermore, several studies have found that primary care physicians (PCPs) do not routinely screen patients for sleep disorders (Senthilvel, Auckley, & Dasarathy, 2011) and that a high number of family medicine database batteries do not include questions related to sleeping symptoms (Sorscher, 2011). For those practices whose review of systems (ROS) does include questions about sleep, often these questions may have poor sensitivity, or due to competing complaints a physician might not follow up on responses related to sleep (Grover, Mookadam, Chang, & Parish, 2016). A recent survey of 239 medical professionals revealed that therapies other than CBT-I were most commonly recommended for patients with insomnia despite it being recommended as the first line of treatment by the American College of Physicians (Qaseem et al., 2016) and a blue ribbon panel of scientists as far back as 10 years ago (National Institutes of Health (NIH), 2005).

On a positive note, continuing medical education focusing on the assessment of excessive sleepiness and OSA has been associated with improved patient-reported outcomes (Johnson et al., 2015). As this literature grows, the call to action among primary care communities has become louder, and the necessity for primary care to become more engaged in the field of sleep medicine has become more urgent (Sorscher, 2011). The next section will outline the behavioral principles that serve as the foundation for behavioral sleep medicine treatments and then highlight ways in which these BSM services can be best integrated within the primary care setting.

**Nuts and Bolts of BSM**

Competency in behavioral sleep medicine includes being well-versed in fundamental aspects in the science of sleep: the two-factor model of the sleep system, classical and operant conditioning, and the diathesis-stress model.

**Two-Factor Model**

The two-factor model of sleep posits that sleep and wakefulness are dynamic neurobiological states driven by homeostatic and circadian processes (Borbély, 1982; Kaplan & Harvey, 2014). The homeostatic process determines sleep likelihood and functions like a spring: pressure increases when awake and releases when asleep (Kaplan & Harvey, 2014). The circadian process determines alertness and is like a clock: it develops a rhythm that renews daily based on “zeitgebers” or cues such as routine, light, and temperature that entrain the body’s internal clock (Kaplan & Harvey, 2014; Lack & Bootzin, 2003). The homeostatic and circadian processes converge to shape sleep; specifically, high sleep pressure and low alertness are conducive to sleep (Kaplan & Harvey, 2014). Cognitive-behavioral approaches to insomnia, which will be discussed in greater detail later in this chapter, work to align patients’ behaviors with these processes. For example, sleep restriction builds homeostatic pressure by increasing time spent awake, thus increasing the likelihood of sleep. Wind-down and wake-up routines provide cues to increase or decrease alertness via circadian processes. These two processes thus work in tandem to drive natural sleep processes.

**Classical and Operant Conditioning**

Classical and operant conditioning are also central components of behavioral sleep medicine (Lichstein & Nau, 2003). Such conditioning posits that if a stimulus and response continually co-occur, the two will become paired, so that the stimulus elicits the response. This phenomenon ties into factors that perpetuate sleep disorders. For example, in insomnia, the bed can become paired
with activities other than sleep as well as the frustration and anxiety that comes from not being able to fall asleep. A CBT-I treatment component that targets this perpetuating factor is stimulus control, whereby patients only go to bed when sleepy and get out of bed if they cannot sleep (Kaplan & Harvey, 2014). In this context, the goal of stimulus control is to strengthen the association between the bedroom and sleep. As this association becomes stronger, the bedroom will again become a cue for sleep. Classical and operant conditioning are also at the core of other treatments in behavioral sleep medicine, which include treatment for sleep-onset association disorder, nighttime awakenings, compliance with treatment for sleep apnea, and bedtime refusal in pediatric populations (Stepanski, 2003).

**Diathesis-Stress Model**

The diathesis-stress model is a third mainstay of behavioral sleep medicine. This model, proposed by Spielman, Caruso, and Glovinsky (1987) regarding insomnia, contends that predisposing, precipitating, and perpetuating factors lead to insomnia. **Predisposing** factors (i.e., diatheses) are characteristics that make an individual vulnerable to insomnia, such as a genetic predisposition or trait characteristics such as hyperarousal. **Precipitating** factors are stressors that disturb sleep initially, such as pain or loss of a loved one. **Perpetuating** factors are behaviors or patterns that maintain insomnia, such as increasing time in bed to make up for lost sleep, napping during the day, or engaging in stimulating activities in bed when not able to sleep. CBT-I targets perpetuating factors to restore normal sleep (see Figure 18.1).

The field of behavioral sleep medicine is particularly vital in light of research linking poor sleep to a host of physical and mental health problems. Sleep deprivation is a risk factor for hypertension, diabetes, obesity, heart attack, and stroke (Colten & Altevogt, 2006). Recent estimates state that 28% of adults get 6 or fewer hours of daily sleep and that such sleep deprivation has cumulative impact on emotional and physical health. In 2013, the Centers for Disease Control and Prevention identified insufficient sleep as a “public health epidemic” and teamed up with the American Academy of Sleep Medicine (AASM) and the Sleep Research Society (SRS) to create the National Healthy Sleep Awareness Project (Morgenthaler et al., 2015).

While insufficient and disrupted sleep can be the result of various sleep disorders, the central focus of this chapter will be on insomnia, given that it is the most common sleep disorder and the one in which BSM has made the most advances. Insomnia overlaps with mental and physical health disorders: in clinical settings prevalence rises to approximately 50% (NIH, 2005). Insomnia is commonly comorbid with major depressive disorder, generalized anxiety, substance abuse, dementia, pain, diabetes, and cardiovascular disease as well as other sleep disorders (NIH, 2005). While in the past, comorbid insomnia has been viewed as secondary to other diagnoses, there is growing evidence that (1) the mechanisms perpetuating insomnia are consistent regardless of the initial cause; (2)
sleep disturbance often predates and predicts the development of other disorders; (3) insomnia and other disorders frequently have a reciprocal rather than a unidirectional relationship (i.e., a comorbid condition can make sleep worse and similarly poor sleep can exacerbate symptoms associated with comorbid condition); and (4) sleep problems often persist once the concurrent condition has been treated (Stepanski & Rybarczyk, 2006). For example, studies have shown that even with the successful treatment of depression, patients continue to struggle with sleep difficulties (Carney, Segal, Edinger, & Krystal, 2007; Manber et al., 2003; Mouchabac, Ferreri, Cabanac, & Bitton, 2003; Nierenberg et al., 1999).

The American Academy of Sleep Medicine identifies CBT-I as a standard treatment for both primary and secondary insomnia, the highest level of recommendation (Morgenthaler et al., 2006). Recently, the American College of Physicians, an international organization of internal medicine specialists, published guidelines recommending that all patients receive CBT-I as the initial treatment for chronic insomnia (Qaseem et al., 2016). These guidelines mark a significant turning point in the treatment of insomnia within the medical setting. Multiple meta-analyses support the efficacy of CBT-I (Geiger-Brown et al., 2015; Irwin, Cole, & Nicassio, 2006; Montgomery & Dennis, 2003; Morin, Culbert, & Schwartz, 1994; Murtagh & Greenwood, 1995; Wu, Appleman, Salazar, & Ong, 2015), and multiple randomized controlled trials indicate that the efficacy of cognitive behavioral therapy is similar to that of pharmacotherapy posttreatment with fewer side effects (for a review, see Perlis, Smith, Cacialli, Nowakowski, & Orff, 2003). Moreover, the treatment effects of cognitive behavioral therapy appear to be more durable than those of pharmacotherapy, suggesting that cognitive behavioral therapy is a more efficacious treatment in the long term (Perlis et al., 2003). Although pharmacotherapy is markedly more efficient than CBT-I in the short term, the greater long-term efficacy of cognitive behavioral therapy likely makes it more cost-effective over time (Perlis et al., 2003).

Additionally, a paradigm shift has occurred in the last 15 years with respect to insomnia that is comorbid with a medical or psychiatric condition and CBT-I (Stepanski & Rybarczyk, 2006). It was previously assumed that individuals with insomnia that was “secondary” to a comorbid medical or psychiatric conditions were not good candidates for CBT-I because it would not address the root cause of the problem. We now have 15 years of research (see 3-P model in Figure 18.1 and the earlier discussion) and over 35 randomized clinical trials (see Wu, Appleman, Salazar, & Ong, 2015 for meta-analysis) demonstrating that comorbid insomnia is not only as responsive to CBT-I as primary insomnia but also that CBT-I often leads to improvements in the comorbid medical or psychiatric condition. For example, the first two authors of this chapter and colleagues (Margolies, Rybarczyk, Vrana, Leszczyszyn, & Lynch, 2013) found that CBT-I led to improvements in both sleep and PTSD symptoms and depression in individuals with comorbid insomnia and PTSD. These studies have fully supported the 3-P model of insomnia, showing that the primary precipitating and perpetuating factors are behavioral ones regardless of whether the insomnia is a medical or psychiatric condition overlaid on the insomnia. It may be even more important to treat insomnia in the context of comorbid insomnia because it will enhance sleep processes vital for emotional functioning as well as growth and rejuvenation of immune, nervous, muscular, and skeletal systems. Therefore, in assessing and treating insomnia in primary care, it is absolutely essential to convey to other healthcare providers as well as patients that the comorbid insomnia can and should be treated with CBT-I.

**Integration of BSM Within Primary Care**

Although the efficacy of BSM approaches has been well established and has solid empirical support, the frequency in which patients are referred to BSM providers does not match the state of the science.
Behavioral interventions for sleep are primarily delivered in sleep clinics or mental health settings and thus miss a significant percentage of patients who might benefit from such treatment (Goodie, Isler, Hunter, & Peterson, 2009). In fact, a recent study investigating referral practices and attitudes of medical professionals within two large academic medical centers in the United States found that of the many patients seen in a one-month period, only 1% were referred for CBT-I treatment (Conroy & Ebben, 2015).

The question of how to better address this gap in science and real-world practice is an important area of investigation. Multiple collaborative care models have been proposed to tackle this discrepancy in knowledge and treatment. One integrated care model involves a trained mental health clinician who acts as a behavioral health consultant (BHC) and sees patients in the context of the primary care setting after being consulted by a PCP. Consistent with the primary care visit, patients are typically seen for 15–30 minute appointments with a brief number of follow-up visits to assess treatment intervention and maintenance. Studies have demonstrated that such an approach is effective and that after treatment, patients report improved quality and greater satisfaction with sleep as well as less interference with daytime functioning (Goodie et al., 2009). A handful of studies have also shown that this model can be effective when non-sleep specialists (i.e., primary care providers and social workers) are trained in CBT-I (Bothelius, Kyhle, Espie, & Broman, 2013; Espie et al., 2007). This will be addressed in more detail later in this chapter. A one-time sleep consultation by a sleep medicine specialist is another promising approach that has shown to facilitate healthcare providers’ attention to sleep problems and improved sleep/wake symptoms that might not have otherwise been treated (Edinger, Grubber, Ulmer, Zervakis, & Olsen, 2016; Khawaja, Hurwitz, Herr, Thuras, & Cook, 2014).

BEHAVIORAL SLEEP PROBLEMS AMONG ADULTS IN PRIMARY CARE

Sleep disturbances present in a wide range of conditions seen in primary care and insomnia is thought to be the most prevalent sleep complaint in this setting. The 2005 Sleep America Poll (National Sleep Foundation, 2005) reported that 33% of adults report experiencing at least one sleep problem every night. This section details the impact of insomnia on physical and emotional health and alternative sleep complaints that should be differentiated from insomnia.

Insomnia

Insomnia is a highly prevalent sleep disorder with pernicious consequences on physical and mental health (Hossain & Shapiro, 2002). Insomnia is characterized by insufficient or poor quality sleep, accompanied by difficulties with sleep onset, sleep maintenance, and/or early morning awakenings (American Psychiatric Association, 2013). For diagnosis, these symptoms must persist for at least three nights a week, despite adequate effort and opportunity for sleep, and cause significant distress or impairment on functioning (i.e., occupational, behavioral). Risk factors for insomnia include concurrent medical or psychological disorders, female gender, lower socioeconomic status, older age, and minority race (Ohayon, 2002). Insomnia warrants targeted treatment because it is (1) highly prevalent, (2) unlikely to remit spontaneously without intervention, (3) an independent risk factor for psychiatric or medical illness, and (4) detrimental to occupational performance and to interpersonal interactions (Findley & Perlis, 2014).
Insomnia is often characterized as being either primary or comorbid. In primary insomnia (also called “idiopathic insomnia”), the symptoms are not likely caused or comorbid with another physical or mental health condition. In contrast, comorbid insomnia (previously called secondary insomnia), includes symptoms of insomnia that are related to a concurrent medical problem (e.g., chronic pain), psychological condition (e.g., anxiety, schizophrenia), substance use (e.g., alcohol use, caffeine use), or another sleep disorder (e.g., obstructive sleep apnea). Of note, the newer term, “comorbid insomnia,” is recommended by the National Institutes of Health (2005), as this term better reflects the dynamic, bidirectional relationship between sleep symptoms and comorbid conditions.

The majority of insomnia cases are comorbid with other mental or physical health conditions (NIH, 2005). In fact, the prevalence of comorbid insomnia magnifies with the number of co-occurring conditions. The more severe the insomnia, the more at-risk individuals are for developing other disorders. Conversely, the greater the number and severity of medical and mental health conditions an individual has, the more likely they are to experience disturbed sleep. Insomnia is a ubiquitous condition that plays a role in many common medical and mental health conditions. Cancer, chronic kidney disease, Parkinson’s disease, hypertension, diabetes, and pain (to name a few) are all associated with negative effects on sleep. As medical conditions become more severe and chronic, the greater the likelihood that sleep will be more profoundly and negatively affected. For these patients, problems with sleep and the associated deleterious consequences of insomnia add to the complexity of disease management and their overall quality of life.

Impaired sleep also has a profound role in many psychiatric disorders such as depression and anxiety, and it is considered a core and consequent symptom of many of these conditions (e.g., Riemann & Voderholzer, 2003; Staner, 2010). Clinical and empirical observations indicate that disturbed sleep can have a powerful and harmful effect on the course and severity of a psychiatric disorder, and that left untreated, it will persist (Ohayon, 2002; Reddy & Chakrabarty, 2011).

Insomnia can be further subtyped by the duration of symptoms. Acute insomnia refers to symptoms lasting less than a month, while chronic insomnia symptoms endure for more than one month. If untreated, acute insomnia may evolve into chronic insomnia. Several studies have demonstrated that between 16% and 42% of insomnia patients develop chronic insomnia (Buysse et al., 2008; Hohagen et al., 1994; Skapinakis et al., 2013). Further, worse incident insomnia symptoms increase the likelihood of chronicity, underscoring the importance of screening and targeted intervention (Morin et al., 2009; Roberts, Roberts, & Chen, 2002).

Additionally, practitioners may characterize insomnia by the patient’s predominant symptom (i.e., sleep onset vs. sleep maintenance), which may help to inform treatment aims and priorities. For patients with insomnia, sleep maintenance symptoms are the most common (approximately 50%–70%), followed by difficulty initiating sleep (35%–60%) and nonrestorative sleep (20%–25%), according to epidemiological research (Morin et al., 2009). Of note, it is most common for patients to have multiple symptoms of insomnia (Hohagen et al., 1994).

While insomnia is considered to be one of the most prevalent sleep disorders, precise estimates vary widely throughout the literature, due to inconsistent research definitions. A systematic review of 50 epidemiological studies provided aggregate prevalence ratings for the most common diagnostic categorizations (Ohayon, 2002). Common diagnostic categorizations included (1) difficulty with sleep initiation/maintenance or nonrestorative sleep, regardless of duration or daytime sleepiness (~33% of adults present at least one); (2) any insomnia symptoms and daytime sleepiness (~9%–15% of adults); (3) dissatisfaction with the quality or quantity of sleep (~8%–18% of adults); and (4) official ICD/DSM diagnostic criteria (~6% of adults). All of these prevalence rates...
are higher in women than men, with women having greater than a twofold likelihood of insomnia (Buysse et al., 2008).

**Behaviorally Induced Insufficient Sleep Syndrome**

Insomnia should be distinguished from behaviorally induced insufficient sleep syndrome (BISS), a disorder that describes *volitional*, chronic sleep restriction (American Psychiatric Association [APA], 2013). In contrast, patients with insomnia struggle with sleep despite adequate opportunity. Diagnostic criteria for BISS include (1) excessive sleepiness for at least three months, (2) habitual sleep episode is curtailed, and (3) when the habitual sleep/wake schedule is disrupted, patients sleep longer than usual (American Academy of Sleep Medicine, 2005). Of note, for patients with BISS, sleep quality as well as the ability to fall asleep or maintain sleep is unimpaired. Thus, primary care staff should assess whether the short sleep duration reflects inadequate opportunity for sleep or may be symptomatic of insomnia.

**Delayed Sleep Phase Disorder**

It is also critical to distinguish insomnia from delayed sleep phase disorder (DSPD), wherein the primary complaint is abnormal sleep timing. Delayed sleep phase disorder is characterized by a sleep onset two or more hours later than the patient’s desired bedtime (between 1 a.m. and 6 a.m.). DSPD is only considered a disorder when accompanied by daytime sleepiness and impairments in daily function. Delayed sleep phase is most common in young adults, affecting 1%–16% of young adults or adolescents (Gradisar & Crowley, 2013). While difficulty with sleep onset is prevalent in DSPD, this disorder differs from insomnia in that sleep quality and duration are unimpaired, when patients are allowed to wake up as late as needed for adequate duration (Gradisar & Crowley, 2013). However, delayed sleep phase disorder may be perpetuated by similar cognitive biases to patients with insomnia, including sleep-related attentional bias, distorted perceptions of sleep and daytime functioning, dysfunctional beliefs, and safety behaviors (Richardson, Gradisar, & Barbero, 2016).

**ASSESSMENT OF SLEEP DISORDERS IN PRIMARY CARE SETTINGS**

The incentive to ask patients about their sleep and to assess sleep in its own right should be high and considered an important area of discovery within the primary care setting. The AASM (2016) has designated screening for sleep disorders in the electronic health record (EHR) as a health priority and has a task force developing pertinent sleep-related data fields (Strollo et al., 2011). Consequences of sleep disorders include mood disturbances, medication habituation, memory impairment, daytime fatigue, vocational and interpersonal difficulties, increased healthcare utilization, impaired health status, and accidents. As mentioned in the previous section, sleep disorders can have a profound impact on concurrent psychiatric and medical diagnoses as well as on daytime functioning.

The aim of this section is to provide clinicians with a guide and resources for assessing sleep difficulties within the primary care setting. As Klingman and colleagues (2016) address, integrating
assessments of sleep difficulties within the medical exam is complicated by the fact such assessment
does not focus on one specific illness. Rather, sleep difficulties encompass a class of disorders includ-
ing insomnia, sleep-disordered breathing, circadian rhythm disorders, restless leg syndrome/periodic
limb movements of sleep, parasomnias, and narcolepsy. Assessment, thus, should be considered a
means of better understanding the nature of patients’ sleep and its role in their physical and emo-
tional well-being. To date, there are no official standards or consensus statements for screening for
sleep disorders within the context of primary care and current national screening guidelines do not
address sleep.

Initial Screening

Assessment of sleep disorders is a multidimensional process and can be conducted using a combina-
tion of methods including retrospective and prospective self-report (questionnaires and sleep diary),
objective measures of rest–activity patterns (actigraphy), and physiological measures (polysomnog-
raphy). A thorough sleep history that addresses sleep and wake function and attempts to identify
maladaptive behaviors and cognitive processes is an important evaluative process for diagnosis and
treatment (Schutte-Rodin, Broch, Buyse, Dorsey, & Sateia, 2008). Such a comprehensive undertak-
ing, however, is often beyond the scope of what one clinician can undertake within a primary care
visit. Providers have a small window of time in which to manage a variety and increasing number
of medical problems (Abbo, Zhang, Zelder, & Huang, 2008; Gottschalk & Flocke, 2005; Tai-Seale,
McGuire, & Zhang, 2007), and patient-centered guidelines emphasize the timeliness and efficiency of
medical visits (IOM, 2001; Russell, Ibuka, & Carr, 2008).

Given this pressure of time, a general rule of thumb is that assessment should focus on factors
that are maintaining or exacerbating sleep difficulties and the daytime costs of such impairment
(Goodie & Hunter, 2014). In this section, we will first outline categories of sleep that are important
to address in the sleep history for proper diagnostic and treatment purposes. The second section will
focus on specific assessment tools that can facilitate screening for insomnia and other sleep disorders
as well as track change pre- and posttreatment. It should be noted that the integration of sleep assess-
ment within the primary care setting is a relatively recent emphasis, and thus many of these instru-
ments have not been well researched in the context of primary care.

In primary care settings, assessment of sleep problems can consist of several components. Perhaps
the most important component is a thorough sleep history obtained as part of the clinical interview.
Other components may include the administration of one or more patient self-report measures and/
or the obtaining of physiological data.

Sleep History

A complete sleep history will serve as the best foundation for case conceptualization based on the 3-P
Model of Insomnia. A better understanding of the predisposing and precipitating factors that contrib-
uted to the onset of insomnia will help elucidate important etiological considerations and possible
barriers to treatment that are not within the patient’s control. For treatment purposes, it is the per-
petuating factors that maintain the sleep difficulties that will serve as targets for intervention. Specifi-
cally, the clinician should identify the maladaptive behaviors and cognitions that perpetuate chronic
insomnia (Schutte-Rodin et al., 2008).
Based on Goodie and Hunter’s (2014) comprehensive guideline for categorically assessing and treating insomnia in the primary care setting, the following is an outline that can serve as a guide to obtaining a sleep history during the course of a primary care clinical interview.

**Predisposing and Precipitating Factors.**
1. Onset of the sleep difficulties and how often sleep difficulties occur;
2. Possible major cause(s) of sleep difficulties;
3. Periods of time when the patient did not experience sleep difficulties.

**Perpetuating Factors**
Understanding the patient’s behaviors prior to going to bed can provide insights into what may be contributing to the sleep problem. Important to consider are:

1. Alcohol consumption and tobacco use (frequency, amount, and time of day);
2. Caffeine consumption including coffee, soda, tea, and chocolate (frequency, amount, and time of day);
3. Stimulating activities, such as exercise, close to bedtime;
4. Computer, cell-phone, tablet use prior to bedtime and in bed;

Assessing the patient’s behavior in bed is similarly important. Factors to consider include:

1. Time to bed versus the time lights are turned off for sleep;
2. Difficulty falling or staying asleep;
3. Number of awakenings after falling asleep and how long they last;
4. Watching television in bed or leaving the television on while sleeping;
5. Any activities other than sleep and sex that are performed in bed (e.g., paying bills, worrying);
6. Sleep related thoughts and frustrations;
7. Use of alarm clock.

**Sleep hygiene.** The physical environment in which the patient sleeps is also important to consider. Among the common variables that might contribute to sleep disturbance are noise, room temperature, bedding, lighting, and bed partner’s behavior.

**Daytime behavior and functioning.** Behaviors, activities, and events that occur during the day can also contribute to a sleep disturbance. Factors to consider include:

1. Napping (duration, time of day, planned/unplanned);
2. Work (shift work, responsibilities);
3. Lifestyle (sedentary/active, light exposure, exercise);
4. Travel (change in time zones);
5. Daytime dysfunction, as evidenced by:
   a. Fatigue and sleepiness. Fatigue (low energy, physical tiredness) is a more commonly reported symptom for people with insomnia than sleepiness.
   b. Mood disturbance. Common symptoms include irritability, anhedonia, mild depression, and anxiety.
c. Cognitive difficulties. Common complaints include difficulty concentrating, remembering, and focusing attention.

**Quality of life.** How is the patient’s overall functioning in areas of work, school, and family? Interpersonal difficulties may arise due to the impact of poor sleep on a patient’s emotional and physical functioning. Patients may also experience a decrease in certain activities including socializing, exercising, intimacy, and work, to name a few.

**Medical conditions and medications.** An understanding of current medications and medical complications will help determine other factors that may be contributing to sleep disturbances. During this conversation, it is also important to consider and rule-out other sleep disorders including sleep apnea (central or obstructive), restless leg syndrome (RLS), periodic leg movements, and narcolepsy.

**Psychiatric concerns.** Screening for mental health concerns is important given the intimate relationship between sleep and mental health. Symptoms associated with anxiety, depression, and posttraumatic stress disorder are often reported in the context of primary care and all include impaired sleep as part of the symptom profile.

### Self-Report Screening Measures

Self-report measures are another important tool in the evaluation and assessment of sleep disorders. Either used to enhance a clinical interview or as screening tools, these instruments can help identify patients with sleep problems as well as track baseline and posttreatment changes (Edinger et al., 2015; Goodie & Hunter, 2014). Self-report measures listed here can be easily administered within the primary care setting.

**Epworth Sleepiness Scale (ESS)**
The ESS (Johns, 1991) is an eight-item, self-report questionnaire used to assess subjective daytime sleepiness or sleep propensity in adults and identify comorbid disorders of sleepiness. Patients rate their level of sleepiness on a scale of 0–3 (0 = none, 3 = high chance of dozing) during eight daily activities. Higher scores indicate greater levels of daytime sleepiness (score range 0–24; normal < 10). It has a high level of internal consistency as measured by Cronbach’s alpha (0.88) (Johns, 1992) and validity as measured by significant correlation with sleep latency measured during the day with a Multiple Sleep Latency Test (Johns, 1991). The ESS has been studied in the primary care setting, in which it has been determined that although the measure captures sleepiness, it is not the best predictor of moderate or severe sleep disorder breathing (i.e., sleep apnea). Overall, it is considered a reliable and internally consistent measure (Johns, 1992) with inconsistent and specificity values (Miller & Berger, 2016; Silva, Vana, Goodwin, Sherrill, & Quan, 2011).

**Pittsburgh Sleep Quality Index (PSQI)**
The PSQI (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989) is a 24-item, self-report measure that assesses sleep quality during the past month (poor sleep: global score > 5). Higher scores are a reflection of poorer sleep quality. The PSQI has a high test-retest reliability (.87) and homogeneity (Cronbach’s alpha of .85). It also has good validity in particular for patients with primary insomnia as demonstrated by high correlations between PSQI and sleep log data ($r = .81$) (Backhaus, Junghanns, Broocks, Riemann, & Hohagen, 2002). A recent meta-analysis (Mollayeva et al., 2016)
supports the strong validity and reliability of the PSQI. Furthermore, Buysse and colleagues (2006) recommended that the PSQI be used in all treatment outcome studies as a means of standardizing insomnia research.

**Dysfunctional Beliefs and Attitudes About Sleep 16-Item Scale (DBAS-16)**

The original DBAS (Morin, 1994) is a 30-item, self-report scale designed to assess sleep-related beliefs, expectations, attributions, and attitudes. Higher scores on the DBAS reflect more dysfunctional beliefs about sleep and misattributions of the consequences of insomnia. It is thought that such thinking can contribute to the maintenance of insomnia. The DBAS has reported good internal consistency and has demonstrated measurement sensitivity to cognitive behavioral treatment (Espie, Inglis, Harvey, & Tessier, 2000). The DBAS-16 represents a more streamlined version of the original DBAS with subscales assessing expectations for sleep, worry/helplessness about sleep, consequences of insomnia, and beliefs about the importance of sleep medication. Similar to its parent measure, the DBAS-16 is a reliable (Cronbach’s alpha = .821) and a valid tool demonstrated by receiver operating characteristic (ROC) curve analyses (area under the curve, or AUC = .86) (Carney et al., 2010). The DBAS-16 is now more commonly used in the literature than the DBAS (Morin, Vallières, & Ivers, 2007).

**Consensus Sleep Diary**

The Consensus Sleep Diary (Carney et al., 2012) is a 2-week sleep log to identify general patterns of sleep-wake times and day-to-day variability. A consensus diary was recently developed with the goal of standardizing a basic assessment tool that is used across research and clinical settings. The Consensus Sleep Diary was developed with the input of an expert panel, focus groups, and lexile analyses. Measures derived from diaries include bedtime, sleep latency (SL), number of awakenings, wake time after sleep onset (WASO), time in bed (TIB: time from bedtime to getting out of bed), total sleep time (TST: time in bed minus SL and minus WASO), sleep efficiency (SE: \( \frac{TST}{TIB} \times 100\% \)), total wake time (TWT = SOL + WASO), medication use (frequency, time, and dosage) and nap times (frequency, times, and duration; Edinger et al., 2004; Schutte-Rodin et al., 2008).

**Global Sleep Assessment Questionnaire (GSAQ)**

The GSAQ (Roth et al., 2002) is an 11-item, self-report questionnaire used to screen for multiple sleep disorders including symptoms associated with insomnia, OSA, restless leg syndrome/periodic limb movement, and parasomnias. The measure also assesses mood, life activities, and medical issues as they relate to sleep. Circadian rhythm disorders and narcolepsy are not directly measured; however, the header does include a question about shift work and jet lag. The GSAQ has demonstrated acceptable test-retest reliability (range: .51–.92). Analyses reveal the GSAQ differentiates insomnia from OSA, PLM, and parasomnias with a sensitivity of 79% and specificity of 57% (Roth et al., 2002). The GSAQ is on a single page and has recently been identified as the most suitable instrument for use as a screener for primary care (Klingman et al., 2016).

**PROMIS Short Form Sleep Disturbance (SD) and Sleep-Related (SRI) Item Banks**

The PROMIS (Patient-Reported Outcomes Information System) Sleep Disturbance (SD) and Sleep-Related Impairment (SRI) Item Banks (Buysse et al., 2010) were designed to improve self-report
measures of sleep–wake function. Based on the model used for the broader PROMIS initiative, these item banks were developed through an extensive process of “literature reviews, expert consensus, qualitative research methods, classic test theory (CTT) methods and, item response theory (IRT) analyses” (Buysse et al., 2010, p. 782). The PROMIS SD and SRI have been shown to have “excellent” measurement properties and are anticipated to be useful assessment instruments. (For more information about the development and validation of the PROMIS Item Banks for sleep and about the larger PROMIS initiative, see Buysse et al. 2010).

Short Forms (eight items) of the full SD and SRI item banks have shown to correlate strongly with the original instrument (Yu et al., 2012). While these related measures only contain eight items each, they have been shown to correlate strongly with their longer counterparts and to have greater precision than the PSQI and the ESS.

**Insomnia Severity Index (ISI)**
The ISI (Bastien, Vallières, & Morin, 2001) is a seven-item rating scale used to assess a patient’s perception of the severity of daytime and nighttime symptoms of insomnia, yielding a global score of sleep impairment. Patients rate sleep difficulty in terms of its severity, degree of interference with daily functioning, noticeability of such impairment to others, level of distress, and overall satisfaction with sleep. The following scoring guidelines can be used to qualify insomnia severity: score of 0–7 (no significant insomnia), 8–14 (subthreshold insomnia), 15–21 (moderate insomnia), and 22–28 (severe insomnia). This scale has good internal validity and appropriate test-retest reliability over a 2-week interval. A recent study examining psychometric indicators of the ISI concluded that the ISI is a reliable instrument with high internal consistency in a community and clinical sample (Cronbach alpha of .90 and .91, respectively). IRT analyses demonstrated adequate discriminatory validity and convergent validity was supported with significant correlations between ISI and measures of fatigue, quality of life, anxiety, and depression (Morin, Belleville, Belanger, & Ivers, 2011).

**Patient Health Questionnaire-9 (PHQ-9)**
The PHQ-9 (Kroenke, Spitzer, & Williams, 2001) is a nine-item self-report measure used to assess depression symptoms for diagnostic and treatment monitoring purposes. Many primary care offices use this measure to screen for depression. A recent study (MacGregor, Funderburk, Pigeon, & Maisto, 2012) found that Item #3, which asks the patient to rate how often they have “Trouble falling or staying asleep or sleeping too much,” was highly correlated with the ISI. These findings suggest that the PHQ-9 can provide not only information about depression but may also be an efficient means of screening for insomnia or hypersomnia. The PHQ-9 was developed from the diagnostic criteria for major depressive disorder in the *Diagnostic and Statistical Manual of Mental Disorders, fourth edition* (American Psychiatric Association, 2000). The PHQ-9 is considered a valid and reliable measure of depression (Kroenke et al., 2001) and is addressed in more detail in Chapter 9 of this book.

**Assessment of Obstructive Sleep Apnea**
The gold standard for diagnosis of sleep apnea is the overnight polysomnography (PSG), which is conducted by highly trained specialists in a specialized sleep center. Given the current rates of OSA, PSG is neither a time-efficient nor cost-effective method of screening for OSA; thus, more practical approaches should be considered within primary care. A recent integrative review of studies evaluating the screening and assessment of OSA in primary care settings determined that the STOP-Bang and
Assessment of Sleep Disorders

Berlin questionnaires serve as the current best measures to predict the presence of moderate to severe OSA (Miller & Berger, 2016).

**Berlin Questionnaire**

Designed by primary care and pulmonary physicians, the Berlin Questionnaire (Netzer et al., 1999) includes three symptom categories used to predict the presence of OSA. The first category has five questions addressing frequency of snoring and apneas while sleeping. The second category has four questions addressing daytime functioning and sleepiness. The third category has two questions focused on blood pressure and BMI. Patients are classified into High Risk or Low Risk based on their responses to individual items and overall scores in symptom categories.

The Berlin has been validated within primary care (Netzer et al., 1999) and has been shown to have higher levels of sensitivity compared to specificity (86% and 77%). Researchers note certain limitations including a complicated scoring system and its less than ideal psychometric properties (Chung, 2011; Chung, Abdullah, & Liao, 2016; Miller & Berger, 2016).

**STOP-Bang**

A shorter and more straightforward questionnaire, the STOP-Bang (Chung et al. 2008a) includes a four-item questionnaire (STOP) and a four-item portion informed by demographics and measures. It was initially developed to screen for OSA in preoperative surgical patients for which it has good sensitivity (Chung et al., 2012) but relatively low specificity. It is suggested that the STOP-Bang may be useful in primary care settings where it would be used for widespread screening (Boynton et al., 2013). Results from the Sleep Heart Study demonstrate the sensitivity of the STOP-Bang score to be 89% to detect moderate to severe OSA and 93% to detect severe OSA (Silva et al., 2011).

**Actigraphy**

Actigraphy is a wrist monitor that records both intensity and frequency of movement to determine sleep disturbance and fragmentation throughout the night (Lichstein et al., 2006). Based on difference in movements associated with wakefulness and sleep, actigraphy can provide an estimate of sleep–wake data. While actigraphy does not track all of the physiological data that a PSG captures, it is considered a more efficient and less invasive alternative. It is also considered a good alternative for patients who have difficulty completing sleep diaries. Data available include SE, SOL, TST, and WASO, which are computed using the software analysis program provided with the device. Actigraphy has been shown to have reasonable reliability and validity in normal individuals with relatively good sleep patterns (Sadeh, 2011), but is inconsistent in individuals with poor sleep. Overall, actigraphy is considered a decent tool to be used in conjunction with complementary assessment approaches.

**Assessment Summary**

A comprehensive and efficient assessment such as the one described in this section is critical in developing a treatment plan for primary care patients presenting with sleep disorders. Clinical interviews as well as subjective and objective findings can help tailor treatment to meet the needs of patients as well as inform important referral information. Such an approach also emphasizes to the patient the importance of sleep and its impact on emotional and physical functioning. Treatment should not proceed without the completion of such an assessment.
The remainder of this chapter will focus on treatment approaches that address factors identified in the assessment that are believed to perpetuate patient’s quality of sleep.

**OVERVIEW OF COGNITIVE BEHAVIORAL THERAPY FOR INSOMNIA**

As noted previously, CBT-I is the recommended first line of treatment for chronic or episodic insomnia, according to the American Psychological Association, American Academy of Sleep Medicine, and American College of Physicians (2016). CBT-I is a nonpharmacological, multicomponent treatment designed to combat thoughts and behaviors that perpetuate symptoms of insomnia. Traditionally, CBT-I is administered in six to eight 50-minute sessions by a mental healthcare professional (Kaplan & Harvey, 2014). However, recently, CBT-I has been effectively adapted for a shorter administration (one and two sessions) and for treatment across a range of settings, patient populations, and providers (e.g., Buysse et al., 2011; Edinger & Sampson, 2003; Morgenthaler et al., 2006; Rybarczyk, Lund, Garroway, & Mack, 2013). Although CBT-I’s protocol is flexible, it contains several core components that remain integral throughout any CBT-I administration, including psychoeducation, sleep restriction, stimulus control, sleep hygiene, cognitive components, and relaxation training.

**Sleep Education**

Sleep education provides patients with CBT-I’s treatment rationale, increases their self-efficacy, and promotes healthy sleep habits. Sleep education includes basic information about the sleep system, as well as information about how health behaviors and environmental factors impact sleep. Clinicians may provide a general overview of the biological processes that influence sleep, often citing Borbély’s Two Process Model of Sleep Regulation (1982), wherein sleep–wake patterns are regulated by the homeostatic drive and the circadian rhythm. Patients will be asked to manipulate these two drives throughout their treatment, so a rudimentary understanding of the two processes can help improve treatment adherence, particularly during difficult elements of treatment like sleep restriction.

Learning about and being able to teach patients about the science of sleep is a critical component of CBT-I because the public in general has a very limited understanding of how the sleep system operates. Yet for the subset of the population with insomnia, filling in the vacuum of understanding as to how their sleep operates and why it has gone awry, is almost always a central feature of the disorder. As thinking beings, people have a dire need for predictability and controllability when it comes to something that is vital to their well-being. It is a big part of why insomnia is referred to as a 24–7 disorder, as individuals who suffer from the disorder are preoccupied on a daily basis with causes, consequences, and new remedies to try.

**Stimulus Control**

Stimulus control is grounded in classical conditioning theory, with the primary goal of strengthening the association between the bed and sleep (Sharma & Andrade, 2012). Sleep can be thought of as a “learned response,” one that is subject to environmental cues. Patients with insomnia have created
an association between the bed and lying awake at night unable to sleep, which results in an association between bedtime and hyperarousal/distress. Using stimulus control, patients can break this cycle using guidelines designed to pair the bed with sleep. Common guidelines include: (1) go to bed only when sleepy; (2) if unable to sleep for more than 15 minutes or when frustrated, get out of bed and do a relaxing activity; (3) use the bed only for sleep (e.g., watch TV or read in a different room); (4) get out of bed at a regular time each morning; (5) do not take daytime naps in bed; and (6) sleep only in the bedroom (Findley & Perlis, 2014; Morin et al., 2006).

**Sleep Restriction**

Sleep restriction limits patient’s time in bed in order to curtail sleep duration, promote better sleep quality, and reduce nightly variability in duration/quality (Spielman, Saskin, & Thorpy, 1987). By shortening sleep duration, patients experience mild sleep deprivation. This sleep deprivation “jump-starts” their homeostatic drive (i.e., “sleep drive”), thereby making it easier to fall and remain asleep. Prior to treatment, patients keep a sleep log that details their hours slept per night, awakenings, how long it takes to fall asleep at night, and sleep hygiene factors (e.g., caffeine use, exercise). The sleep log is used to calculate the individual’s average nightly sleep duration. Patients and clinicians work together to plan bed/wake times. Consistent with the aims of stimulus control, a consistent wake time is determined and is used as an anchor throughout treatment. Bedtime is determined and adjusted by restricting total sleep duration to their initial average nightly sleep duration. For example, a patient with average sleep duration of 5 hours and an ultimate desired wake time of 8 a.m. would set 3 a.m. as his initial treatment bedtime. Patients are encouraged not to nap during the day but may engage in any wake-promoting activities to ensure that they stay up until their bedtime.

Patients continue to keep a sleep log as a means of tracking quality and quantity of sleep. Sleep quality specifically refers to a patient’s sleep efficiency (i.e., nighttime sleep duration/time in bed multiplied by 100), with a targeted sleep efficiency of greater than 85% (Trauer, Qian, Doyle, Rajaratnam, & Cunningham, 2015). Patients gradually move their bedtime 15–30 minutes earlier, depending on their sleep efficiency. Together, stimulus control and sleep restriction are considered the strongest components of CBT-I; at the same time they also can be the most challenging components given the behavioral changes that are required and the sleep deprivation that is experienced.

**Sleep Hygiene**

Additionally, the clinician provides sleep hygiene education about health behaviors that influence sleep (e.g., exercise, diet, caffeine use, alcohol use) and assists the patient with making lifestyle changes as indicated. Environmental factors that influence sleep (e.g., light and noise from the TV, uncomfortable temperature, excessive ambient noise) are also reviewed, and the clinician works with the patient to identify and modify such factors. Sleep hygiene is particularly important during the beginning of treatment, but it remains critical to incorporate sleep hygiene information throughout the various phases of treatment. In the primary care setting, patients often receive a sheet on sleep hygiene as a treatment approach. However, sleep hygiene education alone is not considered to be an effective therapy for chronic insomnia and works best when integrated within stimulus control and sleep restriction (Findley & Perlis, 2014).
Cognitive Component

CBT-I also includes a cognitive component, wherein patients and clinicians work together to challenge and reframe thoughts that maintain sleep, including faulty beliefs about sleep, attentional biases, excessive worry, and safety behaviors (Espie, 2002; Harvey, 2005; Morin et al., 2006). Clinicians often use measures like the DBAS or the DBAS-16 to assess sleep myths and to correct these misconceptions about insomnia and daytime functioning (Morin, Vallières, & Ivers, 2007). For example, a patient may catastrophize daytime consequences of insufficient sleep or believe that waking up in the night is aberrant. Correcting these misconceptions normalizes the experience of occasional poor sleep, which can alleviate nocturnal distress. Clinicians will also highlight patients’ attentional biases regarding their sleep. For example, it is common for patients to overestimate nocturnal wakefulness while underestimating sleep duration (Kaplan & Harvey, 2014; Spiegelhalder et al., 2010). Further, it is common for patients to engage in counterproductive “safety behaviors” that provide a sense of control over their sleep, like leaving the TV on at night or frequent time monitoring. Finally, clinicians may challenge excessive sleep-related worries and teach relaxation strategies to reduce anxiety surrounding sleep (Morin & Espie, 2003).

In a similar vein, patients frequently fall back on correlational evidence to develop theories about triggers for a poor night of sleep, such as “I never sleep well when there is a full moon,” “whenever I eat shellfish I am guaranteed to have a bad night,” “my gut instinct is that I was not going to have a good night of sleep,” or, more broadly, “something happened to my brain during my pregnancy that ruined my ability to sleep.” One analogy that the authors frequently use for patients to help them understand why they do this are the mythologies that developed around understanding, predicting, and even controlling the weather in the prescientific world. Weather is so vital to the well-being of a society, like sleep for the individual, that it spawned myths throughout the ancient world. A similar analogy that fits for insomnia is all of the mythology that surrounded fertility in the prescientific era, another biological function vital to survival that spawned innumerable myths and superstitious interventions.

Relaxation Training

Relaxation training is an ancillary component of CBT-I that is often, although not always, included in treatment (Wang, Wang, & Tsai, 2005). Relaxation techniques are designed to reduce both psychological arousal and somatic tension, especially surrounding bedtime (Jacobs et al. 1993). Techniques vary but may include diaphragmatic breathing, mindfulness techniques, progressive muscle relaxation, or guided imagery. During diaphragmatic breathing, patients focus on slowing and deepening their breaths so that their abdomens expand during breathing, a rhythm that mimics sleep onset (Findley & Perlis, 2014).

Mindfulness techniques (e.g., body scan, grounding techniques, scheduled worry time, walking meditation) have been successfully incorporated into CBT-I treatment protocols to provide long-term reductions in patients’ stress and nighttime arousal (Ong, Shapiro, & Manber, 2009). Patients and clinicians may practice these relaxation and mindfulness techniques in session, with the expectation that the patient will continue to practice at home. More recently components of Acceptance and Commitment Therapy (ACT) have also been integrated into the CBT-I model with the goal of helping patients disengage from their sleep struggle and not become overwhelmed with their fears about not sleeping.
TREATING INSOMNIA IN PRIMARY CARE

As we have noted, despite the strong evidence and expert endorsement of CBT-I as a first-line treatment for insomnia, its uptake does not match its scientific promise in the clinical world. In the primary care setting, this is even more significant due to several barriers including a shortage of trained clinicians, limited patient and clinician time, and reimbursement issues (Fields, Schutte-Rodin, Perlis & Myers, 2013; Morin, 2015; Perlis et al., 2003). Several approaches have been proposed to help remedy this issue.

Who Can Conduct CBT-I and Other BSM Treatments?

Until recently, graduate schools and postdoctoral fellowships have housed the majority of behavioral sleep medicine training programs with much of the treatment occurring within designated sleep centers. In recognition of the CBT-I supply and demand issue, a more concerted effort is being made to offer CBT-I training, apart from BSM training, to non-sleep specialists (i.e., social workers, licensed practical nurses) in a wide variety of clinical settings, where the majority of individuals struggling with sleep difficulties seek care. Experts in the field suggest that a doctoral level background may not be necessary to provide quality care for all patients and that a CBT-I skill set could be delivered without having undergone academic BSM training (Fields et al., 2013; Manber et al., 2012). Several studies have demonstrated that master’s level practitioners (i.e., nurse practitioners) can deliver CBT-I effectively and within the setting in which they are already practicing (Bothelius et al., 2013; Buysse et al., 2011; Espie et al., 2007; Jungquist et al., 2010). The VA system is similarly developing a CBT-I training model for mental health clinicians (psychiatry, psychology, social work, nursing) who work across a variety of settings, including mental health and primary care settings (Manber et al., 2012), with the goal of meeting the needs of the patient’s sleep needs.

Stepped-Care Model

Another and related solution for the lack of accessibility of CBT-I is a stepped-care model, whereby patients are filtered to different levels of CBT-I that rise in cost and intensity according to the patient’s needs (see Figure 18.2; Espie, 2009; Mack & Rybarczyk, 2011). CBT-I is suited to a stepped-care approach due to its adaptability (Espie, 2009); there is already a body of literature supporting modified versions of CBT-I. Mack and Rybarczyk (2011) propose the following levels of CBT-I: self-help CBT-I, manualized group CBT-I, brief individual CBT-I, and sleep clinic referral. There is empirical support for each of these treatment levels.

A stepped-care model of treatment addresses the availability of CBT-I resources by first offering low intensity/highly accessible interventions (abbreviated and self-help) as an initial course of treatment. Self-help CBT-I treatments (implemented via booklet(s), video(s), audio recordings, and the Internet) have demonstrated efficacy and durability but have only been examined in a small number of studies of patients with comorbid medical disorders. Two recent meta-analyses indicate that self-help CBT is an effective treatment for insomnia (Cheng & Dizon, 2012; Ho et al., 2015). Cheng and Dizon (2012) examined studies using computerized self-help specifically and concluded that the treatment had significant effects on sleep quality, sleep efficiency, number of awakenings, sleep-onset latency, and insomnia, but not wake time after sleep onset, total sleep time, or time in bed. Moreover, adherence to this line of treatment was 78%. Ho and colleagues (2015) analyzed studies examining self-help CBT-I more broadly to show that this line of treatment reduced symptoms of insomnia;
effect sizes at posttreatment were 0.80 for sleep efficiency, 0.66 for sleep-onset latency, and 0.55 for wake after sleep onset. Treatment gains remained at 1–3 months follow-up, although more research is necessary to determine long-term efficacy. Supplementary telephone consultation appears to supplement self-help CBT-I.

Another promising self-help approach is Internet-delivered treatment. Such web-based mental health interventions are increasingly gaining clinical and research attention as a novel means of reaching those patient populations that might not otherwise have access to treatment and offering an interactive self-help alternative. In 2011, the US Department of Health and Human Services announced an initiative stressing the importance of “health internet technology” as a means of improving access to health care. Internet-delivered cognitive behavioral therapy (ICBT) has garnered empirical support in both the medical and psychiatric literature (Buhrman et al., 2013; Moock, 2014; Newby et al., 2013). The field of insomnia has similarly seen an emergence of web-based applications. The efficacy of Internet-delivered CBT-I has been examined across several controlled and uncontrolled studies with promising results (Espie et al., 2012; Gosling et al., 2014; Ritterband et al., 2012; Thorndike et al., 2013). These studies consistently demonstrate associated improvements in sleep efficiency and sleep quality in the general population. Cancer survivors also reported improved sleep after participating in ICBT-I (Ritterband et al., 2012). Furthermore, in a pilot randomized controlled trial, Thorndike and colleagues (2013) observed improvements in secondary comorbid psychological symptoms, including depression, anxiety, and mental health quality of life, associated with participation in Internet-delivered CBT-I. In short, while more studies are needed to comprehensively understand the effects and benefits, the literature supports the role of Internet-delivered CBT-I and its durability as an effective insomnia intervention. Several empirically supported resources include Sleepio (Espie et al., 2012), SHUTi (Thorndike et al., 2013) and CBT-i Coach (Kuhn et al., 2016), to be used in conjunction with a skilled clinician.

As discussed earlier in the chapter, despite the ubiquity in which sleep hygiene sheets are handed out in the primary care setting, the empirical literature does not support it as a monotherapy in clinical populations (Lacks & Morin, 1992). However, there are recommendations that a primary care clinician can make that are more consistent with the strongest components of CBT-I. Setting a consistent bed and wake-up time, not compensating for a poor night of sleep by going to bed early,
napping or sleeping late, and not spending a significant amount of time in bed awake are behavioral recommendations that have been identified as effective first-tier interventions (Findley & Perlis, 2014).

**Modified CBT-I (Abbreviated and Group)**

Brief behavioral treatment for insomnia (BBTI) capitalizes on modifying waking behaviors as a means of having a direct impact on the two processes that regulate sleep: homeostatic and circadian. Components of BBTI include sleep and sleep hygiene education, sleep restriction, and stimulus control. Manualized BBTI treatment is typically four sessions (Troxel, Germain, & Buysse, 2012) with the option to conduct some of these sessions over the phone. Others have proposed even fewer sessions with similar outcomes (Edinger & Sampson, 2003; Ellis, Cushing, & Germain, 2015).

Several studies support the efficacy of abbreviated CBT-I. Edinger and Sampson (2003) demonstrated that within a primary care setting, two 25-minute sessions of CBT-I with supplemental take-home materials reduced symptoms of insomnia, as compared to sleep hygiene suggestions alone (2003). While promising, this study only followed participants for 3 months; additional follow-up is necessary to determine if CBT-I remains durable when so dramatically compressed. Buysse and colleagues (2011) developed and evaluated a brief version of CBT-I that led to greater improvement in self-reported sleep, sleep diary, and actigraphy but not polysomnography, as compared to instructions to read three publications from the American Academy of Sleep Medicine. Effects remained at 6 months follow-up. Finally, Wagley and colleagues (2013) showed that a two-session CBT-I intervention reduced insomnia symptoms in psychiatric outpatients, though effects were not significant when compared to treatment as usual. CBT-I has also been shown to be effective when delivered as a group intervention resulting in significant improvements in sleep continuity and sleep quality that are maintained over time (Koffel, Koffel, & Gehrman, 2015).

**Unabridged CBT-I**

At the top of the stepped-care model is psychologist-delivered unabridged CBT-I. This treatment is discussed in depth in the previous section. Formal CBT-I training is typically conducted within psychology internship, medical residency, or sleep medicine fellowship programs. Outside of these degree programs, a growing number of training opportunities are being offered through fellowship programs, the Veterans Affairs (VA) CBT-I dissemination program, and a variety of continuing education and continuing medical education programs offered through universities, AASM, Society of Behavioral Sleep Medicine (SBSM), and private educational programs (Fields et al., 2013).

**Stepped-Care Research Summary**

This research collectively indicates that CBT-I can be modified to reduce costs while remaining an effective treatment for chronic insomnia. There is also substantial evidence that healthcare providers outside of psychology can provide CBT-I. Nurses, trainees, and social workers, typically without prior experience in sleep medicine, successfully administered CBT-I in several studies (Bothelius et al., 2013; Epstein, Sidani, Bootzin, & Belyea, 2012; Espie et al., 2007; Espie et al., 2008; Lichstein, Wilson, & Johnson, 2000). This alternative narrows the gap between the demand for behavioral sleep medicine and the supply of credentialed behavioral sleep medicine specialists (Perlis & Smith, 2008).
Research is necessary to assess patient outcomes and clinic efficiency after the implementation of a stepped-care approach. Vincent and Walsh (2013) examined the implementation of a stepped-care model in an outpatient behavioral sleep medicine clinic. Care options included a 6-week computerized treatment, a single-session consultation, 6-week group CBT-I with 90-minute sessions, and individual psychotherapy. The authors posit that the approach improved service efficiency by 69%, although more detail is necessary regarding the definition of service efficiency as well as how precisely this estimate was calculated. Further study is necessary to assess patient attitudes toward stepped-care, predictors of treatment success at different levels of care, and healthcare costs (short term and long term).

CASE STUDIES

To illustrate how a BSM can be applied within the primary care setting, two case studies are presented here.

Case 1

Ms. H is a 47-year-old female who initially presented to her primary care physician with complications related to gastroesophageal reflux disease (GERD). She had been experiencing a worsening of symptoms over the past three months. The patient reported that her symptoms had started to interfere with her daytime functioning, including concentration at work and ability to enjoy time with her husband and three kids. As an aside, the patient mentioned that the GERD seemed to be exacerbating her sleep difficulties, for which this PCP was already prescribing trazodone (50 mg). After ruling out all other medical issues, her PCP wrote her a prescription for Prilosec and then initiated a handoff with the behavioral health psychologist to address the worsening of sleep issues in addition to issues related to her daytime functioning. In this initial visit with the psychologist, a brief sleep history was conducted, sleep questionnaires were administered and the patient was given 2 weeks of sleep diaries to complete.

Sleep History

The sleep history revealed that Ms. H had been struggling with sleep difficulties since the birth of her 9-year-old daughter. She also attributed worsening of her sleep during a 3-year period in which she worked the night shift as a nurse in the emergency department. The patient reported that she had tried several remedies for her sleep difficulties including Ambien. She noted that while the trazodone seemed to help somewhat with her falling asleep, she still had difficulty staying asleep. Patient reported her sleep–wake schedule includes a bedtime of 10 p.m. and that falling asleep can be immediate or “up to an hour” with trazodone (and “hours” without medication). She additionally reported four awakenings per night and that over the course of an evening she is awake for at least one hour. Ms. H notes that she also wakes up between 3:30 a.m. and 5 a.m. and often has difficulty falling back asleep. If she does fall back asleep, she wakes up at 6:00 a.m. on weekdays and 8:00 a.m. on the weekends.

Impact on Daytime Functioning

As noted with her PCP, patient described feeling fatigued, decreased energy, difficulty concentrating at her job as a nurse in a hospital setting, and increased irritability, especially with her family.
Maladaptive Behaviors and Perpetuating Factors

Patient noted that when she gets to bed, she reads on an electronic device while watching television. She also reported that when she turns the lights out, she thinks about her patients and worries about the care that she provided. Patient also reported drinking two sodas in the late afternoon to keep her alert and two glasses of wine close to bedtime to help with her sleep.

Psychiatric History

Ms. H presented with anxious affect and reported current symptoms of anxiety, including feeling stressed at work and at home. She also reported that at bedtime her mind is racing. She denied current suicidal ideation/intent/plan.

Assessment Data

Sleep: Ms. H completed the Insomnia Severity Index on which her score reflects severe insomnia. Her level of sleepiness as measured by the ESS was a 0. Her reported quality of sleep as measured by the PSQI was poor. Based on the DBAS-16, she may have unrealistic expectations for sleep or her ability to cope with sleep loss may have become a factor in her sleep problem.

Mood: Patient completed the Generalized Anxiety Disorder seven-item (GAD-7) scale on which she endorsed symptoms consistent with moderate anxiety. Patient was also given the PHQ-9 on which her scores did not reflect elevated levels of depression.

Conceptualization: Ms. H reported a 9-year history of difficulty initiating and maintaining sleep accompanied by impact on her daytime functioning and mood. Her insomnia appears to be precipitated by both the demands of new motherhood and working a night-shift position for 3 years. Her sleep problems appear to be perpetuated by a number of factors including reading her Kindle and worrying in bed. Ms. H’s alcohol and caffeine use close to bedtime likely contributes to the maintenance of her sleep problems.

Treatment

Education about sleep was provided to establish rationale for treatment. Because Ms. H worked at a hospital close to the clinic, she was able to commit to returning for four sessions of treatment. During this initial visit, patient was provided with the Consensus Sleep Diary and scheduled to return to the clinic in 2 weeks.

Sessions 1–4

Ms. H returned 2 weeks later. After calculating her sleep diaries, a clearer and more specific picture of her sleep/wake schedule emerged. On average, she reported that it took her about one and a half hours to fall asleep (SL), approximately two hours awake after falling asleep (WASO), and an average sleep efficiency (SE) of 55%. Patient’s average total sleep time (TST) was 4.5 hours and total time spent in bed was 8 hours.

At the end of session, Ms. H set the following goals:

1. Set a consistent wake-up time: 5:45 a.m.
2. Sleep schedule: 11:45 p.m.–5:45 a.m.
3. If not tired, do not go to bed, and if awake in bed for more than 15 minutes or when she gets frustrated, get out of bed.
4. Decrease caffeine and alcohol use close to bedtime.
5. Start reading a book in bed and stop using Kindle and watching TV in bed.
6. Schedule a worry time to think about patients during the day.

Over the course of three more sessions, Ms. H continued to complete sleep diaries and engaged in CBT-I treatment including sleep restriction, stimulus control, and mindfulness techniques to address her worry both during the day and at night. Of note, she also decided to go off her medication and rather than do a slow taper, she stopped taking her trazodone altogether at the onset of treatment.

**Results of Treatment**

Over the course of four sessions, Ms. H made significant improvement in enhancing her quality of sleep as evidenced by better sleep efficiency and sleep quantity. Patient also described feeling less fatigued and less irritable during the day. Figures 18.3–18.5 demonstrate patient’s improvements throughout treatment.

![Figure 18.3 Case Study 1 Sleep Latency (SL) and Wake After Sleep Onset (WAS)](image1)

![Figure 18.4 Case Study 1 Sleep Efficiency (SE) Throughout Treatment](image2)
Case 2

Mr. M is a 62-year-old male who presented to his primary care physician with continued difficulty with his sleep. He had been diagnosed with obstructive sleep apnea (OSA) and restless leg syndrome (RLS) over the past year. He was using CPAP and taking Requip 0.5 mg for the respective sleep disorders. The behavioral health psychologist was consulted to help get a better understanding of why this patient would continue to have difficulties with his sleep given that he was following the proper treatment for OSA and RLS.

Sleep History

Mr. M reported onset of sleep difficulties about two years prior and reported that at the same time he experienced several stressors, including kidney failure and the death of his wife. The patient had tried several remedies for his sleep difficulties, including Ambien, Lunesta, and melatonin. More recently he had been treated with triazolam for sleep maintenance, which he describes as being somewhat effective. He continued to take this medication on the nights before he had dialysis. He discussed use of CPAP and reported that although he puts on the mask when he goes to bed at night, he frequently notices that the mask is no longer on when he wakes up in the middle of the night.

Mr. M reported that he goes to bed at 10 p.m. and watches television in bed until he falls asleep. He states that it takes him 15 minutes to fall asleep and that he wakes up four to five times per night for about 20 minutes each awakening. He wakes up at 4 a.m. on dialysis days and 9:00 a.m. on all other days. He estimated his total sleep time to be four to five hours per night.

Maladaptive Behaviors and Perpetuating Factors

Mr. M also reported that he takes 30-minute naps every day of the week and described not being able to control falling asleep during the day. He spends time during the day in his bed and reports that at night he often gets in bed and talks on the phone and pays bills. He notes that he cannot fall asleep at night unless the television is on; however, he believes this is not a problem because he sets the timer to turn off the television after 2 hours.

Recent Medical History

Mr. M was diagnosed with kidney failure 2 years ago and now goes to dialysis on Monday, Wednesday, and Friday mornings. Patient also has a diagnosis of fibromyalgia for which he takes Cymbalta.
Daytime Functioning

He reported daytime impairment attributed to sleep difficulties, including fatigue, malaise, irritability, and some difficulty with memory.

Health Behaviors

Mr. M denies use of nicotine, alcohol, or illicit drugs. Caffeine consumption consists of one cup of coffee in the morning. Current stressors include his medical condition and continued grief over the death of his wife.

Assessment Data

Sleep: Mr. M completed the ISI on which his score reflects moderately severe clinical insomnia. His level of sleepiness as measured by the ESS was a 13. His reported quality of sleep as measured by the PSQI was poor. And based on how he responded on the DBAS-16, Mr. M reported having unrealistic expectations about his sleep (i.e., needing 8 hours of sleep to feel refreshed, inability to function without adequate sleep, sleep medication is the only cure for sleep problems).

Mood: Patient completed the PHQ-9 on which he endorsed some symptoms of depression including anhedonia and feeling down.

Conceptualization: Mr. H reported a 2-year history of difficulty initiating and maintaining sleep, accompanied by daytime sleepiness, low energy, difficulty with concentration, and irritability. Precipitating factors appear to include medical and psychosocial factors maintained by maladaptive sleep-related behaviors, including daily naps, engaging in activities in bed that are not conducive to sleep, and keeping the television on to fall asleep.

Sessions 1–3

Mr. M returned 2 weeks later. Sleep diary calculations revealed that on average it took the patient about one hour to fall asleep (SL) and that over the course of an evening he was awake for more than one hour. His reported total sleep time was 4 hours (TST) and total time spent in bed was 7 hours (TIB), which gave him a sleep efficiency of 54% (SE).

Sleep education was reviewed with patient and the following goals of treatment were established:

1. Because of Mr. M’s dialysis schedule, treatment was modified to include two sleep schedules: one for dialysis days and one for nondialysis days. His wake-up times were 3:30 a.m. and 7:00 a.m. respectively. His sleep window for both schedules was set for 5 hours for the beginning of treatment.
2. Mr. M would stop watching television in bed.
3. He would pay bills at the kitchen table.
4. He would try to stop napping by engaging in more stimulating activities during the day and incorporating exercise into his daily routine.

Results of Treatment

Over the course of three sessions spaced out over 2 months, Mr. M made significant improvement in his sleep quantity and quality even with a modified adaptation of CBT-I. At the end of treatment, he reported feeling more refreshed during the day even though he was no longer taking naps. He also reported an increased interest in being active and had stopped taking the triazolam. Figures 18.6–18.8 show Mr. M’s improvements throughout treatment.
Figure 18.7 Case Study 2 Sleep Efficiency (SE) Throughout Treatment

Figure 18.8 Case Study 2 Total Sleep Time and Time in Bed Throughout Treatment

Figure 18.6 Case Study 2 Sleep Latency (SL) and Wake After Sleep Onset (WASO)
**Case Study Summary**

These case studies illustrate the positive impact that a BSM assessment and treatment can have in the primary care setting. The patients discussed here experienced not only improved sleep quality and quantity but also improvement in their daytime functioning. These patients are not unique in how they responded to treatment and represent a typical treatment trajectory and outcomes experienced by patients who work with BSM providers in the primary care setting.

**SUMMARY**

Sleep disorders often go unrecognized in primary care settings despite their prevalence. Behavioral sleep medicine provides an exciting opportunity for primary care providers and psychologists working in primary care settings to address the needs of their patients in a way that is consistent with patient-centered goals and with best practice guidelines. This chapter provides a snapshot into the ways in which BSM can improve the assessment and treatment of sleep disorders within the primary care setting, as well as the outcomes that can occur with the integration of these empirically based approaches and conceptualizations.

**REFERENCES**


CHAPTER 19

Assessment of Eating Behavior in Primary Care
Karen E. Stewart, Autumn Lanoye, Laura Milliken, and Vanessa Milsom

A number of health conditions are treated/managed in the primary care setting that are related to appetite and eating behavior (e.g., type I and type II diabetes, obesity, metabolic syndrome, heart disease, anemia, thyroid disorders, cancer and its treatment, pregnancy, inflammatory bowel disease). Biopsychosocial influences on eating behavior are complex, and there are numerous ways in which psychologists may be called upon in primary care settings to aid patients with improving eating behavior or coping with eating-related challenges posed by their medical condition(s). Common referral problems include obesity, diabetes management, fear/avoidance of eating, and the eating disorders defined in the DSM-5: anorexia, bulimia, and binge eating disorder.

This chapter seeks to provide an overview of the diversity of eating problems that are seen in primary care and to describe methods that psychologists can use to assess and monitor these problems. Detailed treatment recommendations are beyond the scope of this chapter, and the reader is referred throughout to other resources when available for treatment recommendations. However, guidelines are provided for decision-making about discriminating between problems that can be addressed in the primary care setting or that require referral to longer term outpatient therapy and/or specialty medical/psychiatric services.

ASSESSMENT WITHIN THE PRIMARY CARE SETTING

Primary care psychologists are trained to address a wide range of mental and behavioral health concerns, including eating- and weight-related issues. Under optimal circumstances, psychologists function as an integrated part of primary care teams, delivering “co-located, collaborative care” (IOM, 1996; Zeiss & Karlin, 2008). Patients can be seen via a variety of modalities (face-to-face and telephone) and via scheduled appointments, walk-ins, or warm handoffs from the primary care provider (PCP) or other team members (e.g., RNs, LPNs, RDs). Primary care psychologists deliver population-based care and aim to see a high volume of patients daily. Appointments are brief, lasting as little as 20–30 minutes, and treatment is limited in duration, although some clinics may have the capacity to provide longer term services.

Psychologists working in the primary care setting typically conduct functional assessments and deliver time-limited, problem-focused interventions. Many serve in a triage role, and refer patients who require more intensive follow-up to specialty mental health providers, although depending on the training of the psychologist, some specialty mental health treatment can be provided in the primary care setting.

Assessments by primary care psychologists differ significantly in style and content from those conducted in traditional outpatient mental health settings. Functional assessments typically focus on one presenting problem or concern (e.g., weight loss) and understanding its impact on the patient’s quality of life (Zeiss & Karlin, 2008). Developing and maintaining rapport is essential, given that
many patients are seen intermittently and may wish to follow up monthly or even yearly in conjunction with their PCP visits. Documentation is similarly concise and tailored to the PCPs and other team members as the audience. Feedback to the PCP and treatment coordination can be conducted via the chart, email, or face-to-face discussion.

A tremendous amount of information relevant to eating and weight concerns can be gleaned via chart review alone, including the patient’s weight trajectory as well as fasting glucose and hemoglobin A1c levels. Measures designed to screen for eating disorders can be administered by the PCP or other team members as part of the patient’s routine visit paperwork, with positive screens referred to the psychologist if the patient is amenable. Information gleaned via screening measures is then integrated with the brief clinical interview, and treatment goals and plans are developed in collaboration with the patient.

EATING-RELATED PROBLEMS ENCOUNTERED IN PRIMARY CARE

Eating behavior is influenced by a range of biological, psychological, and social factors. Eating problems are at times caused by medical conditions that are seen in the primary care setting (e.g., cancer and its treatment, pregnancy, inflammatory bowel disease), and eating problems contribute to the development and course of other conditions (e.g., diabetes, obesity, metabolic syndrome, heart disease). Integrated primary care provides the opportunity to intervene in both of these scenarios, in that psychologists can help patients overcome the psychological factors that may perpetuate or exacerbate eating problems caused by health problems, and can help patients improve their eating behavior to reduce risk of negative health outcomes.

Obesity

Obesity and overweight are extremely prevalent in the United States, with one-third of adults classified as overweight (body mass index [BMI] = 25–29.9 kg/m²), one-third classified as obese (BMI 30–39.9 kg/m²), and one in 20 suffering from extreme obesity (BMI > 40 kg/m²; Centers for Disease Control and Prevention, 2016). Contributing factors to obesity are multifactorial and some of these will be problems that are beyond the scope of psychological intervention alone; thus, this is an area where multidisciplinary care models and thoughtful assessment are truly valuable. Psychologists are likely to receive a referral for weight management with little information about psychosocial factors contributing to obesity. Assessing and quantifying those factors will be helpful in treatment planning, discerning if any treatment referrals need to be made, and determining which professionals to engage in the care of the individual. For each patient referred for weight management, it would be ideal to assess in some manner each of the following areas: disordered eating (present or past history), motivation for change, health literacy/nutritional knowledge, stress/chronic stress, sleep duration/quality, emotional eating, and depression/anxiety. While it is not feasible to formally assess each of these areas in a single primary care session, it is possible to gain an appreciation for which (if any) of these factors are interfering with healthy eating behaviors through a combination of screening tools and clinical interview.

Disordered Eating

The Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) now formally recognizes binge eating disorder (BED) in addition to the long-recognized anorexia nervosa (AN) and bulimia
nervosa (BN; American Psychiatric Association [APA], 2013). Disordered eating is not always easily recognizable and it is not uncommon for individuals to “shift” between disordered eating behaviors (e.g., a current binge eater may have previously experienced episodes of bulimia and/or anorexia). Even if a patient is presenting without current disordered eating pathology, it is worthwhile to inquire about past history when working with a patient on weight management behaviors. Many of the behaviors required for weight loss (e.g., self-monitoring, counting calories, weighing regularly) resemble those of disordered eating, but in a healthy, balanced form; thus it will be important to know if the patient has a history of disordered eating behavior or underlying psychological attributes (e.g., rigid “dieting” rules, perfectionism, all-or-nothing thinking about diet and body image, body image disturbance, over-representation of body image satisfaction in overall self-esteem).

In addition to using screening instruments, it is advisable to ask questions like “has there been a time in your life when your diet and weight were on your mind constantly?” Such questions allow for a quick determination if further questions are warranted with regard to current or past disordered eating. Studies show that weight loss is related to frequent, consistent self-monitoring of food intake (Peterson et al., 2014), and frequent weighing is associated with long-term weight maintenance (Butryn, Phelan, Hill, & Wing, 2007). Thus, a primary goal of weight management interventions is helping patients find the optimal degree of vigilance to these behaviors that they can maintain in the long term without returning to previous obsessive or rigid dietary behaviors. Assessing and working with patients to modify the psychological attributes associated with disordered eating behaviors will be key in guiding them to a healthy relationship with self-monitoring and feedback tools necessary to achieve and maintain a healthy weight. This is also key in not only preventing the psychological suffering of disordered eating, but also in avoiding potential binge-restrict-purge cycles that ultimately defeat efforts at weight management and deteriorate overall health and well-being.

BED is likely to be encountered with some frequency in the primary care setting, as it is considerably more prevalent than other eating disorders and observed across the age spectrum in both men and women (Grucza, Przybeck, & Cloninger, 2007) and ethnically diverse patient populations (Jennings, Kelly-Weeder, & Wolfe, 2015). Prevalence rates within community samples are estimated at 2%–3% (Hudson, Hiripi, Pope, & Kessler, 2007), with significantly higher rates observed among treatment-seeking overweight or obese individuals. BED is characterized by the consumption of unusually large volumes of food within a 2-hour time period, accompanied by a sense of loss of control (APA, 2013). It has received increased attention in the research literature in recent years, and was added as an autonomous diagnosis to the DSM-5. However, surveys suggest that awareness of BED as a distinct eating disorder remains low among primary care physicians, which likely impacts screening and referral rates (Supina, Herman, Frye, & Shillington, 2016).

BED can be challenging to diagnose within the primary care setting. Patients with BED may be of normal weight, although the majority are overweight or obese (de Zwaan, 2001). Physical symptoms are typically limited to obesity and associated health consequences, while psychological symptoms can be difficult to detect in the early stages of the disorder. Individuals with BED typically present with a long history of failed weight loss attempts, and are more likely to seek treatment for weight loss than for mental health concerns. Distinguishing binge episodes—the hallmark of the disorder—from overeating episodes is critical for diagnostic accuracy.

Treatment-seeking rates for AN and BN are low. In one community-based sample of women with a diagnosable eating disorder, 73% recognized a problem with eating and 69% desired treatment, but only 28% had sought treatment (Cachelin & Striegel-Moore, 2006). When seeking care for disordered eating, approximately 23% seek mental health treatment, while 30%–73% seek medical treatment for weight loss (Hart, Granillo, Jorm, & Paxton, 2011).

Access to proper eating disorder treatment for African American, Hispanic, and Asian populations is limited. People from ethnic minorities with eating disorders are less likely to seek treatment and
<table>
<thead>
<tr>
<th>Instrument</th>
<th>Administration &amp; Scoring</th>
<th>Psychometrics</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Reference(s)</th>
</tr>
</thead>
</table>
| QEWP-R     | 28-item self-report measure  
Combination of yes/no & Likert response options  
Yields categorical outcome re: likelihood of BED diagnosis | Cronbach’s alpha = .75-.79 for Binge Eating Syndrome items  
Poor concordance with structured diagnostic interviews  
Sensitivity: 74%  
Specificity: 35% | Brief  
Cost-effective | Does not reflect DSM-5 criteria  
Assesses only a small subset of problematic eating patterns | Spitzer et al. (1992)  
Spitzer et al. (1993)  
de Zwaan et al. (2001)  
Yanovski, Marcus, Wadden and Walsh (2015)  
Barnes, Masheb, White, and Grilo (2011)  
cotton, Ball, and Robinson (2003)  
Hill, Reid, Morgan, and Lacey (2010)  
Luck et al. (2002)  
Morgan, Reid, and Lacey (1999)  
Mond et al. (2008)  
Pannocchia, Fiorino, Giannini, and Vanderlinden (2011)  
Perry, Morgan, Reid, Brunton, O’Brien, Luck, and Lacey (2002)  
Siervo, Boschi, Papa, Bellini, and Falconi (2005) | |
| SCOFF      | Format: 5 questions administered in oral or written form  
Time: 30 seconds  
Scoring: every “yes” response equals one point (total score: 5) | Reliability: Cronbach’s alpha = .44-.64; test-retest = .66  
Validity: Concurrent acceptable (r = .52-.73)  
Sensitivity: 78%–100%  
Specificity: 87.5%–88% | Brief  
Cost-effective  
Easy to administer  
Good response from staff and patients | Only used as a screener; administrator cannot assume diagnostic impressions | Cotton, Ball, and Robinson (2003)  
Hill, Reid, Morgan, and Lacey (2010)  
Luck et al. (2002)  
Morgan, Reid, and Lacey (1999)  
Mond et al. (2008)  
Pannocchia, Fiorino, Giannini, and Vanderlinden (2011)  
Perry, Morgan, Reid, Brunton, O’Brien, Luck, and Lacey (2002)  
Siervo, Boschi, Papa, Bellini, and Falconi (2005) | |
| EAT        | 40-item version (1979) & 26-item version (1982)  
6-point Likert scale assesses frequency of symptoms  
3 most extreme responses for each item are scored; scores are summed  
EAT-26 scores > 20 and EAT-40 scores > 30 require follow-up | Cronbach’s alpha = .82-.87  
Correlated with estimates of actual & ideal body size, body dissatisfaction (r = .38-.57)  
Sensitivity:  
  o EAT-26: 77%  
  o EAT-40: 77%  
Specificity:  
  o EAT-26: 94%  
  o EAT-40: 95% | Online version available  
Cost-effective  
Relatively brief | Not normed for Western racial/ethnic minorities | Garner and Garfinkel (1979)  
Garner et al. (1982)  
Murelle, Ainsworth, Bulger, Holliman, and Bulger (1992)  
Gleaves, Pearson, Ambwani, and Morey (2014)  
Anstine and Grinenko (2000)  
Mintz and O’Halloran (2010)  
Davis, Olmsted, and Rockert (1990)  
Garfinkel and Newman (2001) |
<table>
<thead>
<tr>
<th>Instrument Administration &amp; Scoring</th>
<th>Psychometrics</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAT-26: 77%</td>
<td>Reliability: Cronbach’s alpha = .70–.93; test-retest = .66–.94</td>
<td>Brief</td>
<td>Should primarily be used as a screener; take caution when making diagnostic impressions</td>
<td>Barnes, Masheb, White, and Grilo (2011)</td>
</tr>
<tr>
<td>EAT-40: 77%</td>
<td></td>
<td>Cost-effective</td>
<td></td>
<td>Becker et al. (2010)</td>
</tr>
<tr>
<td>EDE-Q</td>
<td>Format: 36-item self-report</td>
<td>Time: &lt;15 minutes</td>
<td>Sensitivity: 60%–82.5%</td>
<td>Anstine and Grinenko (2000)</td>
</tr>
<tr>
<td></td>
<td>Scoring: 7-point scale (0–6), with scores of 4 or higher considered clinically significant</td>
<td>Specificity: 71.7%–98.5%</td>
<td>Specificity: 87.5%–88%</td>
<td>Barnes, Masheb, White, and Grilo (2011)</td>
</tr>
<tr>
<td></td>
<td>4 subscales (Restraint, Eating Concern, Shape Concern, &amp; Weight Concern); to score: add the ratings for the relevant items and divide the sum by the total number of items in the subscale</td>
<td>Reliability: test-retest = .27</td>
<td>Sensitivity: 67%–80%</td>
<td>Quick and Byrd-Bredbrenner (2013)</td>
</tr>
<tr>
<td></td>
<td>Prime-MD</td>
<td>PQ brief</td>
<td>Requires time to administer CEG in patient-centered interview format</td>
<td>Bakker, Terluin, van Marwijk, van Mechelen, and Stalman (2009)</td>
</tr>
<tr>
<td></td>
<td>Format: 2 stages, (1) 26-item self-report patient questionnaire (PQ) and (2) the Clinical Evaluation Guide (CEG) in interview format consisting of 5 diagnostic modules (only administered after positive PQ responses)</td>
<td>Reliability: test-retest = .27</td>
<td>Combined PQ &amp; CEG can make diagnostic impressions</td>
<td>Hahn, Kroenke, Williams, and Spitzer (1999)</td>
</tr>
<tr>
<td></td>
<td>Time: 10–15 minutes for PQ; time varies for CEG</td>
<td>Validity: Good concurrent validity</td>
<td></td>
<td>Johnson et al. (1995)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Specificity: 67%–84%</td>
<td></td>
<td>Loerch, Szegedi, Kohnen, and Benkert (2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spitzer et al. (1994)</td>
</tr>
<tr>
<td>Instrument</td>
<td>Administration &amp; Scoring</td>
<td>Psychometrics</td>
<td>Advantages</td>
<td>Disadvantages</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>--------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------------------------</td>
</tr>
<tr>
<td><strong>Disordered Eating</strong></td>
<td>• Scoring: If the one eating disorder screening item (#17) on the PQ is answered “yes,” then administer the eating disorder module of the CEG to determine if patient meets diagnostic criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fear of hypoglycemia</strong></td>
<td><strong>Hypoglycemic Fears Scale-II</strong></td>
<td>• Format: 33 self-report items with Likert scale</td>
<td>• Reliability: Cronbach’s alpha  &gt; .85</td>
<td>• Responsive to treatment</td>
</tr>
<tr>
<td></td>
<td>• Time: 15 minutes</td>
<td>• Validity: good concurrent and divergent validity</td>
<td>• Cost-effective</td>
<td>• Short form needs further validation research</td>
</tr>
<tr>
<td></td>
<td>• New short form cuts down to 11 items</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Scoring: Mean item responses (0–4) for three scores: Total score, Behavior subscale, and Worries subscale scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nutritional literacy</strong></td>
<td><strong>Newest vital sign</strong></td>
<td>• Format: Nutrition label with six standard questions</td>
<td>• Reliability: Cronbach’s alpha = .76 in English, .69 in Spanish</td>
<td>• Brief and simple to administer</td>
</tr>
<tr>
<td></td>
<td>• Time: 3 mins</td>
<td>• Validity: good criterion validity with measure of overall health literacy ($r = .59$)</td>
<td>• Can be administered by any staff member</td>
<td>• Available in English and Spanish</td>
</tr>
<tr>
<td></td>
<td>• Scoring: &lt; 4 correct answers strongly indicative of poor nutritional literacy</td>
<td></td>
<td>• Psychometrics for Spanish language version less strong than English language version</td>
<td>• Available in English and Spanish</td>
</tr>
<tr>
<td>Instrument</td>
<td>Description</td>
<td>Advantages</td>
<td>Disadvantages</td>
<td>Reference(s)</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td><strong>Disordered Eating</strong></td>
<td>• Scoring: If the one eating disorder screening item (#17) on the PQ is answered “yes,” then administer the eating disorder module of the CEG to determine if patient meets diagnostic criteria.</td>
<td>Reliability: Cronbach’s alpha &gt; .85. Validity: good concurrent and divergent validity. Responsive to treatment. Cost-effective.</td>
<td>• Scoring: Mean item responses (0–4) for threescores: Totalscore, Behaviorsubscale, and Worries subscalescores. Costs associated with treatment outcomes.</td>
<td>Cox, Irvine, Gonder-Frederick, Nowacek, and Butterfield (1987) Gonder-Frederick, Schmidt, Vajda, Greear, Singh, Shepard, and Cox (2011)</td>
</tr>
<tr>
<td><strong>Fear of hypoglycemia</strong></td>
<td>Fear of hypoglycemia is a common fear among individuals with diabetes. It involves a feeling of panic and anxiety when the body does not have enough sugar in the blood.</td>
<td>Reliability: Cronbach’s alpha &gt; .85. Validity: good concurrent and divergent validity. Responsive to treatment. Cost-effective.</td>
<td>• Scoring: Format: 33 self-report items with Likert scale. Time: 15 minutes. New short form cuts down to 11 items.</td>
<td>Once validated, it is a cost-effective method to assess fear of hypoglycemia.</td>
</tr>
<tr>
<td><strong>Nutritional literacy</strong></td>
<td>Nutritional literacy is the ability to understand and apply nutrition information to improve health.</td>
<td>Reliability: Cronbach’s alpha = .76 in English, .69 in Spanish. Validity: good criterion validity with measure of overall health literacy (r = .59). Brief and simple to administer.</td>
<td>• Scoring: &lt; 4 correct answers strongly indicative of poor nutritional literacy. Available in English and Spanish.</td>
<td>Psychometrics for Spanish language version less strong than English language version.</td>
</tr>
<tr>
<td><strong>Emotional eating</strong></td>
<td>Emotional Eating Scale includes 25 items, each representing a mood state. Participants rate their desire to eat while experiencing each mood (1 = no desire to eat, 5 = overwhelming urge to eat).</td>
<td>Reliability: Cronbach’s alpha = .81 total scale, .72-.78 for subscales; 2-week test-retest reliability r = .79. Validity: good construct, discriminant, and criterion validity including associations between changes in EES scores and response to treatment for binge eating</td>
<td>Brief, easy to administer and score. Associated with treatment outcomes.</td>
<td>Meaning of total scale score is uncertain; recommended to use subscale scores only.</td>
</tr>
</tbody>
</table>
are less likely to be diagnosed and referred for treatment compared to Caucasians (Sinha & Warfa, 2013). Primary care-based interdisciplinary care models may thus represent opportunities to bridge this “gap” in treatment for nonwhite patients. Of particular concern, 42% of women treatment seekers reported that their initial treatment contact had made them unwilling to seek help again in the future (Cachelin & Striegel-Moore, 2006). This underscores the importance of proper diagnosis and treatment when disordered eating is encountered in the primary care setting, particularly for minority patients who may be less inclined to seek out or have limited access to mental health services.

Medically stable individuals with disordered eating can be treated in the primary care setting, ideally with a collaboration between physician, nutritionist, and psychologist. Cognitive behavioral and interpersonal therapies have demonstrated efficacy for disordered eating (Williams, Goodie, & Motsinger, 2008). Patients with life-threatening medical complications or who fail to improve with outpatient treatment should be considered for inpatient treatment, and psychologists are encouraged to develop connections with specialty eating disorder clinics and hospitals in their region. Many of these specialty treatment centers offer inpatient as well as intensive outpatient care options. Screening measures for disordered eating are described later in the chapter and in Table 19.1.

Low Health Literacy/Nutritional Knowledge

Psychologists may encounter patients who have limited health literacy or nutritional knowledge. Health literacy has been defined as the “capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” (Nielsen-Bohlman, Panzer, & Kindig, 2004). Approximately 80 million adults (36%) have low health literacy (Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011; Kutner, Greenberg, Jin, & Paulsen, 2006). Low health literacy has been associated with adverse health outcomes, poor utilization of healthcare services, increased use of emergency services and hospitalizations, underutilization of preventative health services, and increased physical and mental health problems (Berkman et al., 2011; Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011; DeWalt, Berkman, Sheridan, Lohr, & Pignone, 2010; Halverson et al., 2013; Herndon, Chaney, & Carden, 2011; Koay, Schofield, & Jefford, 2012; Wolf, Feinglass, Thompson, & Baker, 2010). Socioeconomic status (SES), education level, cultural and linguistic barriers, cognitive ability, and age can all affect low health literacy (Mottus et al., 2014; Parnell, 2015).

Lack of nutritional knowledge is a major barrier to learning healthy eating behaviors and can result in poor health outcomes (Cha et al., 2014). Research suggests that individuals with low health literacy were significantly more likely to have a higher body mass index (BMI) than those with adequate health literacy (James, Harville, Efunbumi, & Martin, 2015; Lasserter et al., 2015). In addition, in a population of patients with poorly controlled diabetes, low health literacy was associated with misperceptions regarding diabetes control (Ferguson et al., 2015). Many patients also struggle to understand food labels or interpret information related to caloric, carbohydrate, or sodium content because they have not received proper nutritional education on the importance of these topics (Cha et al., 2014). These issues become problematic in terms of making unhealthy food choices, potentially causing chronic health problems later in life, and being unable to effectively manage those health problems when they arise.

A brief nutritional literacy screening tool is described in Table 19.1 that may be helpful to psychologists either in planning treatment or determining the need for referral. While some psychologists may be comfortable educating patients on basic nutritional information (e.g., how to read a food label) patients with a need for significant or disease-specific education (e.g., diabetes) should be referred to a nutritionist or certified diabetes educator. Psychologists can coordinate care with these professionals and implement behavioral strategies to assist the patient with successfully implementing the new
knowledge. Some primary care clinics have nutritionists and/or diabetes educators on staff, but some settings may require identifying outside referral sources for patients with these needs. Psychologists are eligible to train and test for diabetes educator certificates as are many other health professions (National Certification Board for Diabetes Educators, 2016), and this may be desirable for psychologists working extensively with diabetic patients.

**Emotional Eating as Maladaptive Coping**

Short-term negative emotional states such as fear and stress are associated with a physiological response that is typically expected to result in reduced appetite and decreased food intake (Schachter, Goldman, & Gordon, 1968). For some individuals, however, these and other undesirable emotions promote an increase in eating behavior, which often occurs in the absence of hunger and leads to the consumption of foods high in fat and sugar (Bennett, Greene, & Schwartz-Barcott, 2013; Konttinen, Mannisto, Sarlio-Lahteenkorva, Silventoinen, & Haukkala, 2010; Oliver, Wardle, & Gibson, 2000). Emotional eating is theorized to serve as a maladaptive coping response by providing a means of avoidance, distraction, and temporary comfort in the face of negative affect (Heatherton, Herman, & Polivy, 1991; Spoor, Bekker, Strien, & van Heck, 2007). While the tendency to eat in response to perceived negative emotion varies as a function of both individual characteristics (Greeno & Wing, 1994) and specific emotional state (Bennett, Greene, & Schwartz-Barcott, 2013; Macht, 2008), findings consistently highlight high rates of emotional eating among those with overweight and obesity (Ganley, 1989). As such, the relationship between emotions and eating behavior is especially important to explore for these patients.

Assessing problems with emotional eating can be beneficial in treatment planning. Moreover, patients struggling with emotional eating will likely benefit from learning emotion regulation skills, relaxation methods, urge surfing skills, and developing alternative pleasurable activities, all of which are appropriate for primary care treatment delivery. An emotional eating screening tool that may be useful in primary care is described in Table 19.1.

**Chronic Stress Exposure**

It is important to consider the impact of chronic stress that may contribute to unhealthy eating patterns. Chronic stress—distinct from stress as a temporary emotional state—results in physiological effects such as the release of glucocorticoids, which simultaneously stimulate appetite and reduce appetite suppression (Sominsky & Spencer, 2014). Among a community sample of women, chronic stress has been linked to perceived lack of control over eating, binge eating episodes, perceived hunger, and greater intake of non-nutritious foods (Groesz et al., 2012). In addition, the experience of adverse life events likely to prompt a chronic stress response (e.g., death of a close friend or family member, excessive credit card debt) is associated with binge eating episodes across genders (Loth, van den Berg, Eisenberg, & Neumark-Sztainer, 2008). For patients with a high level of chronic stress, treatment of problematic eating should also include training in behavioral stress management as an alternative coping method.

While many patients may endure periods of chronic stress throughout their lives, those of low SES may be at especially high risk for these experiences (Baum, Garofolo, & Yali, 1999). Low-income primary care patients report chronic stressors at high rates; not surprisingly, these typically center on financial concerns (Scarinci, Ames, & Brantley, 1999). These stressors introduce competing demands for limited resources (both financial and cognitive/emotional), which may further reduce ability to regulate eating behavior for these patients. Low SES is also associated with other factors related to
problematic eating behavior, such as food insecurity (Castillo et al., 2012; Dinour, Bergen, & Yeh, 2007) and poor sleep quality (Chaput, 2014; Jarrin, McGrath, Silverstein, & Drake, 2013). For patients with low resources, referrals to social services, religious/civic organizations, and food pantries may be critical in improving eating behavior through the establishment of stable housing, access to healthy food, and community connectedness for coping with environmental stressors.

Insufficient Sleep

Insufficient sleep has been linked to obesity, and overeating has been proposed as a potential mechanism for this pathway (Chaput & Tremblay, 2012); thus, primary care patients with sleep problems (e.g., apnea, insomnia) or those who engage in shift work may be at increased risk for problematic eating behavior. Some evidence suggests that inadequate sleep interrupts hormonal regulation processes, resulting in increased cravings and appetite (Morselli, Guyon, & Spiegel, 2012). Further, imaging studies have demonstrated greater neuronal activation in response to food following acute sleep deprivation, reflective of stronger associations with reward pathways (Benedict et al., 2012; St-Onge et al., 2012). Despite these findings, the association between shortened sleep and increased energy intake is inconsistent, which suggests that individual characteristics may also play a role. Research indicates that those with a disinhibited eating style may be particularly vulnerable to the effects of insufficient sleep (Chaput, Despres, Bouchard, & Tremblay, 2011). Taken together, these findings suggest that sleeping habits should be assessed and, if necessary, treated with a behavioral approach for those primary care patients presenting with problematic eating patterns. Chapter 18 of this book provides guidance on assessing sleep problems in primary care settings.

Other Types of Eating Problems

Fear of Hypoglycemia

Severe hypoglycemic episodes can be life-threatening and are characterized by intensely unpleasant physiological and psychological symptoms. Early signs of mild hypoglycemia include shakiness, dizziness, sweating, hunger, irritability, anxiety, and headache. More severe episodes can lead to clumsiness, muscle weakness, difficulty speaking/slurred speech, blurry/double vision, drowsiness, confusion, convulsions/seizures, loss of consciousness, and even death (American Diabetes Association, 2016). Symptoms generally begin to emerge at blood sugar levels below 70 mg/dl, although individuals with chronically high blood sugar levels may begin to experience symptoms at higher blood glucose levels as their bodies have adapted to higher concentrations of glucose. Nocturnal episodes can also occur.

These symptoms can develop and progress rapidly with little warning, thus some degree of vigilance about hypoglycemia is warranted for individuals at risk of these episodes. Both type I and type II diabetics can experience hypoglycemia, and there are also nondiabetic individuals with reactive or fasting hypoglycemia. Among diabetics, tighter blood glucose control (i.e., lower/normalized glycosylated hemoglobin; HbA1c) increases risk of hypoglycemic episodes. For some individuals, this vigilance can become excessive and lead to anxiety-related maladaptive behaviors that result in poor diabetes control, impaired metabolic function, and adverse health outcomes (Wild et al., 2007). Fear conditioning is likely augmented by similarities between autonomic symptoms of hypoglycemia and anxiety (Anderbro et al., 2015). Maladaptive responses can include excessively measuring blood glucose, keeping blood glucose levels high, limiting exercise, eating large snacks, and keeping others close by due to fears of being alone during an episode (Gonder-Frederick et al., 2011). High levels of
fear of hypoglycemia are associated with frequency of prior severe hypoglycemic episodes, nocturnal hypoglycemia, number of symptoms experienced during mild episodes of hypoglycemia, depression, and non-diabetes-related anxiety (particularly sensitivity to physiological sensations of anxiety as measured by the Anxiety Sensitivity Index [ASI]; Anderbro et al., 2015).

When treating individuals with problematic diabetes self-management behaviors in the primary care setting, psychologists should be prepared to assess this construct and incorporate findings into a treatment plan, particularly if the patient has a history of severe hypoglycemic episodes and nondiabetic anxiety. The Hypoglycemia Fears Survey-II (HFS-II; Cox, Irvine, Gonder-Frederick, Nowacek, & Butterfield, 1987; Gonder-Frederick et al., 2011; Grabman et al., 2016) can be used to assess this construct and is described later in the chapter and in Table 19.1. As nondiabetic anxiety, particularly sensitivity to physiological sensations of anxiety, is commonly comorbid to fear of hypoglycemia, an anxiety tracking tool such as the ASI may also be useful in monitoring improvements in general coping with physiological symptoms of anxiety. Cognitive behavioral therapy is recommended and would include skills for restructuring maladaptive beliefs, anxiety reduction, relaxation training, exposure therapy, and coping skill development (Anderbro et al., 2015). Blood glucose awareness training methods can be helpful in improving patients’ ability to recognize, predict, and appropriately prevent or respond to fluctuations in blood glucose levels (Gonder-Frederick, Cox, Clarke, & Julian, 2000).

Avoidant/Restrictive Food Intake Disorder
Avoidant/restrictive food intake disorder (ARFID) is a new diagnosis in the DSM-5 that describes eating problems that result in inadequate nutritional intake and that are not accounted for by lack of access, cultural practices, medical condition, or drive for thin body type (APA, 2013). Inadequate nutritional intake can present as either inappropriately restricted range of food choices (e.g., avoiding certain textures) or eating very small portions, possibly related to limited appetite or fear of eating following traumatic or painful experience(s) with eating. Individuals with ARFID may present with weight loss or nutritional deficiencies, and in some severe cases, individuals with ARFID may require enteral feeding or oral nutrition supplements (APA, 2013).

Eating problems and/or painful/fear-inducing gastrointestinal symptoms can be caused by numerous medical conditions (e.g., pregnancy, chemotherapy, Crohn’s disease). Through conditioning mechanisms, some individuals may develop perpetual difficulties with eating that either exceed what would be expected for their medical condition and/or last beyond the expected recovery period. These authors are not aware of any screening measures specific to ARFID, although similar to individuals with fear of hypoglycemia, the ASI may be beneficial in tracking general anxiety over physiological sensations. Cognitive behavioral therapy can be effective in treating ARFID and would typically include a combination of psychoeducation, cognitive restructuring of catastrophic interpretations of bodily sensations, anxiety management skills, and exposure methods (King, Urbach, & Stewart, 2015).

MEASURES FOR ASSESSMENT AND MONITORING TREATMENT PROGRESS

A number of assessment tools for problems in eating behavior are available, although very little research has focused on identifying which of these measures are best suited to the primary care setting. This chapter provides an overview of the existing measures and discussion of the potential benefits and drawbacks of each in the primary care setting, although more research is needed to determine whether these or other instruments best suit the needs of primary care psychologists.
General Considerations

When administering psychological measures, they generally should not be used as stand-alone diagnostic tools. It is important to combine data obtained from these measures with knowledge of the patient’s medical and psychiatric history in determining a case conceptualization and developing a treatment plan. Assessing for any medical complications, weight, motivation to recover, current or historical suicidal ideation or attempts, comorbid mental health problems, purging or restricting behavior, and environmental stressors are important to help guide diagnoses and treatment recommendations (Pritts & Susman, 2003).

Measures of Disordered Eating

In addition to screening when treating patients for weight-related conditions, the presence of other risk factors for disordered eating also may warrant screening, as not all individuals with disordered eating have abnormal body weight. Such risk factors include participation in activities that promote thinness (e.g., ballet dancing, athletics, modeling) and certain personality traits (e.g., low self-esteem, difficulty expressing negative emotions and resolving conflict, and being a perfectionist; Kreipe & Birndorf, 2000; Pritts & Susman, 2003). Eating disorders share high comorbidity with other psychological disorders, thus psychologists may also want to screen individuals with disordered eating for affective disorders, obsessive-compulsive disorder, somatization disorders, substance abuse, and some personality disorders. These can further complicate the illness and further affect the test results, so it is important to assess for these problems as well (Kreipe et al., 1995; Westen & Harnden-Fischer, 2001).

The Eating Disorder Examination-Questionnaire (EDE-Q)

The Eating Disorder Examination-Questionnaire (EDE-Q; Fairburn & Beglin, 1994) is the self-report version of the Eating Disorder Examination semi-structured interview (EDE; Fairburn & Cooper, 1993). Both are widely considered the gold standard for eating disorder assessment (Berg, Peterson, Frazier, & Crow, 2012). The EDE-Q is a 36-item self-report questionnaire focusing on the past 28 days and is scored using a 7-point scale from 0 to 6, with higher scores indicating greater eating disorder pathology symptom levels. It is administered in paper-and-pencil format and takes less than 15 minutes to complete, which makes it easy to administer in primary care settings. This questionnaire includes four subscales related to the cognitive features of eating disorders (Restraint, Eating Concern, Shape Concern, and Weight Concern), as well as items to assess specific behavioral symptoms (e.g., frequency of binge eating, self-induced vomiting; Berg, Peterson, Frazier, & Crow, 2012).

Each item of the EDE-Q was adapted from a corresponding EDE item and the same 7-point rating scale was used. The four subscales and global scores are derived in the same way as the EDE, with scores of 4 or higher on key items considered clinically significant. The EDE and the EDE-Q produce significantly correlated findings; however, the EDE-Q tends to yield significantly higher scores on the scales reflecting behavioral and cognitive features of eating disorders than the EDE (Barnes, Masheb, White, & Grilo, 2011). Both the EDE and EDE-Q produce two types of data: frequency data on the key behavioral features of eating disorders (e.g., number of episodes of behavior) and subscale scores that reflect the severity of aspects of eating disorder psychopathology (Fairburn, Cooper, & O’Conner, 2008). To score each subscale, the ratings for the relevant items are added together and the sum is divided by the total number of items in the subscale; these are reported as means and standard deviations (Fairburn, Cooper, & O’Conner, 2008).
Table 19.2a Norms from Community-Based Sample of Young Women (n = 241)

<table>
<thead>
<tr>
<th>EDE-Q Measure</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global EDE-Q</td>
<td>1.554</td>
<td>1.213</td>
</tr>
<tr>
<td>Restraint subscale</td>
<td>1.251</td>
<td>1.323</td>
</tr>
<tr>
<td>Eating Concern subscale</td>
<td>0.624</td>
<td>0.859</td>
</tr>
<tr>
<td>Shape Concern subscale</td>
<td>2.149</td>
<td>1.602</td>
</tr>
<tr>
<td>Weight Concern subscale</td>
<td>1.587</td>
<td>1.369</td>
</tr>
</tbody>
</table>


Table 19.2b Norms for US College Women (n = 1,533)

<table>
<thead>
<tr>
<th>EDE-Q Measure</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global EDE-Q</td>
<td>1.65</td>
<td>1.30</td>
</tr>
<tr>
<td>Restraint subscale</td>
<td>1.35</td>
<td>1.43</td>
</tr>
<tr>
<td>Eating Concern subscale</td>
<td>0.89</td>
<td>1.09</td>
</tr>
<tr>
<td>Shape Concern subscale</td>
<td>2.39</td>
<td>1.63</td>
</tr>
<tr>
<td>Weight Concern subscale</td>
<td>1.98</td>
<td>1.60</td>
</tr>
</tbody>
</table>

From Quick and Byrd-Bredbrenner (2013).

Table 19.2c Norms for US College Men (n = 915)

<table>
<thead>
<tr>
<th>EDE-Q Measure</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global EDE-Q</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>Restraint subscale</td>
<td>0.96</td>
<td>1.28</td>
</tr>
<tr>
<td>Eating Concern subscale</td>
<td>0.40</td>
<td>0.71</td>
</tr>
<tr>
<td>Shape Concern subscale</td>
<td>1.36</td>
<td>1.36</td>
</tr>
<tr>
<td>Weight Concern subscale</td>
<td>1.07</td>
<td>1.18</td>
</tr>
</tbody>
</table>

From Quick and Byrd-Bredbrenner (2013).

**Normative Data**

Tables 19.2a, 19.2b, and 19.2c present the normative data for the EDE-Q across studies.

**Psychometric Characteristics**

In general, the EDE-Q has demonstrated good reliability and validity across studies. The four subscales have produced acceptable to good internal consistency across four studies, with Cronbach’s alphas ranging from .70–.93 (Luce & Crowther, 1999; Mond, Hay, Rodgers, Owen, & Beumont, 2004; Peterson et al., 2007). With the second EDE-Q administered 1–14 days (mean of 4.8 days for one study) after initial administration, test-retest reliability coefficients have ranged from questionable to good across the four subscales (.66–.94) and for the behavior frequency items (.51–.92) (Luce & Crowther, 1999; Reas, Grilo, & MacHeb, 2006). In addition, the EDE-Q has demonstrated criterion validity in its ability to detect group differences; construct validity in its convergence with similar measures (good with Restraint subscale), with daily food records, and with the factor structure of the EDE; and concurrent validity with the EDE in the assessment of attitudinal features (Berg, Peterson, Frazier, & Crow, 2012; Mond, Hay, Rodgers, Owen, & Beumont, 2004).

More specifically, the EDE-Q has been shown to be consistent with DSM-5 diagnostic criteria displaying moderate diagnostic concordance ($K = .60$; Berg et al., 2012). In addition, the EDE-Q
demonstrated acceptable sensitivity (60.0%–82.5%) and positive predictive power (64%–82.4%), with the exception of binge eating disorder, and good to excellent specificity (71.7%–98.5%) and negative predictive power (73.8%–98.5%) for all DSM-5 diagnoses (Berg et al., 2012). Overall, the EDE-Q’s sensitivity, specificity, positive predictive power, and negative predictive power in detecting EDE diagnoses were equivalent or improved when using DSM-5 versus DSM-IV-TR criteria (Berg et al., 2012). These results suggest that although it should only be used as a screener, the EDE-Q is diagnostically consistent with DSM-5 eating disorder diagnoses.

Use as an Instrument for Treatment Monitoring and Outcome Assessment
The EDE-Q has been utilized to measure reliable and clinically significant change via the reliable change index (RCI) and clinical significance (CS), with the resulting index called CS/RCI. One study aimed to test the validity of the CS/RCI using change in DSM-IV eating disorder diagnosis as comparison and found that CS/RCI explained more variance in gain scores for the EDE-Q than diagnostic change (Ekeroth & Birgegard, 2014). In addition, the average agreement between diagnostic change and CS/RCI was 60% for the EDE-Q, meaning that the EDE-Q might serve as a reliable measure to assess outcome assessment in terms of a change in DSM-IV psychopathology (Ekeroth & Birgegard, 2014).

Strengths and Limitations of Use in Primary Care
Using the EDE-Q in primary care settings has more strengths than limitations. A major strength of the EDE-Q is that both staff and patients seem to like the EDE-Q because of its brevity, taking less than 15 minutes to complete. In addition, the EDE-Q is also cost-effective and easy to administer. In terms of limitations, as previously discussed, the EDE-Q should primarily be used as a screener in primary care settings, so the clinician should take caution when making diagnostic impressions. In addition, although the EDE-Q has been identified as a “global measure of eating pathology,” it is important to consider each patient’s culture when interpreting the results because of different cultural norms and operational definitions of eating disorder terminology (Becker et al., 2010).

Questionnaire on Eating and Weight Patterns-Revised (QEWP-R)
The Questionnaire on Eating and Weight Patterns-Revised (QEWP-R; Spitzer et al., 1993) is a self-report questionnaire designed to identify individuals at risk for BED, a relatively new diagnosis first introduced in the DSM-IV. The QEWP-R consists of 28 items designed specifically to map onto DSM criteria including symptomatology, distress severity, and time frame. This measure yields a categorical outcome regarding likelihood of BED diagnosis, including information to differentially rule out bulimia nervosa.

The DSM-IV (APA, 1994) outlines six criteria for a BED diagnosis: episodic overeating; loss of control; associated symptoms (e.g., eating rapidly, eating alone or in secret, feeling guilty after eating large amounts of food); marked distress; duration of at least six months with binge eating episodes occurring at least twice weekly; and exclusion of bulimia nervosa. The initial QEWP was developed by Spitzer and colleagues (1992) in conjunction with the Eating Disorders Work Group of the DSM-IV Task Force responsible for the development of these criteria. A revised version, the QEWP-R, reflected the Task Force decision to include diagnoses for both purging and nonpurging subtypes of bulimia nervosa by altering the wording of questions related to differential diagnosis. The 28 items consist of both yes/no and Likert scale response options, and respondents can be categorized into one of several eating pattern profiles based on number of diagnostic criteria met (see Table 19.3).

Initial studies conducted by Spitzer and colleagues (1992, 1993) administered the measure to large samples across a variety of settings, such as weight-related treatment programs, self-help programs
Assessment of Eating Behavior

for overeating, and the general community. Cronbach’s alpha for the eight items included in Binge Eating Syndrome ranged from .75 to .79, with each item demonstrating strong correlations with the calculated eight-item total score (Spitzer et al., 1992). Cumulated across weight control samples, prevalence of BED ranged from 28.8% to 30.1%; across community samples, prevalence ranged from 3.3% to 4.6%. Associated clinical features unique to BED—providing discriminant validity with respect to bulimia—were found to include history of severe obesity and significant history of weight cycling (Spitzer et al., 1993).

Considerations

Designed for use as a screening tool, the QEWP-R does not provide the examiner with definitive conclusions regarding eating disorder diagnoses. In fact, comparisons between the QEWP-R and structured diagnostic interviews for eating disorders (i.e., SCID) have found relatively low interrater agreement (de Zwaan et al., 2001; Dymek-Valentine, Rienecke-Hoste, & Alvery, 2004). In light of this, respondents who screen positive for BED or bulimia should be further evaluated by clinical interview. The subjective nature of constructs such as “large amounts of food” and “loss of control” may contribute to these discrepant results. Compared to the gold standard EDE semi-structured clinical interview, the QEWP-R was found to predict cases with 74% sensitivity and 35% specificity (Celio, Wilfley, Crow, Mitchell, & Walsh, 2004), thus identifying a large proportion of false negatives.

DSM-5 guidelines present slight changes to diagnostic criteria for BED, which include assessing binging episodes rather than binging days, change in episode frequency from at least twice weekly to once weekly, change in duration from 6 months to 3 months, and the addition of qualifiers for severity and course (APA, 2013). The QEWP-5 reflects these changes, revises scoring rules accordingly, eliminates questions not related to diagnostic criteria, and includes additional questions for the assessment of perceived loss of control over eating in the absence of an objective binge episode.

### Table 19.3 Eating Pattern Classifications of the QEWP-R

<table>
<thead>
<tr>
<th>Diagnostic Criteria (DSM-IV)</th>
<th>Eating Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Episodic Overeating</td>
</tr>
<tr>
<td>Frequent consumption of a large amount of food in a 2-hour period</td>
<td>x</td>
</tr>
<tr>
<td>Frequent experiences of loss of control of eating</td>
<td></td>
</tr>
<tr>
<td>At least 3 of 6 associated symptoms (e.g., eating in secret, feelings of guilt)</td>
<td></td>
</tr>
<tr>
<td>Marked distress regarding overeating or being unable to control eating</td>
<td></td>
</tr>
<tr>
<td>Frequency of episodic eating at least 2 days per week on average</td>
<td></td>
</tr>
<tr>
<td>Compensatory behaviors (e.g., purging, overexercising, fasting)</td>
<td></td>
</tr>
</tbody>
</table>

for overeating, and the general community. Cronbach’s alpha for the eight items included in Binge Eating Syndrome ranged from .75 to .79, with each item demonstrating strong correlations with the calculated eight-item total score (Spitzer et al., 1992). Cumulated across weight control samples, prevalence of BED ranged from 28.8% to 30.1%; across community samples, prevalence ranged from 3.3% to 4.6%. Associated clinical features unique to BED—providing discriminant validity with respect to bulimia—were found to include history of severe obesity and significant history of weight cycling (Spitzer et al., 1993).
(Yanovski, Marcus, Wadden, & Walsh, 2015). Additional research is needed in order to determine the psychometric properties and clinical utility of this revised measure.

**Use in Primary Care Settings**

The QEWP-R is routinely used in specialty clinics as part of a comprehensive presurgical psychological assessment with bariatric patients, but its utility in primary care has not been widely assessed. To the authors’ knowledge, only one study has assessed administration of the QEWP-R in a primary care setting, with findings suggesting that this measure is an adequate screening tool, despite discordance with gold standard clinical interviews (Barnes, Masheb, White, & Grilo, 2011). Barnes et al. raise the important distinction between the use of this measure as a research tool—which demands greater concordance with diagnostic interviews, and as a clinical screening tool—which requires only a “starting point” (2011, p. 161) for additional evaluation and potential specialist referral. The QEWP-R is available for free by contacting the authors, and requires minimal time to complete, making it an efficient tool for use in this setting. Of note, this measure only assesses a small subset of possible problematic eating patterns requiring intervention that might be seen in a primary care clinic.

**SCOFF**

The SCOFF (Morgan, Reid, & Lacey, 1999) is a quick and simple screening questionnaire for eating disorders intended for administration by any health professional, even those who are not specialists in the field (Pannocchia, Fiorino, Giannini, & Vanderlinden, 2011). It consists of five questions that form the basis for the acronym SCOFF: (1) Do you make yourself sick because you feel uncomfortably full? (2) Do you worry that you have lost control over how much you eat? (3) Have you recently lost more than one stone (14 lbs.) in a 3-month period? (4) Do you believe yourself to be fat when others say you are too thin? (5) Would you say that food dominates your life? Morgan et al. developed these five questions using focus groups that consisted of patients with eating disorders and specialists in the field.

The SCOFF can be administered in both oral and written forms, and studies have demonstrated consistency in responses between the two, with percentage agreement above 90% for each question (Perry et al., 2002). In addition, the original SCOFF was created in the UK, and that version has different wording than the one used in the United States. For example, the phrase “make yourself sick” in the UK version was changed to “make yourself vomit” in the US version (Perry et al., 2002). To score the SCOFF, every “yes” response equals 1 point and if the patient responds positively (says “yes”) to two or more questions, the administrator should conduct or arrange for further evaluation (Pannocchia, Fiorino, Giannini, & Vanderlinden, 2011).

**Normative Data and Psychometric Characteristics**

No norms were identified for the SCOFF; a standard cutoff of two or more positive responses is used in all available studies. The SCOFF has demonstrated fairly low reliability across populations. However, this low internal consistency may be due to the low number of items and/or the yes/no answer format, as its primary purpose is for screening. In a school setting, the measure demonstrated questionable test-retest reliability ($r = .66$) and unacceptable to poor internal consistency (Cronbach’s alpha = .44–.57; Leung et al., 2009). In clinical and primary care populations, the SCOFF also demonstrated poor to questionable reliability (Cronbach’s alpha = .44, .47, .64; Mond et al., 2008; Pannocchia, Fiorino, Giannini, & Vanderlinden, 2011; Siervo, Boschi, Papa, Bellini, & Falconi, 2005). Regarding validity, the SCOFF has shown to have acceptable convergent validity and good validity with both the written and oral forms (Pannocchia, Fiorino, Giannini, & Vanderlinden, 2011; Siervo, Boschi, Papa, Bellini, & Falconi, 2005).
After developing the SCOFF, Morgan et al. (1999) conducted a subsequent validation study in which a group of women with eating disorders was compared with a group of women without eating disorders (these subjects also completed the Eating Disorder Inventory and the Bulimic Investigatory Test). After setting the threshold at two or more positive answers to any of the five questions, results suggested 100% sensitivity for anorexia and bulimia separately and combined, with a specificity of 87.5% for controls (Pannocchia, Fiorino, Giannini, & Vanderlinden, 2011). In a primary care sample, the sensitivity for the SCOFF was 100% for all cases of anorexia and bulimia, but reduced to 84.6% for cases of eating disorder not otherwise specified (Luck et al., 2002). This same study found the positive predictive value to be 24.4% and the negative predictive value to be 99.3% (Luck et al., 2002).

Additionally, the SCOFF has demonstrated good diagnostic efficiency with sensitivity ranging from 78% to 100% and specificity ranging from 87.5% to 88% with the cutoff of two positive responses across studies (Cotton, Ball, & Robinson, 2003; Hill, Reid, Morgan, & Lacey, 2010; Mond et al., 2008). Moreover, the positive predictive value for the SCOFF in these studies was wider and ranged from 47.7% to 90.6% (Cotton, Ball, & Robinson, 2003; Hill, Reid, Morgan, & Lacey, 2010; Mond et al., 2008).

Considerations
Although based on DSM-IV-TR (and currently DSM-5) diagnostic criteria, it is important to recognize that the SCOFF is only used as a screening tool designed to detect a possible eating disorder (Pannocchia, Fiorino, Giannini, & Vanderlinden, 2011). Thus, after two or more positive responses, the administrator should refer the patient to more specialized treatment so the patient can be further evaluated for a true eating disorder diagnosis and obtain the appropriate treatment. (See “General Considerations” at the beginning of this section.)

Strengths and Limitations of Use in Primary Care
There are multiple strengths to using the SCOFF in primary care settings. A major strength of the SCOFF is that both staff and patients respond well to it because of its brevity, as it takes approximately 30 seconds to complete. Comprising only five questions, the SCOFF is also cost-effective and easy to administer. In terms of limitations, specifically the low reliability demonstrated across studies, as previously discussed the SCOFF should only be used as a screener in primary care settings, so the administrator should take caution to not assume diagnostic impressions.

Eating Attitudes Test
The Eating Attitudes Test (EAT) is a self-report questionnaire designed to identify individuals at risk for eating disorders. This measure has both 40-item (EAT-40; Garner & Garfinkel, 1979) and 26-item (EAT-26; Garner, Olmsted, Bohr, & Garfinkel, 1982) versions, both of which present respondents with a 6-point Likert scale ranging from Never to Always in order to assess symptom severity. The EAT was designed to minimize administrator burden and evaluate a broad range of symptoms characteristic of eating disorders. As such, it lends itself well to screening for the similarly broad range of disordered eating diagnoses found in the DSM-5.

The original EAT-40 was developed in 1979 in response to the perceived reliance on observer ratings and narrow focus of extant eating disorder (specifically, anorexia nervosa) assessments (Garner & Garfield, 1979). The authors aimed to develop an efficient tool that could be used to screen wider populations in order to increase likelihood of early intervention and thus improve treatment success. An initial factor analysis revealed seven factors: food preoccupation, body image for thinness, vomiting and laxative abuse, dieting, slow eating, clandestine eating, and perceived social pressure to gain weight. A subsequent factor analysis performed with a larger sample indicated that 26 items
loaded onto three factors—dieting, bulimia and food preoccupation, and oral control—and retained adequate psychometric properties to be administered as a short-form version of the measure, the EAT-26 (Garner et al., 1982).

The EAT is available in both pencil-and-paper and computerized forms (Murelle, Ainsworth, Bulger, Holliman, & Bulger, 1992). Scoring is similar across both the EAT-26 and EAT-40 such that only the three most extreme responses for each item are scored. Item 26 on the EAT-26 and items 1, 18, 19, 23, and 39 on the EAT-40 are scored such that responses of “Always,” “Usually,” and “Often” receive a score of 0, “Sometimes” = 1, “Rarely” = 2, and “Never” = 3. All other items on the EAT-26 and EAT-40 are scored in the reverse direction (i.e., “Never,” “Rarely,” and “Sometimes” are scored as 0; “Often” = 1, “Usually” = 2, “Always” = 3). Scores are summed across all items for a total score. Total scores greater than 20 on the EAT-26 and greater than 30 on the EAT-40 are suggestive of eating disorder and require additional follow-up.

**Normative Data and Psychometric Characteristics**

Normative data for the EAT-26 and EAT-40 is reported in Garner et al. (1982). In general, the English-language EAT has largely been administered to white female respondents; as such, little is known regarding normative scores for men and racial/ethnic minorities.

In an analysis of 15 studies using the English-language EAT-40 and 54 studies using the English-language EAT-26, both measures were found to have acceptable test score reliability using reliability generalization procedures (Cronbach’s alphas = .82 and .87, respectively; Gleaves, Pearson, Ambwani, & Morey, 2014). Both measures demonstrate significant correlations in the expected direction with body size estimate, ideal size estimate, and body dissatisfaction (r values range from .38 to .57; Garner et al., 1982). At the same time, discriminant validity analyses indicate that the EAT measures constructs distinct from dieting, difficulty controlling weight, and neuroticism (Garner & Garfinkel, 1979).

**Considerations**

The EAT was explicitly designed for use as a screening instrument and is not meant as a stand-alone measure for drawing clinical conclusions. Anstine and Grinenko (2000) identified four questions—common in routine primary care practice—that significantly correlate with above-threshold EAT-26 scores: “How many diets have you been on in the past year?”; “Do you feel you should be dieting?”; “Do you feel dissatisfied with your body size?”; and “Does your weight affect the way you feel about yourself?” Thus, if patients endorse these sentiments or other high-risk behaviors (e.g., vomiting after eating large amounts of food, extreme caloric restriction), it may be appropriate to administer the EAT in order to collect additional information.

Initially, the EAT was thought to have high false positive rates and low predictive power for detecting anorexia nervosa when used in nonclinical samples—a problem further compounded by the changes in diagnostic criteria for eating disorders in the decades since its development (Mintz & O’Halloran, 2010). Changes to the DSM in the fourth and fifth editions, however, created broader classifications for eating disorders. The DSM-5 now includes diagnoses of anorexia nervosa, bulimia nervosa, binge eating disorder, other specified (i.e., subthreshold) eating disorder, and unspecified eating disorder. When implemented as a screening tool for differentiating between any DSM eating disorder diagnosis and no DSM eating disorder diagnosis (as opposed to differentiating among diagnoses), the predictive power of the EAT significantly improves (Mintz & O’Halloran, 2010).

**Guidelines for Decision-Making**

The online version of the EAT (available at http://www.eat-26.com) generates a “referral index” by interpreting EAT score information in the context of BMI and responses to six behavioral questions.
The authors suggest that a specialist referral be made if any of the following criteria are met: (1) BMI < 17.5; (2) EAT-26 score > 20 or EAT-40 score > 30; (3) endorsement of at least one high-risk disordered eating behavior in the past six months. This algorithm recognizes that one might score below the threshold on the EAT but still engage in potentially unsafe eating-related behavior. In each of these cases, a behavioral health clinician can further assess for clinically disordered eating and collaborate with the primary care physician in order to determine the most appropriate treatment plan.

The EAT is not widely used as an ongoing measure of treatment progress, but rather, is most commonly administered pre- and posttreatment in clinical settings. In a sample of women in treatment for bulimia, Davis, Olmsted, and Rockert (1990) determined that a change in score of 15 or more on the EAT-26 reflects reliable change, with a recommended cutoff score of 19 to indicate clinically significant change.

Use in Primary Care Settings
Despite its widespread use (Koslowsky et al., 1992), the authors are not aware of any published reports of EAT implementation in a primary care setting. From a provider perspective, the EAT-26 is relatively economical to administer in terms of time; however, this may depend on the other issues that need to be addressed at any given patient's appointment. Using a computerized version would reduce scoring time and is more easily integrated into electronic medical records. Cost of administration is not an issue, as the EAT is available for free for those who request permission (http://www.eat-26.com). From a patient perspective, perceived social desirability and/or denial of maladaptive eating attitudes might result in artificially deflated scores on the EAT (Garfinkel & Newman, 2001). As such, it is crucial that the provider establish rapport and explain the purpose of the measure in a nonaccusatory manner so as not to raise patient defenses.

The Primary Care Evaluation of Mental Disorders (PRIME-MD) Eating Disorder Module
The Primary Care Evaluation of Mental Disorders (PRIME-MD; Spitzer et al., 1994) is one of the most widely used measures in clinical research for diagnosing mental health disorders by primary care physicians (Bakker, Terluin, Marwijk, Mechelen, & Stalman, 2009). The PRIME-MD consists of a two-stage procedure: (1) a 26-item patient questionnaire (PQ) used for initial symptom screening, and (2) the Clinical Evaluation Guide (CEG) given for positive responses on the PQ (Loerch, Szegedi, Kohnen, & Benkert, 2000). The CEG is a structured clinical interview consisting of five diagnostic modules (mood, anxiety, alcohol abuse, somatoform, and eating disorders), with each module only being administered after it is triggered by positive responses on the PQ.

The instrument was originally created from DSM-III-R criteria and was developed over an 8-month period of administering preliminary versions to 450 patients at seven primary care sites; the measure has been updated for the DSM-IV (Hahn, Kroenke, Williams, & Spitzer, 1999). One item (#17) on the PQ screens for eating disorders, and if it triggers a “yes” response, the clinician should administer the eating disorder module of the CEG to determine if the patient meets DSM-IV criteria (Spitzer et al., 1994). The CEG eating disorder module comprises diagnostic questions regarding bulimia, purging, and nonpurging type corresponding directly to DSM-IV criteria; subthreshold categories of binge eating disorder and eating disorder not otherwise specified are also included (Loerch, Szegedi, Kohnen, & Benkert, 2000).

Multiple forms of the PRIME-MD exist. There are Dutch and German versions and two automated forms (The Patient Problem Questionnaire: A Self-Administered Version of the PRIME-MD [PPQ] and a Computer-Administered Telephone PRIME-MD; Bakker, Terluin, Marwijk, Mechelen, & Stalman, 2009; Hahn et al., 1999; Kobak et al., 1997). The PPQ is a shorter version based on patient self-report,
and the Computer-Administered Telephone PRIME-MD uses interactive voice response (IVR) technology to administer a computerized version via telephone (Hahn et al., 1999; Kobak et al., 1997).

**Normative Data and Psychometric Characteristics**

In general, the PRIME-MD has displayed good concurrent validity with DSM-IV diagnostic criteria (Loerch, Szegedi, Kohnen, & Benkert, 2000) and fair test-retest reliability (k = 0.27; Bakker, Terluin, Marwijk, Mechelen, & Stalman, 2009). The PRIME-MD is also considered diagnostically efficient across forms. For instance, a study examining the German version found good sensitivity (.73), specificity (.67), and overall accuracy (.70) for all the psychiatric disorders (Loerch et al., 2000). Further, according to the diagnostic categories of mood disorders, anxiety disorders, eating disorders, and alcohol-related disorders, the sensitivity of PRIME-MD ranged from .67 to .80 (Loerch et al., 2000). In addition, in a sample of 1,000 patients from four academic medical center primary care clinics, the specificity and positive predictive values of the PRIME-MD were excellent (.84 and .99, respectively; Johnson et al., 1995).

**Considerations**

When administering the PRIME-MD, it is important to remember that the first component (PQ) is used for screening purposes and the second component (CEG) is used for diagnostic purposes based on DSM-IV criteria. However, even if PQ screen items are negative, clinicians are instructed to follow their own clinical judgment in choosing to administer a CEG module. Noteworthy is that the overall sensitivity of the PRIME-MD is limited by the sensitivity of the PQ, and the efficiency of the PRIME-MD will be determined by the specificity of the PQ (Hahn et al., 1999).

The three most strongly substantiated characteristics within the literature that should trigger administration of the PRIME-MD for physicians are (1) the presence of multiple unexplained physical symptoms, (2) functional impairment out of proportion to the patient’s nonpsychiatric medical problems, and (3) physician-experienced difficulty in caring for the patient (Hahn et al., 1999).

**Guidelines for Decision-Making**

Of importance when administering the PRIME-MD is to conduct a patient-centered interview rather than a symptom-driven one; this will help mediate any emotional responses that may arise during the assessment. The PRIME-MD PQ is usually given to the patient prior to seeing the physician, but also may be mailed to the patient’s home before the appointment. Alternatively, clinicians may choose to administer the PQ orally to the patient during the visit and then the clinician can move directly to the corresponding CEG module. However, before moving on to the CEG modules, the clinician should assess for the patient’s presenting concerns and discuss an agenda for the visit. If triggered, the CEG modules should be administered in the order they are arranged (mood, anxiety, eating, alcohol, and somatoform, respectively), as this sequence maximizes efficiency. Of importance is ensuring that the patient understands the duration and severity criteria for each symptom. The last page of the PRIME-MD CEG is a checklist of diagnoses with ICD-9-CM codes to use as a summary sheet for the patient’s records, as well as CPT and EM codes that can be used for visits (Hahn et al., 1999).

In addition to being a diagnostic instrument, the administration of the PRIME-MD CEG is also useful as a form of feedback to the patient. This is because the physician inquires about the patient’s responses on the PQ to ask the more specific symptom questions on the CEG. Physicians can then provide feedback about each diagnosis as it is indicated during the interview or discuss diagnoses at the end. It is recommended that physicians provide information about the diagnoses made and discuss any potential questions or concerns their patients have. Assessing one’s mental status, psychosis, and drug abuse should be addressed in any patient with a PRIME-MD diagnosis, as these conditions are not included in the measure. In addition, clinical judgment should be used for assessing other
diagnoses that are not included in the PRIME-MD (e.g., using the PCL-C to assess the presence of PTSD symptoms). Lastly, the patient’s social system is an important component of this evaluation; assessing it allows for the determination of the amount of the patient’s social support (Hahn et al., 1999).

**Use as an Instrument for Treatment Monitoring and Outcome Assessment**

It is important to monitor the treatment of patients with PRIME-MD diagnoses in order to assess symptom severity and determine whether conditions persist. Administering the PQ physical symptom checklist and the symptom lists in the mood and anxiety modules can be done at sequential visits in order to monitor treatment. However, the PRIME-MD is not designed for the graded assessment of symptom severity or one’s response to treatment (Hahn et al., 1999).

**Strengths and Limitations of Use in Primary Care**

As previously mentioned, the PRIME-MD was developed for usage in a primary care setting and has been found to be valid and reliable when used by primary care providers, including psychologists. The PRIME-MD is a cost-effective instrument that may qualify for third-party reimbursement (Hahn et al., 1999). However, the CEG does require the physician to have the time to administer the specific modules in a patient-centered interview format, as well as provide appropriate feedback when diagnoses are present; this may be difficult in some primary care settings that abide by the traditional 30-minute visit model.

**Other Measures**

**Hypoglycemic Fears Scale**

The first version of the hypoglycemic fears survey (HFS) was developed in the 1980s with a goal of identifying patients with excessive fear and phobic avoidance of symptoms of hypoglycemia. Cox and colleagues (1987) developed 34 items based on interviews with diabetes healthcare providers and insulin-dependent diabetic patients. The initial items were administered to a small group of patients and care providers and redundant items were revised or removed, leaving 29 items that were divided into two subscales: affective fear and behavioral avoidance of hypoglycemia (Cox, Irvine, Gonder-Frederick, Nowacek, & Butterfield, 1987).

The HFS-II retains only 18 of the items from the HFS in their original form, with the remaining 15 reflecting significant changes based on a number of studies. The HFS-II appears to better reflect the intended constructs of fear and behavioral avoidance compared to HFS in that factor analysis indicated a two-factor solution for the HFS-II with items loading on the expected subscales and with both subscales representing distinct, but correlated constructs. Rasch PCM analyses also supported a two-factor solution (Gonder-Frederick et al., 2011).

The HFS-II has been used in studies with both type I and type II diabetics, and is administered as a self-report questionnaire with patients indicating how often during the past six months they experienced symptoms or behaviors on a Likert scale ranging from 0 (“never”) to 4 (“very often”). Alternative versions have been developed for use with pediatric patients and their parents (Gonder-Frederick et al., 2000). Also, the scale has been translated and validated in Swedish (Anderbro et al., 2008), and the worry subscale was translated and used for an Asian-Pacific region study including patients from China, Korea, Malaysia, Thailand, and Taiwan. Findings of this study support initial validity of this subscale in these languages (Sheu et al., 2012). An 11-item short form of the HFS-II has recently been developed that retains the factor structure and good internal reliability (Grabman et al., 2016). Further validity testing is needed for the short version, but this represents a promising development in efficient assessment of this construct for the primary care setting.
Normative Data and Psychometric Characteristics
The HFS-II demonstrates good internal consistency (> .85) for the total score and both subscale scores. Test-retest reliability was adequate (.63–.81; Gonder-Frederick et al., 2011). The HFS-II correlated positively with measures of other types of psychological distress (Modified State-Trait Personality Inventory Anxiety, Anger, and Depression subscales, Beck Depression Inventory, Beck Anxiety Inventory) and negatively with health-related quality of life (i.e., SF-12 Mental and Physical Component Summary scores). HFS-II scores were also higher in individuals with episodes of severe hypoglycemia in the previous year (Gonder-Frederick et al., 2011). Higher scores indicate higher fear of hypoglycemia, and scores are typically higher in type I (total scale $M = 1.34$, $SD = 0.66$) than type II (total scale $M = .80$, $SD = 0.70$) diabetics in US patient groups, and women generally score slightly higher than do men on the HFS-II (Gonder-Frederick et al., 2011).

Use as an Instrument for Treatment Monitoring and Outcome Assessment
The HFS-II demonstrates good criterion validity in that it is sensitive to change in interventions that target fear of hypoglycemia or that reduce hypoglycemic episodes in both type I and type II diabetics (Cox, Gonder-Frederick, Polonsky, Schlundt, Kovatchev, & Clarke, 2001; Johnson, Kotovych, Ryan, & Shapiro, 2004; Schachinger et al., 2005).

CASE EXAMPLE

Mrs. Anderson is a 52-year-old Caucasian married female with insulin-dependent type II diabetes, hypertension, and chronic low back and knee pain. She has been referred due to difficulty improving her diabetes management despite extensive counseling with a nutritionist (who is also a certified diabetes educator) and her primary care physician. Most recent HgA1c = 11.3, weight = 230 lbs., height = 63 inches, BMI = 40.7. (Note that names and case details have been altered to protect the identity of the patient.)

Review of medical records indicates a past history of dependence on prescription pain medication that continued for 3 years after knee surgery. She has been substance-free now for 15 years. No history of previous psychological treatment was noted in her record, but she is prescribed 75 mg of sertraline, an antidepressant.

At her initial appointment with the primary care psychologist, Mrs. Anderson reported concern about her difficulty in changing her eating habits and worry over her health. She noted that despite feeling highly motivated to improve her eating habits, she found herself repeatedly losing control and eating more than she intended on a near daily basis. On a typical day, she would wake around 10 a.m., and eat her first meal around noon. She described this meal as “something healthy” like a salad with grilled chicken and two tablespoons of ranch dressing. Her next meal would be around 6:30 p.m. when she would typically eat pork chops, seasoned rice, and canned green beans. At dinner, she typically would eat large portions and experience difficulty stopping herself from going back to get second and third portions. She reported that “on a good day” she stopped eating after dinner, but that most of the time, she finished dinner feeling as if she “blew it” with her diet and would follow dinner with a series of snacks including ice cream, chips, cookies, and salted nuts. She would go to bed around midnight with a painfully full stomach, feeling guilty, ashamed, and hopeless about ever being able to meet her goal of getting her diabetes under control.
In addition to her eating problems, she reported experiencing social anxiety and depression since her mother’s death 7 years earlier and feelings of shame over her inability to cope with this loss. She tearfully noted that “I should be over it by now.” Her eating has gotten more difficult to manage since losing her mother and she has gained 60 pounds. She reported limited social or physical activity, no longer attending her church that she once loved, and rarely leaving her home other than coming to her medical appointments. She was prescribed sertraline 3 years ago by her primary care provider and reported that initially it seemed to be helping; presently, however, she does not think it is really helping with her depression. With regard to her diabetes management, she reported checking her blood sugar four times daily and injecting her insulin once daily at dinnertime, as instructed. She denied any difficulty maintaining these self-management behaviors.

At the conclusion of her first session, Mrs. Anderson was provided with feedback that her scores on testing were consistent with her self-reported depression, anxiety, and problems with binge eating (see Table 19.4). She agreed with the psychologist’s perception that her difficulties with eating were largely related to her grief, sadness, anxiety, and loss of happiness in her life, and that binge eating had become an overlearned response to difficult emotions. She agreed it was likely that she’d be better able to change her eating if she felt happier, enjoyed life more, and felt more confident in herself. She noted that she had realized for several years now that pulling away from people and her church was something she wanted to change; however, she found it difficult to gather the energy to go back even though friends still sometimes call and ask her to go with them.

Mrs. Anderson was given a brief overview of the types of skills she would learn by attending primary care psychology sessions, and she developed a plan to attend church the following weekend. She was also instructed to record her food intake using an adapted version of the food records given to her by her nutritionist. The form included space to record her food intake, time of day, mood at time of eating, intensity of mood (1 = mild mood, 10 = extreme mood), and whether she considered the eating episode to constitute a binge, an overeating episode, or neither. She was also agreeable to discussing a possible medication adjustment with her primary care provider. Her next appointment with the psychologist was scheduled in 2 weeks. The psychologist mentioned the medication adjustment to Mrs. Anderson’s primary care provider in the clinic workroom and also copied her on her notes in the medical record. Her primary care provider increased Mrs. Anderson’s sertraline dose to 150 mg.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ-9</td>
<td>18, no suicidal ideation</td>
<td>Moderately severe depression</td>
</tr>
<tr>
<td>GAD-7</td>
<td>17</td>
<td>Severe anxiety</td>
</tr>
<tr>
<td>EDE-Q Global</td>
<td>4.2</td>
<td>Clinically elevated</td>
</tr>
<tr>
<td>EDE-Q Shape</td>
<td>3.7</td>
<td>Not clinically significant (NCS)</td>
</tr>
<tr>
<td>EDE-Q Weight</td>
<td>5.7</td>
<td>Clinically elevated</td>
</tr>
<tr>
<td>EDE-Q Eating</td>
<td>5.1</td>
<td>Clinically elevated</td>
</tr>
<tr>
<td>EDE-Q Restraint</td>
<td>2.2</td>
<td>NCS</td>
</tr>
<tr>
<td>HbA1c</td>
<td>11.3</td>
<td>High daily blood glucose levels</td>
</tr>
<tr>
<td>Weight/BMI</td>
<td>230/40.7</td>
<td>Clinically obese</td>
</tr>
<tr>
<td>Number of binge episodes during previous 28 days</td>
<td>24</td>
<td>Meets DSM-5 frequency criteria for BED</td>
</tr>
</tbody>
</table>
Mrs. Anderson attended bimonthly 30-minute sessions during the first 3 months of treatment and then attended monthly for 3 additional months. She attended a total of 9 sessions. GAD-7 and PHQ-9 were administered monthly and the EDE-Q and were readministered at 3 months and at 6 months (see Tables 19.5 and 19.6, respectively). Session content included traditional cognitive behavioral therapy as well as dialectical behavior therapy methods. Specific intervention components included:

- Relaxation skills;
- Psychoeducation about the nature and function of emotions, grief, and healthy dietary and body image attitudes;
- Behavioral activation;
- Self-monitoring, breaking the chain analysis;
- Cognitive restructuring of rigid dietary rules, distorted body image beliefs;
- Urge surfing skills (skills for coping with urges to binge eat);
- Mindfulness and acceptance-based emotion regulation skills.

At her 3-month appointment, Mrs. Anderson reported signing up for a grief support group starting that month at a church in the community. She had successfully attended church 75% of the time over the past three months and had reconnected with several friends in so doing. She

<table>
<thead>
<tr>
<th>Measure</th>
<th>Score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ-9</td>
<td>12, no suicidal ideation</td>
<td>Moderate depression</td>
</tr>
<tr>
<td>GAD-7</td>
<td>12</td>
<td>Moderate anxiety</td>
</tr>
<tr>
<td>EDE-Q Global</td>
<td>3.6</td>
<td>NCS</td>
</tr>
<tr>
<td>EDE-Q Shape</td>
<td>3.1</td>
<td>NCS</td>
</tr>
<tr>
<td>EDE-Q Weight</td>
<td>5.0</td>
<td>Clinically elevated</td>
</tr>
<tr>
<td>EDE-Q Eating</td>
<td>4.2</td>
<td>Clinically elevated</td>
</tr>
<tr>
<td>EDE-Q Restraint</td>
<td>2.2</td>
<td>NCS</td>
</tr>
<tr>
<td>HbA1c</td>
<td>9.8</td>
<td>High daily blood glucose levels</td>
</tr>
<tr>
<td>Weight/BMI</td>
<td>227/40.2</td>
<td>Clinically obese</td>
</tr>
</tbody>
</table>

**Table 19.5** Case Example: 3-Month Data

<table>
<thead>
<tr>
<th>Measure</th>
<th>Score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ-9</td>
<td>8, no suicidal ideation</td>
<td>Mild depression</td>
</tr>
<tr>
<td>GAD-7</td>
<td>6</td>
<td>Mild anxiety</td>
</tr>
<tr>
<td>EDE-Q Global</td>
<td>3.3</td>
<td>NCS</td>
</tr>
<tr>
<td>EDE-Q Shape</td>
<td>2.8</td>
<td>NCS</td>
</tr>
<tr>
<td>EDE-Q Weight</td>
<td>4.1</td>
<td>Clinically elevated</td>
</tr>
<tr>
<td>EDE-Q Eating</td>
<td>3.5</td>
<td>NCS</td>
</tr>
<tr>
<td>EDE-Q Restraint</td>
<td>2.6</td>
<td>NCS</td>
</tr>
<tr>
<td>HbA1c</td>
<td>7.6</td>
<td>High daily blood glucose levels</td>
</tr>
<tr>
<td>Weight/BMI</td>
<td>219/38.8</td>
<td>Clinically obese</td>
</tr>
<tr>
<td>Number of binge episodes during previous 28 days</td>
<td>0</td>
<td>No longer meets DSM-5 frequency criteria for BED</td>
</tr>
</tbody>
</table>

**Table 19.6** Case Example: 6-Month Data
experienced an improvement in her mood fairly early in therapy after discovering that church members did not criticize or judge her for her absence, but instead welcomed her and readily included her in activities and events. She also attributed this improvement in mood to her increased energy with her new sertraline dose.

While this initial improvement in mood was promising, Mrs. Anderson still suffered with moderate depression and anxiety (see Table 19.5). She attributed this primarily to grief, which is why she decided to join the group. She reported that therapy had helped her to develop more self-compassion about her difficulty coping with her mother’s death and had given her skills for coping, but she also felt a need to connect with others who were suffering with loss to gain further perspective. Therapy had also helped her to develop improved self-compassion with her struggle with binge eating and her weight.

Her food records indicated that the frequency of her binge episodes have dropped from approximately six per week prior to treatment to zero episodes in the past four weeks. Although she was initially disappointed to find that her weight had dropped only 3 pounds, she was encouraged by her improved HbA1c and was responsive to feedback that binge abstinence, contrary to what one might expect, does not always lead to weight loss. The psychologist pointed out that over the 3 months prior to treatment, she had gained 4 pounds, and Mrs. Anderson recognized that this meant a very meaningful change had occurred. She reported that she had gained confidence in her ability to prevent an overeating episode from turning into a binge episode due to restructuring her cognitions about overeating. Her testing results also confirmed that her eating concerns were improved. Despite these gains, Mrs. Anderson reported that she still struggled considerably with emotional eating and difficulty resisting specific cravings when she got them, which was two to three times per day.

Mrs. Anderson was overjoyed to learn that her HbA1c level had again dropped considerably and that she had lost 8 pounds over the past three months (see Table 19.6). She attributed her success to significantly improved mood, continuing improved energy levels, and self-confidence. Her food records indicated episodes of overeating and emotional eating several times per week, and only one binge episode over the previous three months. Mrs. Anderson reported that the chain analysis worksheets had been particularly helpful in recognizing her patterns and developing alternative habits and responses to emotionally difficult situations. She noted that therapy was helping her to recognize the relationship between her emotions and her eating habits. She was also gaining an ability to recognize and validate her emotional experiences without judgment, and to choose different responses than binge eating in order to manage her emotions.

**SUMMARY**

Skills for assessing and treating a diversity of problems in eating behavior are valuable for psychologists in integrated primary care settings. While most of these problems can be adequately managed in the primary care setting, this will vary based on the available on-site expertise, structure of the clinic, and nature/severity of the presenting problem. When problems present that are beyond the scope of the primary care model, psychologists can be pivotal in assessing the nature of these problems and identifying appropriate treatment resources in the community.

Assessment of eating problems in the primary care setting ideally integrates information from multiple sources: consultation with collaborating staff members, review of medical records, clinical interview, and standardized assessment measures. Many of the problems encountered in primary care are complex and comorbidity is common. Primary care psychologists are skillful at quick global
assessments that can be honed using specific assessment measures. A number of assessment measures are available that are suitable for use in primary care settings, although limited research is available on the use of these measures specifically in this setting. Given the high likelihood of eating problems presenting in primary care rather than traditional mental health settings, further research identifying which assessment tools work best in this setting should be a priority for primary care psychologists.

REFERENCES


During the past few decades, there has been an increasing interest in the measurement of health status in medical and behavioral healthcare delivery systems. Initially, this interest was shown primarily within those organizations and settings focused on the treatment of physical diseases and disorders. In recent years, behavioral healthcare providers have recognized the value of evaluating the patient’s general level of health.

It is important to recognize that the term “health” means more than just the absence of disease or debility; it also implies a state of well-being throughout the individual’s physical, psychological, and social spheres of existence (World Health Organization [WHO], 1948). Dickey and Wagenaar (1996) note that this concept of health recognizes the importance of eliciting the patient’s point of view in measuring health status. They also point to similar conclusions reached earlier by Jahoda (1958) that are specific to the area of mental health. Here, an individual’s report of his mental health status relative to how he feels it should be is an important component of “mental health.”

Health status can be defined as “‘functional capacity’ or a state of physiological or psychological functioning or well-being” (Berger et al., 2003, p. 131). Measures of health status can be classified into one of two groups: generic and disease-/condition-specific. Generic health status measures are developed to be appropriate for use with patients without regard to the condition, disease, or disorder they may be experiencing, as well as with nonpatients. Condition-specific health status and functioning measures are intended for use with patients suffering from a particular condition, disease, or disorder, with the content of these measures generally assessing symptoms, limitations in functioning, and other aspects of health commonly affected by that particular condition. Condition-specific surveys have been utilized for a number of years. Most have been developed for use with specific physical rather than mental disorders, diseases, or conditions; however, condition-specific measures of mental health status and functioning are available. Burnam (1996) has identified several strengths and weaknesses of generic and condition-specific health status measures. These are summarized in Table 20.1. Use of one type of instrument instead of the other often requires trade-offs. To deal with this issue, Burnam recommends including the administration of both types of measures, at least when assessing outcomes.

Arguably, the most widely used and respected generic health status measure is the 36-item Short Form-36v2 Health Survey (SF-36v2; see Maruish, 2011). It measures eight domains of health—four addressing physical health-related constructs and four addressing mental health-related constructs—that reflect the WHO concept of “health” and thus reflects a biopsychosocial approach to the assessment of health. Many would consider it the gold standard for generic health status measurement instruments as evidenced by its use in studies seeking to validate other health status measures (for example, see Gersh, Arnold, & Gibson, 2011; Hawthorne, Kaye, Gruen, Houseman, & Bower, 2011; Yoshida et al., 2011). In response to concerns that even this relatively brief objective measure is too lengthy for regular administration in clinical and research settings, a 12-item, abbreviated version of the measure was developed. The SF-12v2 Health Survey (SF-12v2; see Maruish, 2012c) was developed for use in large-scale, population-based research where the monitoring of health status at a broad level
is all that is required. It also has been found to be useful for clinical assessment purposes (see Maruish, 2011, 2012c).

It is important to note that some experts refer the SF health surveys as measures of health-related quality of life (HRQOL) rather than measures of health status (for example, see Bowling, 2001; Fayers & Machin, 2007; Walters, 2009). According to Berger et al. (2003, p. 129), HRQOL is “a broad theoretical construct . . . concerned with the evaluation of health status, attitudes, values, and perceived levels of satisfaction and general well-being with respect to either specific health conditions or life as a whole from the individual’s perspective.” Citing McDowell and Newell (1996), Fortin, Dubois, Hudon, Soubhi, and Almirall (2007) indicate that HRQOL “provides a multidimensional perspective that encompasses a patient’s physical, emotional, and social [i.e., biopsychosocial] functioning” (p. 2). Sometimes the two terms are used interchangeably (for example, see Walters, 2009). Even the developers of the instruments refer to the SF-36v2 and SF-12v2 as both health status and HRQOL measures.

The purpose of this chapter is to provide the reader with a brief but comprehensive overview of the development, psychometric properties, basic administration and interpretive guidelines, and potential uses of the SF-36v2 in primary care settings. Concurrently, this same type of psychometric information will be presented for the SF-12v2, as this abbreviated version of the SF-36v2 will likely provide an adequate and more useful assessment for the majority of patients being referred to psychologists in primary care settings.

**WHY MEASURE HEALTH STATUS OR HRQOL?**

Measures of health status and HRQOL are now commonly used in randomized controlled trials (RCTs) for various reasons, including to support labeling claims for pharmaceutical and medical device products (for example, see Optum, 2013). In fact, the US Food and Drug Administration (FDA) now provides guidance for the use of these types of measures, which are referred to as patient-reported outcome...
Assessment of Health Status and HRQOL

Experts have identified several benefits of administering health status/HRQOL measures in clinical settings. Cella and Stone (2015) note several reasons, including that patient self-report of physical and mental health, as well as aspects of their ability to function, is a key component of health surveillance; has been proven to be a powerful predictor of morbidity and mortality; and employs valid measures of health service needs and treatment outcomes. Both Walters (2009) and Fayers and Machin (2007) have noted that assessments employing HRQOL measures facilitate communication with patients, allowing for the identification of problems and their effects on the patient that otherwise might not be reported. HRQOL measures can also identify problems that may continue after treatment termination and require further treatment. In addition, Ware (1992) indicates that this type of information can be used for several purposes. For example, it can be helpful in clinical decision-making; detecting and understanding well-being, functioning, and other aspects of HRQOL that can be affected by disease and its treatment; tracking changes in functioning over time, thus allowing for more information to be considered in choosing among treatment options; and better predicting the course of chronic disease.

In addition, depending on the measure employed, important information about aspects of health not directly queried by the measure’s items can be derived from their empirically demonstrated relationship to the results obtained from the administration of the measure. This in fact is the case for both the SF-36v2 and the SF-12v2, as will be discussed later in the interpretation section of this chapter. It is these Short Form (SF) surveys’ ability to go beyond their health component summary measure and health domain scores to provide important related information that makes them valuable resources for psychologists practicing in primary care and other clinical settings.

DEVELOPMENT OF THE SHORT FORM HEALTH SURVEYS

Each of the original and revised versions of the SF-36 Health Survey contains 36 items that comprise a total of eight independent scales—four having to do with the patient’s physical health and its effects on functioning, and four having to do with the patient’s mental health and its effects on functioning. Two component summary scales also are derived to present an overall summary of the patient’s responses related to each of these two health dimensions.

Comprehensive yet brief overviews of the development of the SF family of instruments are presented in Ware (1999) and Wetzler, Lum, and Bush (2000). Further information is provided in Maruish (2011, 2012c). In summary of these works, development of the SF-36 can be said to have its origins in two large-scale studies. One is the Health Insurance Experiment (HIE; Brook et al., 1979; Newhouse & The Insurance Experiment Group, 1993), which investigated issues related to healthcare financing. The other is the often-referenced Medical Outcomes Study (MOS; Stewart & Ware, 1992; Tarlov et al., 1989), which investigated physician practice patterns and consequent patient outcomes. Both required the use of brief measures of functional and health status. The MOS investigators used a 149-item questionnaire that made use of items from other commonly used instruments, including that used in the HIE. It is from this questionnaire that 36 items were later selected to represent the eight health concepts or domains that were considered to be most affected by disease and treatment,
and that were also the most widely measured concepts in other instruments. The following scales represent these eight health concepts or domains: Physical Functioning (PF), Role-Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social Functioning (SF), Role-Emotional (RE), and Mental Health (MH). Each of the scales is composed of 2 to 10 of the items, and each of the 35 scale-related items scores on only one of the scales. Another item that appears on the SF-36v2—the Self-Evaluated Transition (SET) item—does not score on any of the health domain scales or component summary measures; however, it provides valuable information regarding the patient’s view of how their health has changed over a specific period of time.

Factor analytic work on SF-36 findings from several sources revealed that 80%–85% of the instrument’s reliable variance could be accounted for by physical and mental health factors. These factors led to the development of the Physical Component Summary (PCS) and Mental Component Summary (MCS) measures (Ware, Kosinski, & Keller, 1994) and have consistently been found in subsequent investigations (see Maruish, 2011; Ware, Kosinski, & Dewey, 2000).

Other large-scale studies contributed to the further development of the SF-36 (see Maruish, 2011). During the International Quality of Life Assessment (IQOLA) Project (see Gandek & Ware, 1998), a multinational effort to develop 14 validated translations of the SF-36 was undertaken. Resulting from this endeavor were improvements in the survey’s items. The second version of the SF-36, the SF-36v2, was subsequently developed in 1996. The improvements incorporated into the SF-36v2 include improved formatting and presentation of the items and instructions, wording changes to improve objectivity and understandability of the survey, and increased measurement precision through expanded item response choices. Data collected as part of the 1998 National Survey of Functional Health Status allowed for the development and incorporation of norm-based, standardized scoring (i.e., T-scores, with a mean of 50 and standard deviation of 10) as part of the SF-36v2. Higher T-scores indicate better health in the area measured by the domain scale or component summary measure.

The SF-12 (Ware, Kosinski, & Keller, 1995) was developed in response to the SF-36 being judged as being too long for some applications (e.g., large-scale surveys). Originally, it was developed as a means of predicting SF MCS and PCS scores from only 12 items—one-third of the number of items constituting the SF-36. Development of the SF-12v2 followed the introduction of the SF-36v2 and included some of the same features that were incorporated in the SF-36v2, such as item improvements and scoring of both the component summary measures and health domain scales using norm-based T-scores. As with the SF-36v2, higher SF-12v2 T-scores indicate better health.

The norms developed from the 1998 norming survey were subsequently replaced by updated norms developed from data gathered as part of the QualityMetric 2009 Norming Study, a large-scale, nationwide renorming project for the SF surveys and other health status instruments and health questionnaires developed by QualityMetric Incorporated (now Optum), the survey’s publisher. Data from these other instruments, gathered at the same time as the SF survey data, have enabled an expansion of the interpretation of the SF-36v2 and SF-12v2 component summary measure and health domain scores beyond what was previously possible. The norming study also allowed for the gathering of updated normative and interpretive data for the accompanying questionnaires that were administered.

Note that there are two forms for each of the SF-36v2 and SF-12v2 surveys. The standard form of each survey asks individuals completing the survey to respond to the questions with regard to their functioning during the past four weeks. The acute form asks respondents to consider their functioning during the past week. The acute forms allow the clinician to accurately track patient changes on a more frequent basis than is possible using the standard form 4-week recall instructions. It is beyond the scope of this chapter to address both forms of each of the surveys. Thus, the focus of this chapter will be only on the standard (4-week recall) form for each of the SF-36v2 and SF-12v2 surveys.
Background, psychometric, and interpretive information for each of the SF-36v2 and SF-12v2 acute (1-week recall) forms can be found in their respective User’s Manuals (Maruish, 2011, 2012c).

THE SF-36V2 AND SF-12V2

The SF-36v2 Health Survey is a 36-item, multiple-choice health status survey that represents a second, revised version of the original SF-36. Like the original SF-36, the SF-36v2 was developed and normed for administration to both patients and nonpatients 18 years and older. Briefly, the SF-36v2 health domain scales comprise the following:

- **Physical Functioning** (PF) consists of 10 items that asks the respondent how limited they are in performing each of a wide range of physical tasks (e.g., bathing, lifting heavy objects) due to their health.
- **Role-Physical** (RP) has four items that ask the respondent how their physical health has presented problems for them in terms of the amount of time spent, difficulty, accomplishing as much as they would like, and limitations in performing the kind of work or other activities they were involved in.
- **Bodily Pain** (BP) has two items that ask about how much bodily pain the respondent has been experiencing and the extent to which it interfered with “normal” activities.
- **General Health** (GH) includes five items: two that ask the respondent to rate their health in general, two that ask them to compare their health to that of others, and one that asks if they expect their health to get worse.
- **Vitality** (VT) comprises four items to which the respondent indicates how much of the time they have felt tired, full of life, happy, and nervous.
- **Social Functioning** (SF) has two items that inquire about how much time and to what extent physical or emotional problems interfered with the respondent’s normal social activities.
- **Role-Emotional** (RE) has three items that ask how much of the time emotional problems have caused the respondent to accomplish less than desired, be less careful, and spend less time on work or other activities.
- **Mental Health** (MH) consists of five items, two of which ask how much of the time the respondent has experienced depression-related symptoms and one item each inquiring about feeling nervous, happy, and calm and peaceful.

The **Self-Evaluated Transition** (SET) item records the respondent’s self-rating of their health in general compared to 1 year ago (on the standard form) or 1 week ago (on the acute form).

Scoring the SF-36v2 and SF-12v2 involves a multistep process of transforming health domain raw scores to T-scores. This process includes the application of proprietary algorithms to the health domain scale raw scores in order to transform them to T-scores (mean = 50, standard deviation = 10), and the application of other algorithms to calculate the two component summary measure T-scores. The z-scores (used to generate T-scores) for all eight health domains are entered into each of the two component summary measure algorithms to arrive at PCS and MCS T-scores.

As the products of factor analytic work conducted with SF-36 and SF-36v2 data, the component summary measures provide additional sources of information about the survey respondent. Each SF-36v2 health domain scale z-score is multiplied by a domain-specific factor score coefficient. The resulting eight products are summed and the aggregate of the scores is converted to a T-score. All eight health domain z-scores along with separate sets of factor score coefficients are used in the
calculation of both the PCS score and the MCS score. The PF, RP, GH, and BP z-scores are multiplied by positively weighted factor coefficients, while each of the SF, RE, and MH domain scales are multiplied by negative factor score coefficients in the computation of the PCS score. In the computation of the MCS score, the MH, RE, and SF z-scores are multiplied by positively weighted factor coefficients, while the PF, RP, BP, and GH z-scores each are multiplied by negative factor score coefficients. The VT scale score contributes to the scoring of both summary measures through positive factor weightings. The factor score coefficients used in the algorithms for computing both the standard and acute form PCS and MCS scores are exactly the same as those used to compute the SF-12v2 PCS and MCS scores. The rationale for using the same PCS and MCS scoring algorithms across surveys and forms can be found in Maruish (2011, 2012c).

All 12 items that make up the SF-12v2 are taken directly from the SF-36v2. One or two items represent each of the health domain scales and, like the SF-36v2, each item is scored for only one of the eight health domain scales. The SF-12v2 PF, RP, RE, and MH scales each contain two items while the BP, GH, VT, and SF scales are each represented by one item.

The most current norms that are used by the publisher of the SF-36v2 and SF-12v2 are those based on data from large representative samples of the US general population gathered in 2009 (standard form N = 4,036 and 6,026 respectively; Maruish, 2011, 2012c). In addition to the combined general population norms, the publisher developed age, gender, and gender-by-age supplemental norms based on the 2009 norming sample. Moreover, other data gathered as part of the 2009 norming study allowed for the development of disease-specific benchmarks for 40 chronic conditions (e.g., diabetes, cancer, hypertension, anxiety, depression, asthma).

PSYCHOMETRIC PROPERTIES

The amount of published literature on the SF family of instruments is huge. QualityMetric Incorporated, publisher of the SF family of instruments, reported the number of articles, books, book chapters, guides, manuals, and other documents pertaining to the SF surveys to be more than 19,000 as of January 2016. It is beyond the scope of this chapter to provide a review of all of the reliability and validity information that can be found in the literature to support the use of the SF-36v2 and SF-12v2 surveys. Instead, an overview of the most relevant standard form psychometric data for each of the two surveys derived from the 2009 norming study and reported in their respective User’s Manuals (Maurish, 2011, 2012c) will be presented here.

SF-36v2

Data gathered as part of the 2009 norming study allowed for an extensive investigation of the SF-36v2’s reliability and validity. Some of the most important standard form reliability and validity information that are reported in the latest edition of the SF-36v2 User’s Manual (Maurish, 2011) are discussed next.

Reliability

Reliability of the standard form SF-36v2 was investigated through the computation of the internal consistency and test-retest reliability coefficients and the standard error of estimate (SEM, an alternative means of viewing and evaluating reliability) for the PCS and MCS scores and each of the eight health domain scores. PCS and MCS scores were computed using Nunnally and Bernstein’s (1994) method that takes into account the reliabilities of the eight health domain scales and the covariances among
Assessment of Health Status and HRQOL

All health domain scale estimates are Cronbach’s alpha coefficients. This approach yielded alpha coefficients of .96 and .93 for PCS and MCS, respectively. Alpha coefficients for the eight health domain scales ranged from .82 for GH to .96 for RP. With rare exception, alpha coefficients reported in the literature for various diagnostic and national samples were found to be 70 or greater (for example, see Lam, Lam, Fong, & Huang, 2013; Signorovitch, Brainsky, & Grotzinger, 2011; Stull et al., 2014; Thumboo et al., 2013; Zhou et al., 2013).

Test-retest reliabilities for PCS and MCS were found to be .88 and .79, respectively, while the reliabilities for the health domain scales ranged from .61 for RE to .87 for GH. For this investigation, retest intervals ranged from 80 to 123 days ($M = 106.04$ days, $SD = 5.93$ days). PCS and MCS SEMs were 2.0 and 2.7, respectively, and ranged from 2.09 for RP to 4.2 for GH for the health domain scales. Among other test-retest reliabilities found in the recent literature are those reported by Zhou et al. (2013) for a group of Chinese methadone maintenance patients. Here, health domain scale values for a 1-week interval ranged from .72 (SF) to .87 (GH) while PCS and MCS reliabilities were .86 and .85, respectively. Two-week test-retest health domain reliabilities ranged from .70 (RP) to .88 (PF) in Lam et al.’s (2013) sample of 240 Chinese adults. Results from this same sample yielded retest reliabilities of .87 for PCS and .85 for MCS.

Validity

Several approaches were taken in the investigation of the validity of the standard form SF-36v2. Construct validity was investigated using various methods. First, just as with the original SF-36, factor analytic studies largely confirmed the two major components of health—physical and mental—and the health domain scale content of each component: PF, RP and BP with the physical component; SF, RE, and MH with the mental component; and GH and VT with both components. Second, convergent and discriminant validity were demonstrated in item-scale correlations (e.g., all but one item correlating highest with its parent scale) and interscale correlations where, with some exceptions, scales most closely associated with the mental component (VT, SF, RE, and MH) correlated more strongly with each other than with those most closely associated with the physical component (PF, RP, BP, and GH), and vice versa. Third, in an investigation of know-groups validity, PCS and the domain scales measuring primarily physical health (PF, RP, BP, and GH) were better than MCS and the domain scales assessing primarily mental health (VT, SF, RE, and MH) in discriminating groups of individuals who differed in the presence of a physical condition. The opposite was found to be true with groups differing in the presence of a mental condition.

There are numerous published studies confirming the SF-36v2’s ability to discriminate between groups of individuals differing on known health variables. For example, Gao, Gao, Li, and Hao (2012) compared the SF-36v2 results from a group of patients with chronic liver disease (CLD) to those from a group of healthy controls and found the control group’s mean PCS T-score to be significantly higher than that for the CLD group (54.6 vs. 47.8, $p < .001$). Similar differences were found when comparing the two groups’ mean MCS T-scores (56.4 vs. 51.7, $p < .001$). In investigating the effects of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), Nacul et al. (2011) found that component summary measure and health domain mean T-scores for a group of 170 ME/CFS patients to be less than 37, with the mean PCS score ($T = 26.8$) and mean physical health domain scores tending to be lower than those for MCS ($T = 34.1$) and the mental health domain scores. Moreover, comparison of these ME/CFS patients’ mean PCS and MCS scores to those of 10 other diagnostic groups (e.g., diabetes, heart disease, back pain/sciatica, osteoarthritis; published in Ware et al. [2007]) found the ME/CFS group’s mean PCS score to be between 11.5 and 18.9 T-score points below those of any of the other diagnostic groups, with their mean MCS score ranging from 2.2 to 14.4 points lower than these comparison groups. In addition, in a study of 460 Chinese patients with chronic hepatitis B (CHB), Zhuang et al.
(2014) found that this group had significantly lower mean T-scores \((p < .001)\) on all 10 component summary measures and health domain scales than a group of 460 matched healthy control group subjects. The greatest difference (10.1 points) was observed on the mean GH scale T-score (39.5 vs. 49.6).

**Criterion validity** was investigated through indicators of concurrent validity and predictive validity. Correlations of each of the component summary measures and health domain scales with data on health care, validation, and background criterion variables and data from other measures gathered as part of the 2009 norming study were computed. Several relationships between SF-36v2 variables and the criterion variables were hypothesized \(a\ priori\) to exist. Evidence for the existence of each relationship was dependent on whether the obtained correlation between the two variables was .50 or greater, indicating a large effect size for product moment correlations (Cohen, 1998). Moreover, data collected from patients resurveyed three to four months after their initial survey provided data supporting the **predictive validity** of the PCS and MCS measures. In general, increasingly lower scores on PCS at baseline (i.e., poorer health) were associated with an increasing probability of one or more outpatient visits and one or more bed days due to injury or illness during the 4 weeks preceding the survey’s readministration. They were also associated with not working because of health at the time of readministration of the SF-36v2. Increasingly lower MCS scores were associated with an increasing probability of reporting both feeling down and depressed and having little or no interest in doing things \(several,\ half,\ or\ nearly\ every\ day\) during the 2 weeks prior to readministration.

**Content validity** for an instrument like the SF-36v2 could be examined from any of several perspectives. For the purpose of this chapter, it is best to consider each of the health domains that are surveyed and the detail of their coverage. As noted in the *User’s Manual* (Maruish, 2011), the survey’s eight health domains were selected from among the 40 domains that were included as part of the Medical Outcomes Study (Stewart & Ware, 1992). According to Ware and Ware, Snow, Kosinski, and Gandek (as cited in Maruish, 2011) and as reported earlier, the selected domains included those most frequently assessed by other health surveys that were widely used at the time as well as domains thought to be most affected by disease and other health conditions. As for the extent of measurement for each of the eight health domains, the *User’s Manual* noted that the SF-36v2 was developed with a recognition of the trade-offs that exist between the breadth of the domains represented in the survey and the depth of the measurement these domains required. Such tradeoffs are necessary to arrive at a useful, psychometrically sound measure that is accepted by both patients and their health care providers. Despite its brevity and limited content coverage, research and feedback available to date indicated that the SF-36v2 provides a comprehensive, valid, and reliable assessment of the most important aspects of health status.

(Maruish, 2011, p. 279)

**SF-12v2**

The developers of the SF-12v2 took the same approach to investigating the reliability and validity as was taken with the SF-36v2. And as with its parent instrument, the most important SF-12v2 reliability and validity information derived from the 2009 norming study is reported in the latest edition of the SF-12v2 *User’s Manual* (Maruish, 2012c).

**Reliability**

Internal consistency coefficients for the SF-12v2 PCS and MCS measures were found to be .92 and .88, respectively. Because each health domain scale consisted of only one or two items, alpha coefficients were not computed for these eight scales. Instead, alternate form reliabilities (i.e., correlations of each
health domain scale with its SF-36v2 counterpart with a correction for item overlap) were reported. These ranged from .67 for VT to .93 for RP. Khanna, Jariwala, and West-Strum (2015) reported PCS and MCS alpha coefficients of .84 and .73, respectively, for a sample of autistic adults, while Lam et al.’s (2013) Chinese sample yielded coefficients of .67 and .60 for the PCS and MCS measures, respectively.

Test-retest reliabilities for PCS and MCS measures were .85 and .67, respectively; among the health domain scales, they ranged from .47 for RE to .77 for both PF and GH. Retest intervals were the same as those for the SF-36v2. SEMs were 2.9 and 3.5 for PCS and MCS, respectively, and ranged from 2.7 for the RP scale to 5.8 for the VT scale. For Lam et al.’s (2013) sample, the health domain scale 2-week retest coefficients ranged from .60 (SF) to .80 (PF); PCS and MCS coefficients were .82 and .81, respectively.

Validity
The developers’ approach to investigating the validity of the SF-12v2 was generally the same as that used in their investigation of the SF-36v2’s validity. Factor analytic studies yielded essentially the same type of results as those reported for the SF-36v2. Also, the pattern of interscale correlations was similar to that found with the SF-36v2. Correlations of PCS and MCS with the health domain scales showed that the four mental health domain scales (MH, RE, SF, and VT) correlated highest with the MCS measure and lowest with the PCS measure while the opposite was true for the physical health domain scales (PF, RP, BP, and GH). Similarly, the individual items from the four mental health domain scales correlated highest with the MCS measure and lowest with the PCS measure. Again, the opposite was true for the items for the four physical health domain scales.

As with the SF-36v2, in known-groups validity studies the SF-12v2 PCS measure and the physical health domain scales were better than the MCS measure and the mental health domain scales in discriminating groups of individuals who differed in the presence of a physical condition, with the opposite being true with groups differing in the presence of a mental condition. Among the investigations of SF-12v2 known-groups validity reported in the literature is one by Carris, Ghushchyan, Libby, and Smith (2015), who investigated the impact of treatment-resistant hypertension (TRH) on HRQOL. Compared to nonresistant hypertensives, the TRH group had a significantly lower PCS mean T-score (35.8 vs. 43.2, \( p < .0001 \)) while the mean MCS T-scores were similar (49.1 vs. 50.4). Vietri, Otsubo, Montgomery, Tsuji, and Harada (2015) found significant differences \( (p < .005) \) in both PCS and MCS mean T-scores for a group of 333 depressed patients reporting pain not attributed to obvious physical causes, compared to a matched group of depressed patients not reporting pain, with the pain group reporting lower mean scores on both the PCS (43.0 vs. 47.2) and MCS (29.1 vs. 32.0) measures. In another study, Bhattarai et al. (2011) investigated the quality of life in patients with neurocysticercosis (NCC) compared to that of a matched control group. They found the NCC group’s PCS and MCS mean T-scores both to be significantly lower than those for the control group.

Also, most of the correlations of SF-12v2 variables with validation, healthcare, and background variables hypothesized \textit{a priori} to yield large effect sizes were indeed found to exist. Moreover, increasingly lower scores on PCS at baseline were associated with an increasing probability of one or more outpatient visits during the 4 weeks prior to the survey readministration three to four months later. Increasingly lower MCS scores were associated with an increasing probability of reporting both feeling down and depressed and having little or no interest in doing things several, half, or nearly every day during the two weeks prior to survey readministration. These MCS scores were also associated with the individual not working because of health at the time of the survey readministration. Finally, as a brief derivative of its parent survey, the previous comments regarding the content validity of the SF-36v2 also apply to the SF-12v2.
WHICH SF SURVEY TO USE

After a decision is made to administer an SF survey, the question becomes: Which one? The fact that the SF-12v2 comprises only one-third the number of items of the SF-36v2 immediately makes the SF-12v2 an attractive choice, particularly for use in primary care and other settings where patient time is limited. Completion time is approximately 5–10 minutes for the SF-36v2 and 2–3 minutes for the SF-12v2. In addition to this is the strong, empirically demonstrated relationships between results obtained from the administration of the SF-36v2 and those obtained from the administration of the SF-12v2. The User’s Manual (Maruish, 2012c) reports that the correlations between the SF-12v2 PCS and MCS measures with their SF-36v2 counterparts are both .96. Moreover, the correlations between the two surveys’ counterpart health domain scales ranges from .80 for GH to .99 for RE, with six of the eight health domain scale correlations exceeding .90. Thus, the SF-12v2 is a good, quick survey of overall health status or HRQOL that could be used on all patients seen in primary care practice or selected patients that are referred to psychologists for evaluation. Note that it is generally recommended that interpretation of SF-12v2 results be limited to findings related to the PCS and MCS measures.

For those wanting to take full advantage of all of the interpretive information that can be derived from the administration of an SF survey, the SF-36v2 is the better choice. The SF-36v2 User’s Manual (Maruish, 2011) provides a great amount of content-based and criterion-based interpretive information that goes well beyond the norm-based interpretation available from the component summary measure and health domain scale T-scores.

MODES OF ADMINISTRATION AND SCORING

There are several modes of administration available for both the SF-36v2 and the SF-12v2. These include pencil-and-paper, interview, online, fax, interactive voice response (IVR) technology, smartphone, and tablet/kiosk. A general discussion of the use of these modalities, including advantages and disadvantages, is presented in Chapter 5 of this handbook. Also, numerous studies have investigated the effect of mode of administration on the obtained results of the SF family of survey instruments. Differences in findings due to mode of administrations have been noted in some of these investigations (see Maruish, 2011, 2012c), and Maruish and Turner-Bowker (2009) provide summaries of findings of published studies comparing different modes of administration. The reader is also referred to studies by Whitehead (2011), Turner-Bowker, Saris-Baglama, and DeRosa (2013), and Marsh, Bryant, MacDonald, and Naudie (2014). Perhaps the best recommendation for choosing the mode of administration of the SF-36v2 or SF-12v2 is summarized in Maruish (2011):

In light of the findings of these studies . . . one should be aware that the method by which the SF-36v2 [and SF-12v2] data are collected may impact the obtained results. Consequently, the means of data collection should be considered in all studies involving the SF-36v2 [and the SF-12v2 or any of the Short Form surveys]. . . Ideally, data collection should always be limited to one method if the data are to be aggregated or when an individual respondent’s results are to be compared to [his or her results from an earlier administration of the survey] . . ., from another respondent, or from a group of respondents.

(p. 50)

Also available are numerous translated or adapted versions of each of the two instruments that were developed using a process that is rooted in the one used in the IQOLA project (see Gandek &
Ware, 1998; Optum, 2015). As of December, 2015 there were more than 170 translations/adaptations of the SF-36v2 and over 150 translations/adaptations of the SF-12v2 listed on the publisher’s website (see Optum, Inc., 2016). Note that all modes of administration and scoring of each of the two SF surveys are available only via a license from the surveys’ publisher or one of its approved resellers.

INTERPRETIVE STRATEGIES

The QualityMetric 2009 Norming Study gathered SF-36v2 and SF-12v2 data as well as background, validation, and other survey data from large, representative national samples of US adults. Consequently, each of the two surveys has a significant amount of information to draw upon to make norm-, content-, and criterion-based interpretations of the survey results obtained from individual patient administrations. This section will draw upon the information presented in the SF-36v2 and SF-12v2 User’s Manuals (Maruish, 2011, 2012c) to present an overview of the various interpretive approaches that are available to users of each of the two surveys. Psychologists employing either of the instruments with primary care patients are encouraged to refer to each instrument’s respective User’s Manual for further interpretive information.

SF-36v2

Three approaches to interpreting SF-36v2 results are presented in the survey’s User’s Manual: Norm-based, content-based, and criterion-based (Maruish, 2011). Use of all three approaches enables the user to maximize the extensive interpretive data that was gathered as part of the 2009 norming study.

Norm-Based Interpretation

General population T-score normative data derived from survey data obtained from 4,036 adults who participated in the QualityMetric 2009 Norming Study serve as the basis for the norm-based interpretation of standard form (4-week recall) SF-36v2 results. In general, when considering individual patient data, the survey developers recommend that

- T-scores (mean = 50, SD = 10) that are within five points (i.e., 0.5 SD) above or below the mean should be considered as falling within the average range (i.e., average T-score range = 45–55).
- T-scores greater than 55 should be considered indicative of better than average health while T-scores below 45 should be considered indicative of worse than average health on the component summary measure or health domain scale being considered.
- T-scores below 40 (i.e., more than 1 SD below the mean) should be considered as falling in the significantly impaired range in the indicated general health (physical or mental) dimension or specific health domain.
- T-scores in the 40–44 range should be considered outside of the average range and should be investigated further to determine what limitations may exist in the indicated health dimension or domain.

When considering group-level aggregated data, the developers (Maruish, 2011) recommend that scores falling within three T-score points above or below the mean should be considered as falling within the average range (i.e., average T-score range = 47–53).

It is recommended that interpretation of the SF-36v2 profile of T-scores proceed first by examining the patient’s health status from a broad view, as revealed by their PCS and MCS scores. This can
then be supplemented by examining the health domain scores. In addition to the benefits of knowing the patient’s status on an individual health domain basis, it allows for a determination of which physical health domain (PF, RP, BP, GH) scores may be contributing to a below-average PCS score, or which mental health domain (VT, SF, RE, MH) scores may be contributing to a below-average MCS score. Implications of both high and low scores on the component summary measures and health domain scales are presented in the User’s Manual.

As with other psychometric instruments, it is always helpful to construct a confidence interval (CI) for each of the component summary measure and health domain scale T-scores. The SF-36v2 User’s Manual provides the T-score SEM values necessary to arrive at the 68%, 80%, 90%, and 95% CIs for each of the 10 measures and scales.

In addition to general population norms, the patient’s component summary measure and health domain T-scores can be compared to data from two other sets of norms and benchmarks. The first set consists of norms developed by age, sex, and age-by-sex. Data collected during the 2009 norming study also allowed for the development of the second set of comparison data consisting of benchmark data for 40 physical and mental disease/condition-specific populations. Either set of comparison data is available as a separate product from the survey’s publisher.

Content-Based Interpretation

Content-based interpretation provides another means of using PCS, MCS, and health domain scale T-scores to get the most out of SF-36v2 data. The User’s Manual (Maruish, 2011) reports the percentage of the 2009 norming sample that responded to each of the items from each health domain scale in a manner indicative of problems in functioning in that domain, within each of nine T-score ranges for that domain. Similar types of data are reported for each of nine PCS and MCS score ranges. For example, 58.4% of the norm sample who obtained a VT T-score in the 35.0–39.9 T-score range reported feeling worn out most or all of the time, and the response endorsement rate increases through the lower VT T-score ranges. To facilitate score interpretation, the SF-36v2 standard form item response data presented in Table 20.2 were reported by at least 40% of the 2009 general population sample who obtained a T-score of less than 40 (i.e., more than 1 SD below the mean) on the component summary measures and health domain scales. The cutoffs (T < 40 and response endorsement rate ≥ 40%) used to create Table 20.2 are admittedly arbitrary, but they represent points at which this author considers the responses to be worthy of further inquiry by the PCP or primary care psychologist. Those utilizing the SF-36v2 in their daily practice are encouraged to evaluate the usefulness of these cutoffs for their purposes and, if necessary, apply different score and percentage cutoffs to the content-based interpretation table data found in the User’s Manual.

Criterion-Based Interpretation

Similar to the approach to content-based interpretation, in criterion-based interpretation the obtained component summary measure and health domain T-scores can provide further information about the patient via data collected from non-SF surveys and questionnaires that were administered concurrently with and/or three to four months after the initial administration of the SF-36v2 during the 2009 norming study. Thus, in addition to content-based interpretation data, Table 20.2 presents noteworthy behavioral, emotional, and other variables that were reported by at least 40% of the 2009 general population sample who obtained a T-score of less than 40 on the SF-36v2 standard form component summary measures and health domain scales. For example, at least 75% of the norm sample who obtained a PCS T-score in the 35.0–39.9 T-score range reported that they were not working at a paying job because of their health at the time of the SF-36v2 administration. As with the content-based data,
Table 20.2 Potential Problematic Areas of Functioning Based on SF-36v2 Standard (4-Week Recall) Form Content- and Criterion-Based QualityMetric 2009 Norming Study Data for Component Summary Measures and Health Domain Scales With T-Scores Less Than 40

<table>
<thead>
<tr>
<th>Measure/Scale</th>
<th>Content-Based</th>
<th>Criterion-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical Component Summary</strong></td>
<td>• Limited in vigorous and moderate activities&lt;br&gt;</td>
<td>• Chronic conditions limit usual activities&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Limited in lifting/carrying groceries&lt;br&gt;</td>
<td>• Not working because of health</td>
</tr>
<tr>
<td></td>
<td>• Limited in climbing one or more flights of stairs&lt;br&gt;</td>
<td></td>
</tr>
<tr>
<td>Physical Functioning</td>
<td>• Limited in bending/stooping/kneeling&lt;br&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Limited in walking 100 yards&lt;br&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Health not excellent&lt;br&gt;</td>
<td></td>
</tr>
<tr>
<td>Role-Physical</td>
<td>N/A</td>
<td>• Depression is a chronic condition&lt;br&gt;</td>
</tr>
<tr>
<td>Bodily Pain</td>
<td>N/A</td>
<td>• Anxiety is a chronic condition&lt;br&gt;</td>
</tr>
<tr>
<td>General Health</td>
<td>• Health fair/poor&lt;br&gt;</td>
<td>• Not always feeling happy/satisfied with life&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Not as healthy as others&lt;br&gt;</td>
<td>• Experienced stress/pressure in daily living&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Health expected to get worse&lt;br&gt;</td>
<td>• Stress/pressure has affected health&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Health not excellent&lt;br&gt;</td>
<td>• Trouble falling asleep&lt;br&gt;</td>
</tr>
<tr>
<td>Vitality</td>
<td>• Not feeling full of life&lt;br&gt;</td>
<td>• Feeling drowsy/sleepy during the day&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Not having a lot of energy&lt;br&gt;</td>
<td>• More/less than average number of hours of sleep each night&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Feeling worn out&lt;br&gt;</td>
<td>• Got enough sleep to feel rested&lt;br&gt;</td>
</tr>
<tr>
<td>Social Functioning</td>
<td>• Feeling tired&lt;br&gt;</td>
<td>• Not getting needed amount of sleep&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Physical/emotional health interferes with social activities&lt;br&gt;</td>
<td>• Feeling down/depressed/hopeless&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>• Not feeling calm&lt;br&gt;</td>
<td>• Having little interest/pleasure in doing things&lt;br&gt;</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 20.2 (Continued)

<table>
<thead>
<tr>
<th>Measure/Scale</th>
<th>Content-Based</th>
<th>Criterion-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role-Emotional</td>
<td>N/A</td>
<td>• Not always feeling happy/satisfied with life</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Experienced good deal of stress/pressure in daily living</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Chronic conditions limit usual activities/enjoyment</td>
</tr>
<tr>
<td>Mental Health</td>
<td>• Not feeling calm</td>
<td>• Depression a chronic condition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Feeling down/depressed/hopeless</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Anxiety a chronic condition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Overall quality of life rated fair/poor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not always feeling happy/satisfied with life</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Stress/pressure affected health</td>
</tr>
</tbody>
</table>

Note. Identified potential problem areas are based on a summary measure or health domain scale T-score that is less than 40 and the problem was indicated to be present for at least 40% of those with the patient’s T-score or a lower T-score (see Maruish, 2011).

the response endorsement rate increases through the lower PCS T-score ranges. Again, the T-score and percentage cutoffs employed for criterion-based interpretation information presented in Table 20.2 are based on what this author feels are clinically noteworthy. They can be adjusted and then applied to the data in the criterion-based interpretation tables found in the SF-36v2 User’s Manual.

SF-12v2

The available approaches to interpreting SF-12v2 results are the same as those noted for interpreting SF-36v2 results. The following norm-based, content-based, and criterion-based interpretive information is taken from the SF-12v2 User’s Manual (Maruish, 2012c). Note that the SF-12v2 developers did not develop content- and criterion-based interpretive data for the health domain scales owing to the fact that each of these eight scales contains only one or two items.

Norm-Based Interpretation

General population T-score normative data derived from SF-12v2 survey data obtained from 6,028 adults who participated in the 2009 norming study serve as the basis for the norm-based interpretation of standard form SF-12v2 results. In general, the same norm-based interpretive approach recommended for SF-36v2 results, including the same T-score cutoffs and the application of SEMs to construct CIs around scores, should be used in the interpretation of SF-12v2 results. Separate age, gender, and disease/condition normative data are also available for the SF-12v2 from the survey’s publisher.

Content-Based Interpretation

As with the SF-36v2, content-based interpretation data are presented in the SF-12v2 User’s Manual (Maruish, 2012c). The SF-12v2 PCS and MCS standard form item response data presented in Table 20.3 were reported by at least 40% of the 2009 general population sample who obtained a T-score of less than 40 on the component summary measures.
Criterion-based Interpretation data are also presented in the SF-12v2 User’s Manual (Maruish, 2012c). Table 20.3 presents noteworthy behavioral, emotional, and other variables that were reported by at least 40% of the 2009 general population sample who obtained a T-score of less than 40 on the SF-12v2 standard form PCS and MCS measures and scales.

### Table 20.3 Potential Problematic Areas of Functioning Based on SF-12v2 Standard (4-Week Recall) Form Content- and Criterion-Based QualityMetric 2009 Norming Study Data for Component Summary Measures With T-Scores Less Than 40

<table>
<thead>
<tr>
<th>Measure</th>
<th>Content-Based</th>
<th>Criterion-Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Component Summary</td>
<td>• Limited in moderate activities&lt;br&gt;• Limited in climbing several flights of stairs&lt;br&gt;• Health rated fair/poor</td>
<td>• Number of chronic conditions 1 SD above the general population mean&lt;br&gt;• Chronic conditions limit usual activities&lt;br&gt;• Limited in vigorous activities&lt;br&gt;• Limited in lifting/carrying groceries&lt;br&gt;• Limited in bending/stooping/kneeling&lt;br&gt;• Limited in climbing one flight of stairs&lt;br&gt;• Limited in walking 100 yards&lt;br&gt;• Not as healthy as others&lt;br&gt;• Health rated fair/poor&lt;br&gt;• One or more outpatient visits 3–4 months later&lt;br&gt;• Bothered by emotional problems&lt;br&gt;• Depression is a chronic condition&lt;br&gt;• Anxiety is a chronic condition&lt;br&gt;• Not always feeling happy/satisfied with life&lt;br&gt;• Experienced stress/pressure in daily living&lt;br&gt;• Stress/pressure has affected health&lt;br&gt;• Not feeling full of life&lt;br&gt;• Not having a lot of energy&lt;br&gt;• Feeling worn out&lt;br&gt;• Feeling tired&lt;br&gt;• Health interfered with social activities with family&lt;br&gt;• Trouble falling asleep&lt;br&gt;• Not getting enough sleep to feel rested little or none of the time&lt;br&gt;• Getting the needed amount of sleep&lt;br&gt;• Felt drowsy/sleepy during the day</td>
</tr>
<tr>
<td>Mental Component Summary</td>
<td>• Not having a lot of energy&lt;br&gt;• Not feeling peaceful and calm</td>
<td></td>
</tr>
</tbody>
</table>

Note. Identified potential problem areas are based on a component summary measure T-score that is less than 40 and the problem was indicated to be present for at least 40% of those with the patient’s T-score or a lower T-score (see Maruish, 2012c).

### Criterion-Based Interpretation

Criterion-based interpretation data are also presented in the SF-12v2 User’s Manual (Maruish, 2012c). Table 20.3 presents noteworthy behavioral, emotional, and other variables that were reported by at least 40% of the 2009 general population sample who obtained a T-score of less than 40 on the SF-12v2 standard form PCS and MCS measures and scales.

### A General Approach to Short Form Survey Interpretation

An interpretive approach involving the systematic application of norm-, content-, and criterion-based approaches to analyzing SF-36v2 or SF-12v2 results obtained from primary care patients is recommended. The first step involves determining if either of the component summary measure scores fall below the previously discussed “normal” or average range for the 2009 US general population sample. PCS scores below this range are indicative of problems in physical health status. Examination
of scores for the four physical health-related domain scales—PF, RP, BP, and GH—will help determine which of these scales and their associated problems are likely contributing to the lowered PCS score. Similarly, if the MCS score is below this range, examination of scores for the four mental health-related domain scales—VT, SF, RE, and MH—will help to determine which of these scales are contributing to the lowered MCS score. This will help identify areas of functioning in which the patient is currently experiencing problems and the general level at which they are being affected. Keep in mind that the PCS score and the MCS score are each calculated using regression equations that include scores from all eight health domain scales, not just those associated with physical health or mental health, respectively. It is important to be mindful of this because in some cases, this can result in SF-36v2 and SF-12v2 score profiles that appear inconsistent or otherwise unusual to the survey user.

One can extend the norm-based interpretive approach by comparing the patient’s component summary measure and/or health domain scale scores to relevant age- and/or gender-based norms and, as appropriate, chronic condition benchmark data. Thus, in addition to knowing where the patient stands in relation to the general population, this comparison may permit the clinician to temper the interpretation of scores—particularly those falling outside of the average range—based on demographic and/or health-related variables that may affect Short Form survey results.

With the identification of the health domains in which the patient’s health status is impaired, one should “drill down” further to determine specific aspects within those domains that are particularly problematic. Either or both of two different approaches to a content-based interpretation can be taken. The first approach is the one that was previously described. The other approach involves directly examining each response to all of the items that make up the health domain scale(s) in question. (Each survey’s respective User’s Manual provides the means for identifying the items included in each scale.) This latter approach will require the survey user to draw upon their clinical judgment to determine the meaningfulness/importance of the given responses in order to understand the patient at that particular point in time.

Applying a criterion-based interpretive approach to the SF-36v2 or SF-12v2 data (as previously described) after the content-based approach can provide information about the patient that goes beyond what can be determined by just examining survey item responses. For example, one may be alerted to the potential presence of a mental health disorder (e.g., depression, anxiety), various symptoms (e.g., sleep disturbance, low energy) or dissatisfaction with life, or to the likelihood of future outpatient healthcare visits.

Users of the SF-36v2 or SF-12v2 should refer to the survey’s respective User’s Manual for much more detailed recommendations for interpreting survey results taking the norm-based, content-based, and criterion-based approaches. Moreover, as with any other psychometric instrument, SF-36v2 and SF-12v2 findings should always be viewed within the context of other assessment information obtained from other sources. In this regard, the Short Form survey findings can serve as a source of hypotheses about the patient to be verified by other assessment information. Conversely, they also can serve as a means of verifying hypotheses generated from other sources.

**USE AS A SCREENER**

The original SF-36 was developed to be “a brief, broad, generic measure of eight domains, or aspects, of health status that are considered important in describing and monitoring individuals suffering from a disease or illness” (Maruish, 2011, p. 15). Subsequently, the original SF-12 was developed to meet the demand for an even briefer version of the parent instrument. SF-36v2 and SF-12v2 were developed later to provide updated and improved versions of the two original instruments.
Although not generally viewed as such, from one perspective the SF-36v2 and SF-12v2 also might be considered screeners, or at least capable of serving the purpose of screeners. Both instruments are brief and provide a broad picture of the patient’s physical and mental health status. The availability of sets of normative data based on large, recent samples of the adult US general population enables the clinician to quickly obtain a good picture of the patient’s overall functioning in the eight Short Form health domains. In turn, this allows the clinician to quickly identify potential health-related problem areas that may require further investigation. Added to this is the availability of content- and criterion-based data tables (see Maruish, 2011, 2012c) that provide the percentage of large, national samples of US adults who reported problematic behaviors, feelings, functional limitations, and behavioral health symptoms at each of multiple SF-36v2/SF-12v2 component summary measure and/or health domain scale T-score ranges. This allows clinicians to set T-score cutoff points to meet the screening needs in their particular practice and, if necessary, adjust them later.

**Mental Health Screening Measures Derived From the Short Form Instruments**

There are three screening instruments that have been derived from the SF-36v2. The first is the stand-alone, five-item MH domain scale, often referred to as the Mental Health Inventory-5 (MHI-5). An enhanced version of the SF-12v2 has been developed to allow SF-12v2 users to administer the full five-item version of the SF-36v2 MH scale without having to administer the full SF-36v2. Along with the MHI-5, the SF-12v2-Enhanced MH version also allows for the scoring of the Mental Health Severity Index (Bjorner, Kosinski, & Raju, 2009) that provides an estimated Beck Depression Inventory (BDI)-based depression severity rating. The third screening instrument is the MCS measure.

**Mental Health Inventory-5 (MHI-5)**

The MHI-5 (see Berwick et al., 1991; Veit & Ware, 1983) comprises the five items that make up the SF-36v2 MH scale. These items were taken from the SF-20 (Ware, Sherbourne, & Davies, 1992), which preceded the development of the original SF-36. Together, they were found to be the best predictor of the summary score for the 38-item Mental Health Inventory (MHI-38; see Veit & Ware, 1983; Ware, Johnston, Davies-Avery, & Brook, 1979). The items ask about the respondent’s happiness and nervousness, and about feeling calm and peaceful, “down in the dumps,” and “downhearted and depressed.”

The MHI-5 has been involved in numerous investigations into both physical and mental health issues that were conducted by various research teams over the past several years. For example, it has been used to measure mental health (Fone et al., 2014; Fone et al., 2013; Hewitt, Turrell, & Giskes, 2012; Olesen, Butterworth, & Rodgers, 2012; Rowlands & Lee, 2010; Surtees et al., 2008; ten Have, van Dorsselaer, & de Graff, 2015), identify depressive symptoms (Ramirez-Avila et al., 2012; Trainor, Mallett, & Rushe, 2013), screen for anxiety disorders (Cuijpers, Smits, Donker, ten Have, & de Graf, 2009; Hallgren & Morton, 2007; Means-Christensen, Arnau, Tonidandel, Bramson, & Meagher, 2005), measure psychological distress (Blinderman, Homel, Billings, Tennstedt, & Portenoy, 2009; Khlat, Legleye, & Sermet, 2014; Muhsen, Garty-Sandalon, Gross, & Green, 2010; Ziegler et al., 2011), track changes in mental health over time (Dodwell, Thomas, & Iqbal, 2012; Fone et al., 2014; Rowlands & Lee, 2010), and validate other measures (Barbour, Saadeh, & Salameh, 2012).

Investigated as the SF-36v2 MH health domain scale, the *User’s Manual* (Maruish, 2011) reports the MHI-5 as having a Cronbach’s alpha coefficient of .87, a test-retest coefficient of .76 for an 80- to 123-day retest interval, and an SEM of 3.6. Similar alpha coefficients were found in other studies.
In known-group comparison studies with the 2009 normative sample, MH/MHI-5 was found to be second only to the MCS measure as the best discriminating health domain scale or component summary measure in tests of discrimination of groups of individuals differing in whether they reported having a mental condition or not (Maruish, 2011).

The MH/MHI-5 scale has been shown to correlate highly ($r \geq .50$) with several relevant mental health variables, such as quality of life rating, happiness or satisfaction with life, feeling interest or pleasure in doing things, not feeling down/depressed/hopeless, not experiencing stress or pressure in daily living, and stress or pressure not having an effect on the patient’s life (Maruish, 2011). A strong correlation with the SF-36v2 MCS measure ($r = .92$) has also been demonstrated. Similarly, Kelly, Dunstan, Lloyd, and Fone (2008) obtained a correlation of .88 between MCI-5 and MCS measures. In addition, correlations with other mental health screening instruments have been reported in the published literature. For example, the MHI-5 was found to correlate $-.63$ with the Zung Self-Rating Depression Scale (Yamazaki, Fukuhara, & Green, 2005), $-.61$ with the General Health Questionnaire-12 (GHQ-12) Depression subscale (Damasio et al., 2013), and $-.43$ with the Mini-International Neuropsychiatric Interview Major Depressive Episode module (Friedman, Heisel, & Delavan, 2005). Moreover, the MHI-5 derived from the original SF-36 (SF-36v1) was found to correlate $-.64$ with the Beck Depression Inventory (van den Beukel et al., 2012) and $-.64$ with the number of positive responses to the 12-item GHQ-12 (Kelly et al., 2008).

The usefulness of the MHI-5 as a depression screener has often been investigated through the generation of receiver operating characteristic (ROC) curves and the determination of the resulting area under the curve (AUC; see Chapter 6 of this book for a discussion of ROC curves and AUCs). AUCs have been identified as measures of criterion validity (see Friedman et al., 2005). An AUC of .80 or greater has frequently been cited as the minimum criterion for determining the usefulness of an instrument for screening for depression (see Holmes, 1998; Means-Christensen et al., 2005; van den Beukel et al., 2012). Using criteria such as DSM diagnoses or scores from the BDI or PHQ-9 to identify depressed individuals, the investigations reviewed have revealed AUCs ranging from .81 to .93 (Bell et al., 2015; Cuijpers et al., 2009; Friedman et al., 2005; Means-Christensen et al., 2005; Rumpf, Meyer, Hapke, & John, 2001; Yamazaki et al., 2005).

Investigations of AUCs for the MHI-5 are usually accompanied by estimates of the instrument’s sensitivity and specificity to depressive disorders based on the application of various MHI-5 cutoff scores, as well as the identification of an optimal or otherwise recommended cutoff score. Differences in recommended cutoffs are reported in the literature and likely reflect such things as differences in the populations investigated, the criterion by which depressed subjects were identified (e.g., DSM diagnosis vs. PHQ-9 scores), the version of the index (SF-36v1 vs. SF-36v2) that was used, and the scoring metric ($T$-scores vs. 0–100) and norms (1998 vs. 2009) that were used to score the index. However, when used to screen for depression, Yarlas et al. (2015) used an SF-36v2 MHI-5/MH $T$-score cutoff of 39 while employing the 2009 norms, based on investigations by Berwick et al. (1991) and Ware, Snow, Kosinski, and Gandek (1993). This cutoff is based on studies involving over 3,400 patients diagnosed with or without major depression or dysthymia and finding what equates to a $T$-score cutoff of 39 or below to be optimal, yielding a sensitivity of 66.8% and a specificity of 86.2%. As this $T$-score cutoff represents the one used by the Short Form surveys’ developer and publisher in its investigations, this author recommends that other MHI-5 users also use a $T$-score cutoff of 39. This cutoff may later be adjusted depending on its demonstrated utility in one’s particular patient population in terms of sensitivity and specificity.
Finally, being the SF-36v2 MH health domain scale, the MHI-5 derived from the SF-12v2-MH Enhanced version is scored to yield an SF-36v2 MH scale norm-based T-score (in addition to the SF-12v2 MH scale T-score) and is interpreted in the manner previously indicated for SF-36v2 MH scale, including with regard to the application of the norm-based, content-based, and criterion-based approaches. Thus, when administered as part of the SF-12v2-MH Enhanced, the MHI-5 can provide information that extends its utility beyond its usefulness as just a screener.

**Mental Health Severity Index**

The Mental Health Severity Index (Bjorner, Kosinski, & Raju, 2009) was developed to provide a means of predicting of Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) scores from MHI-5 scores, and then subsequently classifying the BDI-II scores into one of four levels of mental health problem severity: *minimal*, *mild*, *moderate*, or *severe*. Bjorner et al. found the correlation between predicted and actual BDI-II scores to be .71 (N = 471). Moreover, in a sample of 3,445 patients being screened for depression, they found that both the MHI-5 and predicted BDI-II scores correctly classified over 85% of the sample and yielded comparable sensitivity, specificity, and AUC values. Recommendation for the use of this index awaits further validation.

**Mental Component Summary (MCS) Measure**

Both the SF-36v2 and SF-12v2 MCS T-scores have also been investigated as another means of screening for depression. The SF-36v2 MCS measure has an internal consistency of .93 (Maruish, 2011) while the alpha coefficient for the SF-12v2 MCS measure is .88 (Maruish, 2012c). Based on the same sample of patients with or without major depression or dysthymia that were used to determine an optimal MHI-5/MH scale T-score cutoff, Ware and Kosinski (2001) found an SF-36v2 MCS T-score cutoff of 42 to be optimal, yielding a sensitivity of 73.7% and a specificity of 80.6% in identifying clinically depressed patients.

Several other studies have investigated the MCS measure as a tool for screening for depression and other mental disorders, or for documenting the effects of various types of mental disorders. For example, Bell et al. (2015) used SF-36v2 and PHQ-9 results from the 2013 US National Health and Wellness Survey (N = 75,000 adults) to determine the “appropriate” MCS and MH T-scores for identifying major depressive disorder. The AUC for both the MCS and MH measures was .89. The appropriate MCS cutoff score was determined to be 46, yielding a maximized sensitivity of 88.1% and specificity of 76.1%. For the MH scale, a cutoff score of 43 was identified with a sensitivity of 84.9% and a specificity of 77.8%. Also, one investigation of a sample of 272 low-income Chinese adults found mean T-scores to be significantly lower (p < .001) for those classified (by PHQ-9 scores) as having no to mild depressive symptoms (T = 51.98) compared to those with moderate to severe symptoms (T = 41.28; Lao, Chan, Tong, & Chan, 2015).

The MCS measure has also been found to be sensitive to other mental disorders in addition to depression. Mack et al. (2015) found SF-36v2 MCS T-scores to differ (p < .001) for those individuals found to have a 12-month mental disorder (T = 43.3), individuals with a lifetime mental disorder excluding the preceding 12 months (T = 49.9), and those found to be unaffected by mental disorder (T = 52.1) in a representative sample of over 4,400 German adults. Mean T-scores for individual 12-month diagnostic groups included 37.0 for major depressive disorder, 33.4 for dysthymia, 37.1 for GAD, 49.5 for alcohol abuse, and 44.1 for alcohol dependence. Also, compared to those with no diagnoses, those meeting the DSM-IV criteria for two or more mental disorder diagnoses were found to have a mean MCS T-score reduction (−11.40) that was almost twice that of the mean reduction of those with just one diagnosis (−6.08). Moreover, from the results of over 43,000 adults surveyed
as part of the 2001–2002 National Epidemiologic Survey on Alcohol and Related Conditions, Comer et al. (2011) found that those meeting the criteria for a DSM-IV generalized anxiety disorder (GAD) had a SF-12v2 mean MCS T-score of 37.9 and a mean MH/MHI-5 T-score of 38.1. Thus, identification of the risk of the presence of a mental disorder by low SF-36v2 or SF-12v2 MCS scores will require further exploration by the clinician to determine the nature of the disorder.

As with the MHI-5, recent investigations indicate there is a question of what MCS T-score cutoff score should be used to identify patients at risk for a depressive disorder. In general, this author again recommends using the cutoff identified by the developer and publisher of the Short Form surveys (i.e., $T \leq 42$) with MCS scores obtained from the SF-36v2. This same cutoff is also recommended for use with SF-12v2 MCS scores, as the MCS scores from the two surveys correlate .96 (Maruish, 2012c). Again, as with the MHI-5, the MCS cutoff for either survey can be adjusted later depending on its demonstrated usefulness in one’s particular primary care practice.

### Who to Screen

The SF-36v2 and SF-12v2 are appropriate for administration to both patient and nonpatient populations. The only criteria for whom to screen with either of these instruments is that they be 18 years or older and reading English at about the sixth to seventh grade level or higher. Use of the parent-completed SF-10 Health Survey for Children (Saris-Baglama et al., 2007) is recommended for patients aged 5–17 years. For non-English-speaking individuals, more than 170 and 150 translated or adapted versions of the SF-36v2 and SF-12v2, respectively, are available from the survey’s publisher. For English-speaking individuals reading below the recommended minimum grade level, interactive voice response (IVR) or interview versions of each of the two surveys can be used.

Beyond the minimum administration criteria inherent in the surveys themselves, to whom the SF-36v2 and SF-12v2 could or should be administered, is a decision that must be made by the PCP, psychologist, or integrated primary care team. One primary care team may decide that only certain patients should be administered a Short Form survey. These might include those who are suspected of having a significant physical or mental health disease or disorder, as indicated by their history or clinical presentation; or to those who have a diagnosed chronic or other physical or mental health condition that will likely be followed over a long period of time (e.g., one year). Another primary care team may feel that all patients being seen in the practice should be screened for any of several reasons. For example, the team may want to establish a baseline for possible treatment monitoring or outcomes assessment purposes; identify mental or physical health problems or functional limitations that might not otherwise be detected during a visit; provide a means of comparison from one episode of care to another; or determine if the patient’s perceived health status is different from that of their actual presentation.

### Implementation

How an effective screening process can or should be implemented in a primary care practice again will depend on the particulars of the practice. These include how easy it is to administer and score the instrument, the clinician’s ability to appropriately act on the information it provides, and staff acceptance and commitment to the process (see Chapter 1 of this book). Using brief instruments like the SF-36v2 and SF-12v2 can help lessen the impact of all of these issues on the implementation process.

Screening procedures can be designed in several ways to meet what is considered best for the primary care practice in terms of its patient population, standard workflow for providing services to patients, available financial and staff resources, current or planned information technology, and
perhaps other practice-specific variables. Administration of the SF-36v2 or SF-12v2 can take place either offsite before the patient’s scheduled office visit or on-site on the day of the visit. Offsite administration is possible through the use of mail-out/mail-back forms, mail-out/faxback forms, online through Internet or smartphone technology, or through an IVR system. This enables the clinician to have and review the scored Short Form results before the patient’s visit. In office, the patient may complete the survey via a paper form followed by entry into the Internet-based software scoring and reporting program via optical scanning or key entry, or through direct response entry into the scoring software via tablet or online computer administration. Either way, the PCP or psychologist can have the SF-36v2 or SF-12v2 results before the patient is seen. Overall, the availability of several options for administering, scoring, and reporting the results of the SF-36v2 and SF-12v2 should allow for the implementation of an effective and efficient screening process in most primary care practices.

**Guidelines for Decision-Making**

In general, PCS, MCS or health domain scores falling below a $T$-score of 40 should be considered significantly below the US general population mean and therefore worthy of follow-up by the PCP or psychologist. If the primary purpose of screening is to identify the likelihood of a mental health problem, then the focus should be on those health domain scales (VT, SF, RE, MH) and component summary measure (MCS) associated with mental health. Referring to the content and criterion-based interpretation tables for the SF-36v2 (Table 20.2) and SF-12v2 (Table 20.3), as well as examining individual item responses, will provide the clinician with leads to potential mental health-related problem areas. Based on their particular patient population, the primary care psychologist may find that the $T$-score (< 40) and minimum percentage (40%) cutoffs used to construct Tables 20.2 and 20.3 are inadequate; consequently, they may revise the summary measure and health domain scale $T$-score and percentage cutoffs to more accurately identify patients requiring further assessment or follow-up.

**Provision of Feedback to the Patient**

Both the SF-36v2 and the SF-12v2 can provide primary care patients with a wealth of information about their physical and mental health status. With norms based on large, representative national samples, the SF-36v2 or SF-12v2 findings can offer the patient a clear picture of where they stand on each measured health domain and dimension, in comparison to the US general population or, with the use of supplemental norms or benchmarks, to others similar to themselves on one or more variables (e.g., age, gender, presence of a specific chronic condition). The awareness of physical or mental health limitations relative to others may serve to help the patient acknowledge previously denied or overlooked health problems. This, in turn, may help the clinician to more effectively engage the patient in a plan of treatment to ameliorate or resolve problematic health issues. Eliciting feedback from the patient regarding their obtained scores as well as their individual Short Form item responses may also serve to provide further insight into problem areas and help build or strengthen an existing relationship with the patient.

Finn (2007) provides an excellent approach to the provision of feedback to patients regarding their psychological assessment results. This therapeutic assessment approach is a semi-structured, collaborative means of getting the most out of test results and also serves as a therapeutic intervention in and of itself. Although primarily designed for using the MMPI-2 with behavioral health patients, such an approach—perhaps in a modified form that considers in-office time limitations—should also serve as a basis for providing Short Form survey results to primary care patients.
USE OF THE SHORT FORM SURVEYS AS INSTRUMENTS FOR TREATMENT MONITORING AND OUTCOMES ASSESSMENT

Psychologists’ training and expertise in the area of psychological testing make them uniquely qualified to make significant contributions to patient services in primary care settings as well as in other medical and behavioral health settings. One important way this expertise can be applied is through the use of testing for the purposes of monitoring patients for changes in behavioral and/or physical health problems during the course of treatment as well as after the treatment has ended. Over the past two decades, Lambert and his colleagues have published significant contributions in the areas of treatment monitoring, provision of treatment progress feedback to clinicians, and outcomes assessment with psychological measures. Much of this large body of work is presented in an organized form in Lambert (2010), which this author considers a must-read volume for anyone engaged in behavioral health treatment regardless of the treatment setting.

Generic health status measures such as the SF-36v2 and SF-12v2 provide excellent means of tracking primary care and other types of outpatients over time to help determine what, if any, effect the patient-specific plan of treatment is having or has had after the completion of treatment. Either Short Form survey can provide a quick, general measure of any patient’s physical and mental health status as compared to the US general population as well as to others their own age and/or gender, and in many instances, to others experiencing similar physical or mental diseases or disorders. This enables an assessment of change in health status from several different perspectives.

General Monitoring and Outcomes Assessment Processes

In the general processes for treatment monitoring and outcomes assessment using either the SF-36v2 or SF-12v2, baseline measurement at the time of treatment initiation is the critical first step. Data from screening patients with one of the two surveys at the beginning of treatment can serve as the baseline measurement if the screening instrument provides data on the variable(s) one wants to use for monitoring and/or outcomes assessment purposes later on.

The results from the baseline SF-36v2 or SF-12v2 measurement, along with other information, can provide the data for developing a recovery curve. This curve could be generated with statistical procedures (e.g., hierarchical linear modeling) applied to Short Form data gathered from similar patients at specific points in time after initiation of treatment (e.g., every 2 weeks, every three treatment sessions). Alternately, in settings with no statistical resources or adequate data available, one could develop a less empirically based curve based on the clinician’s experience with similar patients. Remeasurement with the same Short Form survey at the predetermined points in time or treatment will yield new data that, if indicating the expected improvement, would lead one to continue on the planned course of treatment. Little or no improvement, or deterioration in health status, would suggest a need to modify treatment and then reassess the patient again at the next designated remeasurement point during the course of treatment. Here again, little or no improvement or further deterioration would indicate a need to modify treatment and then remeasure at the next designated point and so on, until treatment termination or referral to another care provider. Smith, Manderscheid, Flynn, and Steinwachs (1997) recommended that reassessment occur at those points in time that are clinically meaningful from the standpoint of the course of the disorder. That is, reassessment should occur at points in which clinical changes (e.g., remission, relapse) would be expected. They add that “assessment that is simply convenient for assessors will have little clinical utility” (Smith et al., 1997, p. 1036).
Part of the outcomes assessment process may include the treatment monitoring process. Attached to this process is remeasurement of the patient via the same Short Form survey one or more times after the termination of treatment. The purpose of posttreatment remeasurement is to determine if the changes noted at the time of treatment termination are indeed lasting changes. Common intervals of posttreatment remeasurement are 3, 6, and/or 12 months after termination. While not contributing anything to the patient’s treatment outcome, posttreatment data provide the clinician with further information about how successful they have been in the patient’s treatment and thus can aid in the treatment of future patients with similar problems.

There are two matters that bear consideration for those intending to use the SF-36v2 or SF-12v2 for treatment monitoring or outcomes assessment purposes. First, one must decide which of the 10 Short Form variables—two component summary measures and eight health domain scales—one wishes to track. Those employing the SF-12v2 should limit their use of data to the PCS and MCS measures, as the one- or two-item health domain scales may not prove to be reliable enough to serve as monitoring or outcomes variables. Those wanting to also track the MHI-5 (SF-36v2 MH scale) in addition to PCS and MCS can opt to administer the SF-12v2-MH Enhanced version of the survey. Those wishing to track one or more of the other health domain scales in addition to or instead of the PCS and MCS measures should employ the SF-36v2.

Second, a determination of how frequently patients will be reassessed must be made ahead of time. Remeasurement of interval-of-time-specific measures or subsets of items within them should be undertaken only after a period of time equivalent to or longer than the time interval to be considered in responding to the items has passed. For example, an instrument that asks the patient to consider the extent to which certain symptoms have been problematic during the past four weeks should not be readministered for at least four weeks. The responses from a readministration that occurs less than four weeks after the first administration would include the patient’s consideration of his or her status during the previously considered time period. This may make interpretation of the change of health status (if any) from one administration to the next difficult, if not impossible.

If one anticipates that one or more remeasurement intervals in the monitoring or outcomes assessment process will be less than four weeks, the acute (1-week recall) form of the SF-36v2 or SF-12v2 should be employed. This indeed is likely to be the case in many primary care settings. Note that the each of the acute forms for the SF-36v2 and SF-12v2 has its own reliability, validity, and content- and criterion-based interpretive data that are different from those reported in this chapter. Those using the acute form of either survey therefore should consult the appropriate User’s Manual (Maruish, 2011, 2012c) to review the relevant acute form psychometric and interpretive (i.e., content- and criterion-based) information. For those anticipating remeasurement intervals of less than seven days, another survey in the Short Form family of instruments, the eight-item SF-8 Health Survey (Ware, Kosinski, Dewey, & Gandek, 2001), offers a 24-hour recall version that may be suitable.

Determining Change

There are two general approaches to the analysis of treatment of individual patient outcomes data. The first is by determining whether changes in patient scores on outcomes measures are statistically significant. The other is by establishing whether these changes are clinically significant. The issue of clinical significance has received a great deal of attention in psychotherapy research during the past several years. This is at least partially owing to the work of Jacobson and his colleagues (Jacobson & Truax, 1991; Jacobson, Follette, & Revenstorf, 1984, 1986). Their work came at a time when researchers began to recognize that traditional statistical comparisons do not reveal a great deal about the efficacy of therapy.
From their perspective, Jacobson and his colleagues (Jacobson & Truax, 1991; Jacobson et al., 1984) felt that clinically significant change could be conceptualized in one of three ways. Thus, for clinically significant change to have occurred, the measured level of functioning following the therapeutic episode would have to either (1) fall outside the range of the dysfunctional population by at least two standard deviations from the mean of that population, in the direction of functionality; (2) fall within two standard deviations of the mean for the normal or functional population; or (3) be closer to the mean of the functional population than to that of the dysfunctional population. Jacobson and Truax viewed the third option as being the least arbitrary, and they provided different recommendations for determining cutoffs for clinically significant change, depending upon the availability of normative data.

At the same time, these same investigators noted the importance of considering the change in the measured variables of interest from pre- to posttreatment in addition to the patient’s functional status at the end of therapy. To this end, Jacobson et al. (1984) proposed the concomitant use of a reliable change index (RCI) to determine whether change is clinically significant. This index is nothing more than the pretest score minus the posttest score divided by the standard error of the difference of the two scores, with the level of significance set at the 95% confidence level ($p < .05$).

The developers of the SF-36v2 and SF-12v2 provide extensive discussions on different approaches to determining changes in component summary measure and, in the case of the SF-36v2, health domain scores in both individual-patient and group-level data (see Maruish, 2011, 2012c). Limiting the current discussion to determining significant differences in individual patient scores, it is worth noting that the developers provide tables indicating RCI values based on an assumption of uncorrelated baseline-to-follow-up measures as well as values based on alternate 80%, 90%, and 95% confidence levels and a .10 baseline-to-follow-up correlation. Values for all reported confidence levels are based on reliability data derived from the SF-36v2 and SF-12v2 2009 US general population normative data. The RCI 95% confidence level values for the SF-36v2 PCS and MCS measures are $\pm 5.5$ and $\pm 7.5$, respectively; for the SF-12v2, PCS and MCS RCI values are $\pm 9.7$ and $\pm 11.9$, respectively. For the developer-recommended 80% confidence level, the SF-36v2 values are $\pm 3.4$ and $\pm 4.6$, respectively, while the SF-12v2 80% confidence level values are $\pm 6.0$ and $\pm 7.4$, respectively. The SF-36v2 User’s Manual also provides these same confidence level values for the health domain scales.

Application of the cutoff values reported in the two User’s Manuals (Mariusch, 2011, 2012c) provides easy means of determining statistically significant change in a SF-36v2 or SF-12v2 score from one point in time to another (e.g., baseline to treatment session 3, baseline to treatment termination, termination to 3 months posttermination). The individual user must determine which set of values (i.e., which level of confidence) they will be most comfortable with and/or will be most useful to them. Applying one of the three previously indicated criteria options offered by Jacobson and Truax (1991) will also be required for determining clinically significant change. This author agrees with Jacobson and Truax’s recommendation for using the criteria requiring the score be closer to the mean of the functional population than to that of the dysfunctional population.

### DISEASE-SPECIFIC CONSIDERATIONS

One of the primary advantages of employing a generic measure of health status such as the SF-36v2 or SF-12v2 is that it is appropriate for use with all patients regardless of whatever physical or mental disorder they may be suffering from. The responses of anyone completing one of these Short Form surveys are scored using normative data collected in 2009. As such, everyone is compared to a large representative sample of the US adult general population on the eight health domains and two
general health dimensions that are assessed by both the SF-36v2 and SF-12v2. Being able to compare patients—regardless of their specific health problem—against the same standard affords clinicians many types of opportunities (e.g., obtain a general picture of their general patient population, fairly compare the health status of groups of patients to each other on important domains of physical and mental health). However, patient comparisons with similar individuals on important demographic or health variables further extends the usefulness of Short Form survey data.

Supplemental Normative and Benchmark Data

There are times when clinicians or researchers may wish to compare a patient’s Short Form survey results to those of others suffering from the same disease or disorder. This can provide information about whether the medical or behavioral health problem is affecting a particular patient or group of patients more or less than would be expected based on Short Form responses given by one or more relevant subsamples of the 2009 normative sample. To facilitate this, the surveys’ publisher developed benchmark data (means and standard deviations) for each of 40 disease or disorder groups based on members of the 2009 normative sample who indicated they were a member of one or more of those 40 chronic condition groups. Some of the medical condition subgroups for which benchmark data are available include diabetes, osteoarthritis, kidney disease, allergies, and hypertension; some of the behavioral health subgroups are depression, anxiety, and alcohol/drug use. Moreover, the literature is replete with publications that can serve as other sources of benchmark data for conditions. Unfortunately, some investigators have continued to report health domain scale data using the original SF-36 0–100 scoring metric instead of the current $T$-score method (for example, see Lam et al., 2013; Signorovitch et al., 2011; Thumboo et al., 2013), making the findings less useful to the reader.

Although not calculated with the data gathered as part of the 2009 Short Form survey norming study, the unique effects of each of 16 chronic conditions on each of the PCS, MCS, and eight health domain $T$-scores are available for both the SF-36v2 (Ware et al., 2007) and the SF-12v2 (Ware et al., 2010) based on the 1998 norming data. With very few exceptions, these effects are represented by $T$-score decreases on the component summary measures and health domain scores. For example, the estimated unique effect of having depression on the SF-36v2 MCS is a $T$-score that is 14.1 points lower than if depression was not present. Similarly, the estimated unique effect of having osteoarthritis is having an SF-12v2 PF $T$-score that is 5.7 points lower than if osteoarthritis was not present. Using multivariate regression analyses, differences between the 1998 normative sample mean health domain $T$-scores and 2009 normative sample mean health domain $T$-scores were found to not exceed 0.66 $T$-score points (Maruish, 2011). This suggests that the estimated unique effects of each of the 16 chronic conditions on the health domain scores in the 1998 SF-36v2 normative sample, as reported by Ware et al. (2007), may serve as a good approximation of the effects of these same health conditions on 2009 normative sample health domain scores.

Moreover, in addition to the condition-specific benchmarks, separate age, gender and age-by-gender norms are available for the both standard and acute form versions of both the SF-36v2 and SF-12v2. This allows for additional comparisons to others who are similar to the patient on these particular demographic variables.

Comorbid Conditions and Short Form Survey Results

For those working in primary care settings, it comes as no surprise that patients diagnosed with one chronic medical or behavioral health condition often also suffer from a comorbid behavioral health condition or chronic medical condition, respectively. For example, data from the 2001–2003 National
Comorbidity Survey Replication (NCS-R) found that 29% of those reporting a medical disorder also reported having a comorbid mental health disorder while 68% of adults with a mental disorder indicated they also had one or more general medical disorders (Alegria, Jackson, Kessler, & Takeuchi, 2003; Druss & Walker, 2011; Kessler et al., 2004). The literature contains numerous examples of how comorbid mental disorders can exacerbate chronic conditions (for example, see American Psychological Association [APA], 2014). Moreover, the influence of a mental illness and chronic disease on each other is evident. For example, the Centers for Disease Control and Prevention (CDC; 2009) reported that chronic conditions can worsen depressive symptoms while depressive disorders can actually lead to chronic conditions. In one study investigating the impact of various types of chronic conditions on HRQOL, as measured by the SF-36 PCS and MCS measures, Fortin et al. (2007) found that the mental health conditions had more of a negative impact on PCS scores than various physical health conditions had on MCS scores.

The results of another study conducted by Bayliss et al. (2012) demonstrated how comorbid conditions can affect scores on the SF-36v2’s PCS and MCS measures. In this study, Bayliss et al. used self-reported chronic condition data from a checklist of 26 chronic conditions (23 physical, 3 mental) taken from the 2009 SF-36v2 normative sample to identify four study groups: healthy condition (reporting no chronic conditions), physical health condition (reporting one or more physical conditions but no mental condition), mental health condition (reporting one or more mental conditions but no physical condition), and comorbid condition (reporting one or more physical conditions and one or more mental conditions). Relative to the healthy group, the physical condition group’s mean PCS T-scores were 6.4 and 7.5 points lower for males and females, respectively, and for the MCS scores, 1.7 and 2.1 points lower, respectively. Relative to the healthy group, the mental condition group’s mean MCS T-scores were 11.3 and 11.9 points lower for males and females, respectively, and for the PCS scores were 0.7 points lower and 0.3 points higher, respectively.

Additional decreases in SF-36v2 PCS and MCS mean T-scores were noted in both the physical and mental health condition groups when comorbid conditions were present (Bayliss et al., 2012). When a comorbid mental condition was also present, the mean PCS T-score decrement in the physical condition group increased (i.e., was lowered) by 4.5 points to a total score decrement of 10.9 points for males, and by 3.7 points to a total decrement of 11.2 points for females, relative to the healthy group. In this same comorbid condition group, the mean MCS T-score decrement in the physical condition group increased by 13.6 points to a total decrement of 15.3 points for males, and by 13.1 points to a total score decrement of 15.2 points for females, relative to the healthy group. Similarly, when a comorbid physical condition was also present, the mean PCS T-score decrement in the mental condition group increased by 10.2 points to a total score decrement of 10.9 points for males, and by 11.5 points to total score decrement of 11.2 points for females, relative to the healthy group. In this same comorbid condition group, the mean MCS T-score decrement in the mental condition group increased by 4.0 points to a total score decrement of 15.3 points for males, and by 3.3 points to a total score decrement of 15.2 points for females, relative to the healthy group.

Numerous other studies examining the effect of comorbid conditions are reported in the SF-36v2 and SF-12v2 literature. Examples of recent studies include Vietri et al.’s (2015) investigation of incremental burden of pain with comorbid depression; Mack et al.’s (2015) examination of the effect of multiple mental disorders on SF-36v2 domain scale and summary measure scores; Chin, Chan, Lam, Wan, and Lam’s (2015) SF-12v2 findings in depressed patients with and without physical comorbidities; and Fu et al.’s (2015) study of the effect of comorbidities on the quality of life of breast cancer survivors. Among Fu et al.’s findings were significant negative correlations between the number of verified comorbidities and overall quality of life as well as with the SF-36v2 PF and BP scales (p <
.01) and the VT and SF scales \((p < .05)\). Thus, when present, one must take into account the effect of comorbid medical and behavioral health conditions when interpreting SF-36v2 and SF-12v2 results.

### COMMONLY USED COLLATERAL INSTRUMENTS

As generic measures of health status and HRQOL, the SF-36v2 and SF-12v2 have been used in thousands of investigations of numerous physical and behavioral health diseases and disorders. Their widespread use in empirical studies attests to their proven psychometric integrity, their acceptance within the medical and behavioral health communities, and the importance of knowing patient status with regard to the assessed health domains in arriving at a more complete picture of the patient’s overall health status and HRQOL. Often, other psychometric instruments are administered along with SF-36v2 and SF-12v2 in published research studies as well as in clinical practice. Sometimes, these instruments are other generic measures such as the World Health Organization Quality of Life Scale (WHOQOL; World Health Organization, 1998), the Pain Impact Questionnaire-Revised (PIQ-R; Maruish, 2012a), the MOS Sleep Scale-Revised (MOS Sleep-R; Maruish, 2012b), or a visual analogue scale. In many other instances, the collateral instruments that are administered are disease- or condition-specific measures that allow the investigator or clinician to go beyond the patient’s general health status/HRQOL to understand other aspects of health that are commonly impacted by the particular health problem being studied or assessed. For example, patients with asthma may be administered the Asthma Impact Survey (AIS; Kosinski, Turner-Bowker, Bayliss, & Fortin, 2003) along with the SF-36v2. Similarly, the Headache Impact Test (HIT-6; Bayliss & Batenhorst, 2002) may be administered along with the SF-12v2 to patients suffering from debilitating headaches in order to determine the specific effect of their headaches on their functional health and well-being. Examples of other disease-specific measures that are often administered along with the Short Form instruments in published studies are the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC; Bellamy, Buchanan, Goldsmith, Campbell, & Stitt, 1988) and the Oswestry Low Back Pain Questionnaire (Fairbank, Couper, & Davies, 1980; also known as the Oswestry Disability Index [ODI; Fairbank & Pynsent, 2000])

Patients suffering from behavioral health problems may be administered popular disorder-specific psychological instruments along with a Short Form survey. Among those frequently reported in the literature are depression measures such as the Beck Depression Inventory-II (BDI-II; Beck et al., 1996), the Patient Health Questionnaire (PHQ-9; Kroenke, Spitzer, & Williams, 2001), the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977), and the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). The Symptom Checklist-90-Revised (SCL-90-R; Derogatis, 1994) also is reported to be used in several studies found in the literature. Most of these measures are addressed in detail in other chapters in this book.

### PRACTICAL CONSIDERATIONS FOR THE USE OF THE SHORT FORM SURVEYS IN PRIMARY CARE SETTINGS

Any psychological assessment instrument being considered for use in a primary care setting should be evaluated against the general criteria and considerations for selecting psychological measures, found in Chapter 5 of this book. Moreover, the demands made upon psychologists practicing in these settings—a fast-paced daily schedule, on-the-fly consultations, limited time with patients for assessment and treatment, and so forth—make certain considerations particularly important.
Staff Acceptance

As much as in other healthcare settings, the primary care staff’s acceptance of the use of the SF-36v2 or SF-12v2 will be dependent on a number of factors. For those staff members responsible for the administration and scoring of the survey and the reporting of the survey results, one of these factors is the degree to which the use of either survey can be easily integrated into the daily workflow of the practice. Fortunately, several options for Short Form administration, scoring, and reporting (e.g., pencil-and-paper, online, IVR, faxback, mail-out) are available, thus making the integration of the SF-36v2 or SF-12v2 into the practice as easy as possible regardless of the complexities of the practice’s specific workflow.

For the clinical staff, acceptance of the Short Form surveys or any psychometric instrument will be dependent on how useful they find the information that the survey results provide. Being generic health status/HRQOL instruments makes the SF-36v2 and SF-12v2 appropriate for use with all adult patients, regardless of the problems they present. Thus, the results of each of the surveys are easy to understand, help provide a picture of the patient from a biopsychosocial perspective, allow for comparisons of patients or groups of patients, can complement or be complemented with findings from other medical or psychological assessment procedures, and are empirically supported in the scientific literature.

Patient Acceptance

Generally, acceptance is not an issue when it comes to the administration of either survey to either medical or behavioral health patients. Both instruments are relatively brief, taking anywhere from 2 to 10 minutes to complete (depending on the patient and which survey is administered). In addition, the content and language of the items are neither offensive, threatening, nor provocative, and for the most part are in keeping with the types of information that are commonly elicited from patients in primary care and other medical settings.

Cultural Considerations

The patient’s cultural background generally is not a concern when administering either of the Short Form instruments. As indicated earlier, there are over 170 translated or adapted standard form versions of the SF-36v2 and over 150 versions of the SF-12v2. Several acute form versions of each of the two instruments are also available from the publisher. Moreover, the surveys’ specific item wording and content are not likely to be problematic as care was taken during the development and norming of the second version of the original SF-36 to ensure that the second version is indeed what many refer to as the “international version” (see Gandek & Ware, 1998).

Cost Considerations

Both the SF-36v2 and SF-12v2 are available under a license available from the surveys’ publisher. Although relatively inexpensive compared to some other computer-scored and interpreted instruments that the psychologist might use, cost is always a consideration regardless of the setting in which they are used.
**CASE EXAMPLE**

Following is an example of how the SF-12v2-Enhanced MH version was used in the assessment of a patient being seen in a primary care setting. The patient was seen two times during a two-week period. While the accompanying survey and screener data reported in Table 20.4 are the actual results of the two assessments performed during those visits, all other potentially identifying information has been changed to protect the identity of the patient.

<table>
<thead>
<tr>
<th>SF-12v2 Item #</th>
<th>1st Administration Response</th>
<th>2nd Administration Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Health rated very good</td>
<td>Health rated good</td>
</tr>
<tr>
<td>2a</td>
<td>Health limits moderate activities a little</td>
<td>Health limits moderate activities a little</td>
</tr>
<tr>
<td>2b</td>
<td>Health does not limit stair climbing</td>
<td>Health does not limit stair climbing</td>
</tr>
<tr>
<td>3a</td>
<td>Accomplished less than one would like a little of the time due to physical health</td>
<td>Physical health did not prevent accomplishing as much as one would like</td>
</tr>
<tr>
<td>3b</td>
<td>Limited in kind of work/activities some of the time due to physical health</td>
<td>Limited in kind of work/activities some of the time due to physical health</td>
</tr>
<tr>
<td>4a</td>
<td>Accomplished less than one would like a little of the time due to emotional problems</td>
<td>Emotional problems did not prevent accomplishing as much as one would like</td>
</tr>
<tr>
<td>4b</td>
<td>Had problems with doing work/activities less carefully a little of the time due to emotional problems</td>
<td>Emotional problems did not result in problems with doing work/activities less carefully</td>
</tr>
<tr>
<td>5</td>
<td>Pain interfered with work moderately</td>
<td>Pain interfered with work moderately</td>
</tr>
<tr>
<td>6a</td>
<td>Felt calm and peaceful most of the time</td>
<td>Felt calm and peaceful some of the time</td>
</tr>
<tr>
<td>6b</td>
<td>Had a lot of energy most of the time</td>
<td>Had a lot of energy most of the time</td>
</tr>
<tr>
<td>6c</td>
<td>Felt downhearted/depressed some of the time</td>
<td>Did not feel downhearted/depressed any of the time</td>
</tr>
<tr>
<td>7</td>
<td>Neither physical health nor emotional problems interfered with social activities</td>
<td>Neither physical health nor emotional problems interfered with social activities</td>
</tr>
<tr>
<td>8a</td>
<td>Was very nervous some of the time</td>
<td>Was very nervous some of the time</td>
</tr>
<tr>
<td>8b</td>
<td>Never felt so down in the dumps that nothing could cheer one up</td>
<td>Never felt so down in the dumps that nothing could cheer one up</td>
</tr>
<tr>
<td>8c</td>
<td>Felt happy most of the time</td>
<td>Felt happy most of the time</td>
</tr>
</tbody>
</table>

**Table 20.4 Summary of Assessment Results for Case Example**

<table>
<thead>
<tr>
<th>SF-12v2 Measures/ Scales</th>
<th>1st Administration T-Score</th>
<th>2nd Administration T-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td>MCS</td>
<td>52</td>
<td>59</td>
</tr>
<tr>
<td>PF</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>RP</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td>BP</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>GH</td>
<td>58</td>
<td>48</td>
</tr>
<tr>
<td>VT</td>
<td>59</td>
<td>59</td>
</tr>
<tr>
<td>SF</td>
<td>57</td>
<td>57</td>
</tr>
</tbody>
</table>

(Continued)
Olivia E. was a 30-year-old single female who was referred by her family physician to a psychologist who was part of the practice’s primary care team. The presenting problem was identified as anxiety, which she had been suffering from since her days as a college undergraduate. She reported symptoms of sleep problems, numbness in her feet, heart palpitations, feeling fearful that bad things might happen to her, and an exacerbation of the problems when she does not eat. She had been treated for her problem over the past several years with both psychotherapy and pharmacotherapy without experiencing any significant and lasting improvement. In addition to her psychological problems, Olivia also had been dealing with frequent headaches that she manages with acetaminophen. She denied any problems with alcohol or drugs.

Olivia had a bachelor’s degree in exercise science and had worked as a personal trainer at two local health clubs. Feeling that this career path would not lead anywhere, she applied for and was admitted to law school. Recently, she had graduated and began looking for a job, applying for positions as a lawyer as well as a paralegal. She had an interview for a paralegal job with a large legal firm the day before her second session with the psychologist and was waiting to see if she would be offered the position.

Olivia lived at home with her parents. She was the youngest of four children. All of her older siblings had completed college, were gainfully employed, and moved away from home. She had dated occasionally over the past several years but had never been in a relationship that lasted more than a few months. She met men through her friends or her work. Her parents were critical of her, feeling that she should have accomplished more with her life by this time, both personally and professionally. Olivia was disappointed with where she was in her life and accepted at least some of the responsibility for her current situation, citing both her lack of motivation and her often not assuming the responsibilities that come with being an adult.

Before each of her two sessions with the psychologist, Olivia was administered the PHQ-9, the GAD-7, and the acute (1-week recall) version of the SF-12v2-MH Enhanced MH form. As indicated earlier, the Enhanced MH version includes the all of the items from the MHI-5. The results of these measures are summarized in Table 20.4. Examination of baseline test scores suggested that Olivia was suffering from both physical and mental health problems. Her SF-12v2 BP score (40) was on the borderline between significant and questionable problems with bodily pain, which is consistent with her report of headaches. And although her MCS score (52), MH score (47), and Mental Health Severity Estimated BDI-II total score units (8) all fell within the average range, her PHQ-9 score (9) and GAD-7 score (10) indicated the presence of mild depression and moderate anxiety, respectively. Notable among Olivia’s SF-12v2 item responses were her reports of feeling both “downhearted/depressed” and “very nervous” some of the time; pain having interfered moderately with her normal work; accomplishing fewer activities due to both emotional and physical

<table>
<thead>
<tr>
<th>Screeners</th>
<th>1st Administration Scores/Ratings</th>
<th>2nd Administration Scores/Ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ-9</td>
<td>9 (mild)</td>
<td>3 (minimal)</td>
</tr>
<tr>
<td>GAD-7</td>
<td>10 (moderate)</td>
<td>5 (mild)</td>
</tr>
<tr>
<td>Mental Health Severity</td>
<td>8 (minimal)</td>
<td>7 (minimal)</td>
</tr>
</tbody>
</table>
problems and doing activities less carefully due to emotional problems a little of the time; and being limited by physical health in the kind of activities she could do some of the time.

Results from reassessment one week after the baseline session revealed general improvement in Olivia’s mental health status with a mild decrease in her physical health status. Her SF-12v2 MCS score increased by seven $T$-score points to 59 (more than two SEMs above the result from the first testing) and notable increases were seen in her RE and MH scores (10 and 6 points, respectively). Her PHQ-9 score (3) dropped to the minimally depressed range and her GAD-7 score (5) dropped to the mildly anxious level. Her Mental Health Severity Index also dropped by 1 raw score point. On the physical health side, Olivia’s PCS score dropped more than one SEM from the average range (48) to the borderline range (44). The BP score remained the same. Looking at individual aspects of functioning through item analysis, minor changes in a few areas in both the mental and physical health dimensions were apparent.

Unfortunately, Olivia did not return for her next appointment. One might speculate that she was offered the job she applied for and could not make the scheduled appointment. Another possibility was that she was experiencing an acute crisis and sought assistance to help her cope. As she began to feel better (as was indicated by the improvement in her emotional status during the second visit), perhaps because of an intervention provided during the first sessions, she did not feel a need to be seen again, at least not at that time. Regardless, her second assessment suggested the continued presence of mild anxiety that may need to be addressed in the future.

**SUMMARY**

The assessment of health status and HRQOL can provide information that complements other health findings about a patient being evaluated in a primary care setting. Arguably, the most widely used and respected generic measure of health status and HRQOL is the SF-36v2, as attested to by the number of published studies that include its use and its frequently being viewed as the gold standard for this type of measure. Empirical investigations have demonstrated the necessary reliability and validity of both it and its abbreviated version, the SF-12v2, for use for both clinical and investigative purposes. The primary purpose of this chapter has been to demonstrate how each of these Short Form instruments can be useful in primary care and other healthcare settings for identifying, monitoring, and assessing the outcomes of treatment for patients with behavioral health disorders. The surveys’ coverage of the effects of a patient’s health status on his or her physical, mental, and social aspects of functioning makes them ideal for inclusion in the context of a biopsychosocial assessment of primary care patients.

The use of either survey is facilitated by the availability of several administration and scoring options; standard (4-week recall) and acute (1-week recall) versions; and a significant amount of norm-, content-, and criterion-based interpretive information. Considerations for successfully integrating the surveys into primary care practices were addressed in this chapter, a systematic approach to interpreting Short Form survey results was presented, and a case example demonstrating the surveys’ potential utility in a primary care practice was provided.

**ACKNOWLEDGMENTS**

My thanks go to Maria Christoff and John Porcerelli for their assistance in the development of this chapter.
REFERENCES


Assessment of Health Status and HRQOL


Ware, J. E., Kosinski, M., & Keller, S. D. (1995). SF-12: How to score the SF-12 physical and mental summary scales (2nd ed.). Boston, MA: The Health Institute.


Zhou, K., Zhuang, G., Zhang, H., Liang, P., Yin, J., Kou, L., & Hao, M. (2013). Psychometrics of the Short Form 36 Health Survey Version 2 (SF-36v2) and the Quality of Life Scale for Drug Addicts (QOL-Dav2.0) in Chinese mainland patients with methadone maintenance treatment. PLOS ONE, 8(11), e79828. doi:10.1371/journal.pone.0079828


Martha S. Bayliss, Mark Kosinski, and Jakob B. Bjarner

The term “disease impact” has been used for decades as a synonym for health-related quality of life (HRQOL; Ware, 1984). In recent years, the label “disease impact measure” has been used to describe tools for patient-reported outcomes that are disease specific, use attribution to a specific disease in some or all of the questions, and include questions about the impact of the specific disease on role or social function. This chapter will discuss the concepts and measurement ideas underlying measures of disease impact. It will also present examples of disease impact measures: the Headache Impact Test (HIT-6; Kosinski et al., 2003), the Asthma Control Test (ACT; Nathan et al., 2004), and the Pain Impact Questionnaire (PIQ-6; Becker, Schwartz, Saris-Baglama, Kosinski, & Bjarner, 2007). For each measure, we describe the development, the psychometric properties, and the interpretation of scores. Finally, we will discuss the issues of disease impact assessment in case of multimorbidity and the practical applications in population health and primary care.

DISEASE MODEL

Wilson and Clearly (1995) provide a theoretical model for the relationship between different domains affected by disease (see Figure 21.1). Symptoms and well-being reflect the direct experience of the patient. Some symptoms will influence physical and mental function (i.e., the tasks a patient can perform). Typical indicators of physical function would be the ability to walk a mile, to walk up a flight of stairs, or to take a bath. Typical indicators of mental function would be the ability to remember things or to react adequately in a conversation. Physical and mental functions affect the ability to participate in more complex roles in society, that is, role and social function. Finally, symptoms and function will

Figure 21.1 Model for the Relation Between Disease and Different Domains of Disease Impact

influence general perceptions of health and quality of life. At each stage, the “response” to disease processes may be diluted due to the effect of comorbidities or other individual or environmental factors.

**MEASUREMENT ISSUES**

An ideal health measure would be valid, reliable, responsive to treatment effects, and yet relevant for all types of patients and for the general population so that disease impact can be compared across diseases. In practice, it is not always possible to optimize a measure with regard to all these properties. For example, responsiveness to treatment effects may be obtained by focusing the measure at patient experience that is close to the disease process. For example, an item in the ACT asks about an asthma-related symptom: “During the past 4 weeks, how often have you had shortness of breath?” A downside to the focus on specific symptoms is that it makes the item irrelevant to some other diseases (e.g., migraine). Thus, responsiveness may be obtained at the expense of the general relevance of the measure. This is a typical trade-off for disease-specific measures.

**Generic Versus Specific Measures**

Disease-specific measures are typically described as measures that are relevant only for patients with a specific disease. In other words, the measure is designed so that it is responsive to the effect of the particular disease and (to the extent possible) not responsive to anything else. Disease impact tests use different strategies to achieve this goal, including addressing symptoms that directly reflect the disease process and using the attribution to the particular condition in the question text.

**Measuring Symptoms or Impact**

Disease impact tests typically include items measuring symptoms as well as items measuring disease impact. Ideally, the symptoms should be specific to the disease. However, many symptoms such as pain or fatigue can occur with many different diseases. In that sense, they are generic. A disease impact test may sometimes ask questions about generic symptoms with attribution to the specific condition. For example, an item in the six-item HIT-6 asks: “In the past 4 weeks, how often have you felt fed up or irritated because of your headaches?”

For questions concerning function or general health perception, disease impact tests use attribution to the specific condition to focus the question. For example, a question in the HIT-6 asks: “In the past 4 weeks, how often did headaches limit your ability to concentrate on work or daily activities?” This approach relies on the assumption that patients can make correct inferences regarding the specific causes of symptoms and functional limitations. While patients’ assessment of causal mechanisms would not be trusted in the assessment of, for example, the efficacy of a drug, the case can be made that the association between symptoms and function are available for introspection to a very large extent, providing the patient with ample longitudinal data from which to make correct causal inferences. In practice, the use of causal attribution seems to work well for disease-specific scales in terms of validity and responsiveness.

**Assessment for Diagnosis Versus Measuring Impact**

Impact tests aim to be sensitive to differences in disease levels and impact of disease, but they do not usually aim to be a tool for differential diagnosis. A diagnostic test may include symptoms that are
rare, but very specific to the condition in question. An impact test may include symptoms that are
generic, provided they are very sensitive to disease severity. Also, an impact test will often use items
with attribution to a specific condition. Such items assume that the condition has been correctly
diagnosed and that the patient is aware of the diagnosis. Therefore, disease impact tests are usually
applicable only to patients with a diagnosed disease.

**Psychometric Properties**

The psychometric properties of the disease impact measures are specific to each measure and the
populations where it is used. However, many of the measures have used *item response theory* (IRT) for
item selection and development of scoring algorithms. Therefore, a short introduction to IRT and the
use of IRT in the development of disease impact measures is provided here. Also, one characteristic
has emerged across the disease impact measures: responses can be summarized in a single scale of
disease impact.

**Dimensionality**

Disease impact measures typically include items on many different health domains (symptoms,
physical, mental, role, and social function). Generic health measures are typically scores with a
different scale for each domain. However, factor analyses of disease impact tools have found sup-
port for one global scale of disease impact (Bjorner, Kosinski, & Ware, 2003a, 2003b). A possible
explanation for this finding is that construction of disease impact measures causes them to be
essentially unidimensional. By including items concerning disease-specific symptoms and items
with attribution to the particular disease, the disease impact measure may conform to a one-
dimensional model, where the latent dimension is disease severity. In contrast, a generic tool that
covers the same domains of physical, mental, role, and social function, but without disease attri-
bution, may show multidimensionality, because multiple diseases and conditions may influence
each domain.

**Item Difficulty and Measurement Precision**

The construction of several disease impact tools has been informed by insights from IRT. IRT com-
prises psychometric models for categorical items in multi-item scales (for a brief introduction, see
Edelen & Reeve, 2007). Figure 21.2 illustrates some key components of IRT models using two items
from the PIQ-6 as an example. The core element in IRT models are the *item response probability func-
tions* (IRPF), which are shown in the left column. For each level of pain impact, the functions show the
probability of selecting a particular response. For a PIQ-6 score of 50, the average score in the general
population, the most likely response to the first item is “very mild” pain (60% probability), although
the average respondent may also choose “mild” pain (29% probability), “moderate” pain (6% prob-
ability), or “none” (5% probability). Similarly, on the second item, the most likely responses whether
pain has interfered with work are “not at all” (65% probability) or “a little bit” (35% probability).

A second important element of IRT models is illustrated in the Item Information Functions part
of Figure 21.2. According to IRT, the contribution of an item to overall measurement precision is not
constant but depends on the match between the “difficulty” of the item and the health level (in this
case, level of chronic pain) of the respondent. Generally, an item is most useful, when the patient’s
response is not entirely predictable. An item is less useful if it is too “easy” or too “hard” compared
to the respondent’s level of pain. The item’s contribution to overall measurement precision can be
evaluated by the item information function, which is calculated from the IRPFs. For example, for
a respondent with very little or no pain (a PIQ-6 score around 40), the item PIQ-2 provides little
information, because it is almost certain that the response will be “not at all.” On the other hand, the item PIQ-2 is very useful for a person with more than average pain impact.

Information functions can be calculated for each item and also for a complete scale (the information functions for the complete scales are usually called test information functions and are calculated as the sum of the item information functions). The bottom graph shows test information functions for the PIQ-6 scale and for the Bodily Pain (BP) scale from the SF-36 (the BP scale is
constructed from the two items PIQ-1 and PIQ-2). The two horizontal lines illustrate the test information levels equivalent to reliabilities of .70 and .90, respectively. The figure shows that both scales achieve a precision equivalent to a reliability > .70 over a wide range of pain impact. However, only the PIQ-6 achieves high precision (equivalent to reliability > .90) over a wide range. These kinds of analyses have been used with most disease impact scales to select a combination of items that provide high measurement precision over a wide range of disease impact.

EXAMPLES OF DISEASE-IMPACT MEASURES

Headache Impact Test

Historically, episodic headaches have been evaluated using measures of the clinical features of headache such as pain intensity, frequency, duration, and other relevant symptoms such as nausea, aura, sensitivity to light, and sound (Nachit-Ouinekh et al., 2005). A more complete assessment, though, should include other burdensome aspects of the condition, particularly the impact of headache on how patients feel and function (Holmes, MacGregor, & Dodick, 2001). Several patient-completed assessments have been developed to accomplish this. The HIT-6 (Bayliss & Batenhorst, 2002) is a six-item, patient-based assessment designed to measure the impact of headaches on an individual’s functioning and well-being. It is easy for patients and clinicians to use and interpret. The HIT-6 is brief, reliable, and valid, yet comprehensive in its coverage of content areas relevant to the physical, social, and psychological impact of headaches.

Summary of Development

To maximize clinical usefulness, HIT-6 was constructed to have a high level of precision using the fewest questions. Typically, precision in health outcomes measurement is most easily achieved at the expense of brevity—with long surveys that require substantial time to complete. HIT-6 is distinguished from other headache impact instruments by coupling measurement precision with brevity, characteristics that make it ideal for use in busy clinical practices. Precision can be achieved with few questions in a brief, easily administered format (Kosinski et al., 2003).

Total HIT Question Pool

The starting point for selecting the HIT-6 items was the 54 items in the HIT item pool (Bjorner et al., 2003a; Ware, Bjorner, & Kosinski, 2000). The pool of questions derived primarily from widely used headache impact measures including the Headache Disability Inventory (HDI; Jacobson, Ramanadan, Norris, & Newman, 1995), the Headache Impact Questionnaire (HIMQ; Stewart, Lipton, Simon, Korff, & Liberman, 1998; Stewart, Lipton, Simon, Liberman, & Korff, 1999), the Migraine Disability Assessment Questionnaire (MIDAS; Stewart et al., 1998), and the Migraine-Specific Quality of Life Questionnaire (MSQ; Jhingran, Davis, LaVange, Miller, & Helms, 1998; Jhingran, Osterhaus, Miller, Lee, & Kirchdoerfer, 1998). In addition, the HIT question pool includes numerous new questions developed on recommendations from clinical experts in headache. The HIT question pool and the HIT-6 each cover six aspects of functioning that are seriously impacted by headache: pain, role functioning (the ability to carry out usual activities), social functioning, energy/fatigue, cognition, and emotional distress.

The HIT question pool was created to be a source of items for the dynamic, Internet-based HIT (Bjorner et al., 2003a, 2003b; Bjorner, Kosinski & Ware, 2003). Internet-based HIT was an advancement over previous surveys used to evaluate the impact of headache because it quantified, with unprecedented precision, the degree to which headache impacts a person’s life using a few questions
individualized for each respondent. Internet-based HIT uses modern psychometric methods (IRT) to select the most informative questions from the total HIT question pool.

**Psychometric Relevance**
IRT information functions and content validity were considered in selecting a subset of items for a short version of the HIT question pool. Using data from the National Survey of Headache Impact (NSHI), the best candidate items were evaluated on the basis of IRT information functions and content validity (in relation to widely used surveys and clinician judgment). Information functions express the contribution of each item to the overall test precision for various levels of headache impact. Items selected included a subset of 10 candidate items with IRT information functions that spanned a wide range of headache impact defined by the entire item pool, where more than 90% of recent headache sufferers scored. These 10 items also represented the six main content areas covered in widely used surveys (pain, social functioning, role functioning, vitality, cognitive functioning, and psychological distress).

**Clinical Relevance**
The next phase of development consisted of an independent review of the 10 candidate items by a panel of clinicians involved in the treatment of migraine headaches. In addition, the panel recommended 35 new items to be considered for the short form. A survey was fielded by telephone interview (n = 459) and over the Internet (n = 601) that included the original 10 candidate items and the additional 35 items. Participants for the telephone survey were sampled from a prescription database and had a prescription for migraine medication during the previous year.

An evaluation of whether the new items filled measurement gaps or extended the range of impact defined by the original 10 candidate items was undertaken. Also, the items were evaluated for clinical validity in relation to headache severity, the frequency of migraine symptoms, and work loss/productivity. From these analyses, six items were selected for the HIT-6. These six items covered the six content areas represented in the total HIT item pool, covered more than 50% of the range of headache impact measured by the total item pool, and were among the most valid items in tests of discriminant validity involving criterion measures of headache severity, frequency of migraine symptoms, and work loss productivity.

**Linguistic Translation**
The final phase in the development of the HIT-6 consisted of a linguistic translation of the six items into 27 languages. The translation of HIT-6 items followed the same methodology used including forward and backward translations of items and an independent review. This translation process resulted in modifications to all HIT-6 items and response options (Gandek, Alacoque, Uzun, Andrew-Hobbs, & Davis, 2003).

**Available Modes of Administration**
HIT-6 is available in a standard form that includes specific instructions, questions, and response choices for the questions. The standard HIT-6 form should be used because any change to the format could affect the way patients respond to HIT-6 questions. Use of the standard form helps ensure accuracy in the interpretation of HIT-6 scores. The standard HIT-6 survey form is available in many languages.

**Scoring the HIT-6**
HIT-6 is very simple to score, and scoring instructions can be found on the form itself. The Total Score is obtained by adding the numerical weights assigned to the responses indicate by the patients for each of the six items. All six items must be answered to calculate a score.
Basic Psychometric Information

It is important that measurement tools be reliable, that is, they yield stable, reproducible scores when measurements are repeated under identical conditions. Likewise, it is important that measurement tools be valid, that is, they measure the clinical construct they are intended to measure and not other constructs. In the HIT-6 reliability and validity studies, surveys were administered via the Internet to a general population of randomly selected respondents 18–65 years of age who reported a headache not attributed to cold, flu, head injury, or hangover in the past four weeks. Respondents completed questions at baseline (Time 1) and approximately 14 days after baseline (Time 2). The Time 1 survey consisted of questions from HIT-6 as well as questions commonly used in clinical practice as identified by a panel of clinical experts. The Time 2 survey included questions from HIT-6 and three additional questions to assess change in headache impact since Time 1. Questions measuring headache severity and migraine symptoms were also administered. The Time 1 survey was completed by 1,108 respondents, and the Time 2 survey was completed by 540 respondents for a 49.5% follow-up response rate.

Reliability of the HIT-6

Reliability is the degree of stability of a measurement when it is repeated under identical conditions. HIT-6 provides a stable, reproducible measure of the impact of headache on patients’ lives. It was shown to be reliable in each of several investigations.

1. Internal consistency reliability (Cronbach’s alpha) of HIT-6 total scores at Time 1 and Time 2 was .89 and .90, respectively.
2. Alternate-forms reliability was demonstrated by Pearson product moment correlation between HIT-6 scores and HIT-total scores of .78.
3. Test-retest reliability was strong at the level of the total HIT-6 scores. For the total sample, the correlation between Time 1 and Time 2 HIT-6 scores was .81 with an intra-class correlation of .78. For the subsample of patients who reported no change in headache severity between the two time samples, the test-retest correlation coefficient was .83 with an intra-class correlation of .90.
4. Test-retest reliability was also strong at the level of the individual question. Pearson correlation coefficients ranged from .60 to .71, and intra-class correlation coefficients ranged from .57 to .70. (Item-level reliability estimates are expected to be lower than scale-level reliability estimates.)

These data demonstrate that HIT-6 scores provide a stable, reproducible measure of the impact of headache on respondents’ lives.

Validity of the HIT-6

The method of known-groups validity was used to evaluate the validity of the HIT-6 in discriminating the impact of headache by comparing mean scores across groups known to differ in the underlying health concept of interest. The criterion variables used in these analyses included (1) migraine diagnosis based on presence of nausea, light sensitivity, and sound sensitivity; (2) migraine diagnosis based on the more stringent standard derived from all International Headache Society (IHS) symptom criteria (Olesen & Lipton, 1994); and (3) headache severity based on a 0–10 headache pain rating scale with patients classified as mild (score of 0, 1, 2, or 3), moderate (score of 4, 5, 6, or 7), or severe (score of 8, 9, or 10). In these validity analyses, HIT-6 was compared with a headache impact score calculated on the entire available HIT item pool (HIT-total, estimated from the total HIT pool of questions) and to scores from a two-item measure of headache-associated disability comprising two of the six items in the HIT-6 (Disability-2).

The HIT-total was the most valid on all three tests; however, the HIT-6 was also found to have excellent validity in distinguishing among headache sufferers as reflected in the following:
• Migraine diagnosis based on presence of nausea and light and sound sensitivity ($F = 392.36$; relative validity compared with HIT-Total = .87) ($F$ is an index of variability of scores between groups—in this case, between a group of respondents diagnosed with migraine and a group not diagnosed with migraine—divided by the variability of scores within a group. The $F$ statistic reflects the effects of nonrandom differences between groups relative to the effects that could reasonably be attributed to chance. The relative validity coefficient reflects the validity of a measure—HIT-6 in this case—relative to a reference standard, which in this case is HIT-total. The relative validity of the standard is 1.0; the relative validity of the comparison measurement is expressed as a fraction of 1.0.);
• Migraine diagnosis based on all IHS symptom criteria ($F = 405.72$; relative validity compared with HIT-total = .83); and
• Headache intensity of mild, moderate or severe ($F = 232.89$; relative validity compared with HIT-6 = 1.0).

Responsiveness to Change
Responsiveness of HIT-6 was evaluated among headache sufferers sampled from the general population. All study participants met IHS criteria for headache and all completed the total pool of HIT questions during telephone interviews conducted by trained interviewers. Random samples of study participants drawn from mild, moderate, and severe strata (according to IHS criteria) were assessed twice three months apart. In this early study of HIT, only one question from the HIT-6 was administered. A HIT-6 scale score was estimated from this one question by assigning the median HIT-total score to each of the response categories of the one item.

Evaluation of the responsiveness of HIT-6 to change in headache impact over time included (1) measurements of change in the severity of headaches using previously validated self-report scales; and (2) self-evaluations of change (i.e., better, same, or worse).

• For headache severity, three categories (mild, moderate, and severe) were defined as in previous studies. Respondents were classified by comparing these categories at baseline and follow-up.
• For self-evaluated change criteria, respondents were asked at follow-up to compare their headaches now with 3 months ago with respect to five distinct dimensions of headache impact: overall headache impact, physical condition, emotional condition, social/role functioning, and overall quality of life.

For each dimension of headache impact, three change groups were identified (those whose headaches were better, the same, or worse) and mean change in HIT-6 score was compared for these change groups. The results show that HIT-6 scores are responsive to change in headache severity. HIT-6 scores improved on average for those classified as better and declined on average for those classified as worse (Figure 21.3). Changes in HIT-6 scale scores were most responsive to changes in headache severity. The difference between headache sufferers with improved headache severity and those with worsening headache severity was 6.33 points—almost two-thirds of a standard deviation unit difference.

Interpreting HIT-6 Results
In establishing norms for HIT-6, telephone interviews were conducted by trained personnel in April and May 1999. Each interview consisted of (1) questions from four widely used questionnaires measuring headache impact and (2) questions that established headache severity and the likelihood of a migraine diagnosis. Potential interviewees were identified from a randomly generated list of telephone numbers from the 48 contiguous United States. Eligible respondents were 18 to 65 years of age, had suffered at least one headache not attributable to a cold or hangover in the 4 weeks prior to the interview, were conversant in English, and were physically and mentally competent to be
interviewed. Of 7,510 households contacted, 1,533 eligible respondents agreed to participate, with 1,016 (66.3%) completing the interview. The mean interview duration was 21.5 minutes.

Migraine status was based on either IHS symptom criteria or presence of key symptoms including nausea and photo- and phono-sensitivity. Headache pain severity was measured on an 11-point scale (0 to 10, with 10 corresponding to the worst pain). Scores of 1, 2, or 3 were categorized as mild; 4, 5, 6, or 7 were categorized as moderate; and 8, 9, or 10 were categorized as severe.

HIT-6 scores are easily interpreted, as summarized in the following:

- HIT-6 scores range from 36 to 78, with higher scores reflecting greater impact.
- The average score for headache sufferers is 50. Scores above 50 reflect worse than average headache impact and scores below 50 reflect better than average headache impact.
- If a patient’s score is higher than 50, a healthcare provider should discuss the patients’ headaches with them and determine whether intervention is warranted.

What do changes in scores mean?

- A change of 5 points or more is clinically meaningful.
- A 3-point change is noteworthy.
- Each 1-point change or difference in HIT-6 scores corresponds to one-tenth of a standard deviation.

Table 21.1 presents a brief guide to interpreting HIT-6 scores.

<table>
<thead>
<tr>
<th>HIT-6 Score</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 or more</td>
<td>Headaches are severely impacting the patient’s ability to function in everyday life. The patient may be experiencing disabling pain and other symptoms that are more severe than those of other headache sufferers.</td>
</tr>
<tr>
<td>(Severe Impact)</td>
<td></td>
</tr>
<tr>
<td>56 to 59</td>
<td>Headaches are substantially impacting the patient’s ability to function in everyday life. The patient may be experiencing severe pain and other symptoms, causing him or her to miss some time from family, work, school, or social activities.</td>
</tr>
<tr>
<td>(Substantial Impact)</td>
<td></td>
</tr>
<tr>
<td>50 to 5</td>
<td>Headaches are having some impact on the patient’s life but are probably not causing the patient to miss time from family, work, school, or social activities because of headache.</td>
</tr>
<tr>
<td>(Some Impact)</td>
<td></td>
</tr>
<tr>
<td>49 or less</td>
<td>Headaches are having little or no impact on the patient’s life.</td>
</tr>
<tr>
<td>(Little or No Impact)</td>
<td></td>
</tr>
</tbody>
</table>
Applications of HIT-6
HIT-6 is well suited to measure headache impact in a variety of settings ranging from primary care and other clinical practices to disease management and clinical trials. Physicians and other healthcare providers, payers, clinical researchers, public health analysts, and policy officials each will find uses for HIT-6. Brevity and ease of administration are key to the ability of HIT-6 to be integrated easily into normal primary care practice.

HIT-6 can easily be integrated into ordinary clinical practice in several applications. Point-in-time HIT-6 measures can help to identify patients whose headaches are impacting their lives such that they require diagnosis and intervention, facilitate communication between patients and healthcare providers, and tailor treatment strategy to the severity of headache impact.

With repeated use in a given patient, HIT-6 can be used to monitor changes in headache impact over time and determine the effects of a treatment strategy on headache impact.

The brevity of the HIT-6 makes it a very useful survey for use in monitoring the health of general or specific populations. Primary care and other settings can use HIT-6 in conjunction with other measures (such as symptoms, healthcare utilization, and treatment compliance) to assess and track the burden of headache and the benefits of treatment.

Asthma Control Test
Over the past decade an improved understanding of the pathophysiology underlying asthma and the emergence of medications to more effectively prevent acute exacerbations has led clinicians to shift their focus from managing acute attacks to achieving asthma control. Asthma control reflects the level of minimizing the frequency and intensity of symptoms and functional limitations associated with asthma. Despite this shift in focus in the management of asthma, a high percentage of individuals with asthma in the United States continue to have uncontrolled asthma. On average, 38% of children and 50% of adults with asthma had uncontrolled asthma during the years from 2006 through 2010 (Centers for Disease Control and Prevention, 2010). The adverse impacts of asthma have been shown to be most profound among individuals with chronically uncontrolled asthma. These individuals are more frequently admitted to the hospital and have regular absences from work and school (Juniper et al., 2004). Between 80% to 85% of asthma deaths happen among patients with poorly controlled severe disease (Papiris, Kotanidou, Malagari, & Roussos, 2002), and there is a strong association between increased recurrences of hospitalization and uncontrolled asthma (Hartert et al., 2002).

Achieving asthma control requires an intense effort on the part of both physician and patient, and constant monitoring of asthma symptoms and their impact on the patient’s life. In the Guidelines for the Diagnosis and Management of Asthma published by the National Heart, Lung, and Blood Institute (US Department of Health and Human Services & National Institutes of Health, 2007), periodic assessment and ongoing monitoring of asthma are recommended to determine if the goals of asthma therapy are being met. These goals include minimizing asthma symptoms; reducing limitations in work, school, or other physical activities; and preventing acute attacks, emergency treatment, and hospitalization. Both physician assessment and patient self-assessment are regarded as important components for monitoring asthma. Measurements are recommended in these areas: signs and symptoms of asthma; pulmonary function (spirometry and peak flow monitoring); functional status; history of asthma exacerbations; and pharmacotherapy.

According to the Guidelines (NHLBI, 2007), “any assessment of the patient’s symptom history should include at least three key symptom expressions”: (1) daytime asthma symptoms (including
wheezing, cough, chest tightness, or shortness of breath), (2) nocturnal awakening as a result of asthma symptoms, and (3) asthma symptoms early in the morning that are not improved 15 minutes after inhaling a short-acting beta₂-agonist. To assess how the disease expression and the lack of asthma control impact a patient’s functioning, the Guidelines recommend several key areas be periodically assessed, including (1) any missed work or school due to asthma, (2) any reduction in usual activities due to asthma, and (3) any disturbances in sleep due to asthma.

Despite the existence of these guidelines, it has been reported that the level of asthma control continues to be overestimated for many patients, resulting in the undertreatment of the condition and increased occurrences of acute attacks, disruption of the patient’s ability to work or go to school, and a higher probability of emergency care and hospitalization (Nathan et al., 2004; Peters, Ferguson, Deniz, & Reisner, 2006). Worldwide, achieving asthma control continues to be an ongoing problem in the management of asthma (To et al., 2012).

Recognizing that the level of asthma control was often overestimated by both patients and physicians was a sign that asthma treatment guidelines alone were not enough to ensure the proper assessment of asthma control. This deficiency in the assessment of asthma control prompted the development of a simple method for quantifying asthma control by both patients and physicians. The Asthma Control Test (ACT) was developed around the time that the NHLBI Guidelines (2007) made the paradigm shift away from monitoring the severity of asthma to the assessment and monitoring of asthma control in the early 2000s. The ACT is a brief, five-item, patient-completed, clinically validated instrument designed to assess the dimensions of asthma control outlined in the NHLBI Guidelines—asthma symptoms, utilization of rescue medications, and the impact of asthma symptoms on everyday functioning—and support the premise that asthma control is a multidimensional construct (Nathan et al., 2004).

**Summary of Development**

The ACT was designed to produce a valid measure reflecting the multidimensional nature of asthma control that is easy for clinicians to use, and for clinicians and patients to interpret. Leading asthma specialists and primary care physicians treating patients with asthma participated in the development of the ACT by specifying the components of asthma control that should be assessed by the survey and by participating in the definition of a criterion measure of asthma control that was used to evaluate the instrument’s performance. This working group of specialists also helped guide the development of specific survey questions and participated in the design and implementation of the clinical validation study. The final survey instrument fielded in the development and clinical validation study consisted of 22 questions reflecting the multidimensional nature of asthma control and capturing the components of symptoms and impairment consistent with those outlined in the NHLBI asthma treatment guidelines.

The development and clinical validation study was conducted at six asthma specialty practices. Each of the participating practices included one of the asthma specialists involved in the study’s working group. This was done with the belief that on-site availability of a specialist involved with the design of the clinical validation study would help ensure standardized application of the study protocol, including the process used to arrive at the specialist’s rating of asthma control used as the criterion measure for selecting and validating the final set of questions for the ACT. Eligible patients for the clinical validation study were 12 years of age and older who had been diagnosed with asthma and who had been seeing the participating asthma specialist for at least one year. This latter requirement helped to ensure that the specialist’s rating of control, the criterion measure, was as accurate as possible given their familiarity with the patient and his or her asthma condition.
A total of 471 participants completed the survey instrument during a routine, previously scheduled physician office visit. After the participant completed the survey instrument, office staff completed pre-bronchodilator measurements of forced expiratory volume (FEV₁), and the asthma specialist conducted a clinical interview of the patient. At the end of the clinical interview, the asthma specialist rated the patient’s asthma control on a 5-point scale ranging from “not controlled at all” to “completely controlled.” This rating was based on how well the goals of asthma therapy were being met, as outlined in the NHLBI Guidelines (US Department of Health and Human Services, National Institutes of Health, & National Heart, Lung, and Blood Institute, 2002) and as determined by the patient’s history, physical examination, and FEV₁.

Stepwise logistic regression methods were used to select a subset of the 22 survey instrument questions that demonstrated the greatest validity in discriminating between patients who differed in the specialist’s rating of asthma control. Because the distribution of the specialist’s rating of asthma control was skewed toward better control, a dichotomous variable was derived from the specialist’s rating of control and used as the dependent variable for the item-selection analysis. Response values for the 22 questions on the survey instrument were entered into a logistic regression model in a forward stepwise fashion. The criterion for model entry was the ability to significantly discriminate between controlled and uncontrolled patients \((p < .05)\). The results of this analysis identified five items that significantly discriminated between controlled and uncontrolled patients. These five items were selected to represent the ACT. They assess the patient’s self-rating of their asthma’s interference with sleep and ability to get things done, the degree to which it is controlled, and the frequency of shortness of breath and use of a rescue inhaler.

Administering the ACT

The format of the ACT makes it easy to complete. With just five questions on a one-page form, the ACT can be easily integrated into many settings, including routine clinical practice, disease management programs, public awareness campaigns, clinical trials, and quality improvement initiatives. The survey instrument is self-administered and can be completed in a clinical practice setting, at home, or in any location. The ACT can be administered in pencil-and-paper format, during telephone or face-to-face interviews, or by almost any electronic application such as interactive voice response (IVR) systems, tablets, web applications, or through electronic medical record (EMR) systems. No special training is required to administer or complete the ACT.

The standard ACT is intended for use with adolescents (12 years and older) or adults with asthma. Interpretation guidelines developed for the ACT are based on two independent clinical studies conducted with asthma patients 12 years and older (Nathan et al., 2004; Schatz et al., 2006). More recently, a caregiver administered version of the ACT, the Asthma Control Test-Caregiver Report (ACT-CR), has been validated for use in children between the ages of 2 and 11 years of age (Nkoy et al., 2013). The ACT-CR uses a 1-week recall period as opposed to the 4-week recall period used with the standard ACT.

The response choices of each ACT item are scaled from 1 to 5. Scoring of the ACT consists of summing the response choice values across the five items of the ACT to produce a scale score that ranges from 5 (poorly controlled asthma) to 25 (completely controlled asthma). All ACT items must be answered to calculate a score. Methods to score the ACT with incomplete item responses have not been developed. However, it is not recommended that scores be calculated with imputation methods when item responses are missing, as it is not known how imputed scores effect validity and interpretation. Like the ACT, the ACT-CR is scored so that a higher score indicates better asthma control. Scoring the ACT-CR involves the same two steps as scoring the standard ACT.
Basic Psychometric Information

Two pivotal studies were conducted to evaluate the reliability, validity, and interpretation of ACT scores (Nathan et al., 2004; Schatz et al., 2006). One of these studies (Schatz et al., 2006) also included a longitudinal component allowing investigators the opportunity to evaluate the responsiveness of the ACT to changes in asthma control and lung function. A third study was later conducted to determine the minimal clinical important difference (MCID) in ACT scores as a means to facilitate meaningful interpretations of differences in ACT scores at a point in time and changes in ACT scores over time (Schatz et al., 2009).

The reliability, validity, and interpretation of ACT scores were first evaluated in a sample of asthma patients who had been under the care of their asthma specialist for at least one year (Nathan et al., 2004). Reliability was evaluated using internal consistency methods and Cronbach's alpha. Tests of validity were based on the correlations between ACT scores and the asthma specialist's rating of control and percent predicted FEV₁ values. In addition, tests of validity were conducted to evaluate how well ACT scores discriminated between groups that differed in clinician ratings of asthma control, FEV₁ values, and whether the specialist changed the patient's therapy as a result of the visit using the method of known-groups validity. Lastly, tests were conducted to determine the optimal cut-point score on the ACT for screening patients with uncontrolled asthma.

In the total study patient population (n = 436), the ACT was shown to have good internal consistency reliability (r = .84) despite the items being multidimensional (symptoms, functional limitation, rescue medication use, and self-rating of control). Similarly, ACT scores were equally reliable in a sample of patients (n = 133) who had uncontrolled asthma (r = .83) and a sample of patients (n = 303) whose asthma was well controlled (r = .79).

Moderate to low correlations were observed between ACT scores, FEV₁ values, and the asthma specialist's rating of asthma control. The highest correlation observed was between ACT scores and the asthma specialist's rating of control (r = .45, p < .001). The correlation between the asthma specialist's rating of asthma control and percent predicted FEV₁ values was moderate (r = .37, p < .001) and the correlation between ACT scores and FEV₁ values was low (r = .19, p < .001). These correlations supported the contention that asthma control is not inferred from the clinical measure of airway function alone. The higher correlation between the specialists rating of control and the ACT confirms that the other dimensions of control (activity limitation, rescue medication use, symptoms) are all considered in the specialist's determination of asthma control.

Results of the tests of known-groups validity showed significant differences in mean ACT scores across groups of patients who differed on each clinical measure related to asthma control according to hypotheses, thus lending support to the discriminant validity of ACT scores (Table 21.2). Mean ACT scores were significantly lower among patients with poorer control as judged by the specialist than among patients with more favorable control ratings (F = 34.5, p < .001). Likewise, patients with poorer lung function (percent predicted FEV₁) scored significantly lower on the ACT than patients with better lung function (F = 4.3, p = .0052). Finally, patients whose therapy was stepped up as a result of the visit with the specialist scored significantly lower than patients with no change in therapy or a step down in therapy (F = 40.4, p < .0001).

Analyses conducted to identify the optimal cut-point score on the ACT for screening patients with uncontrolled asthma identified a score of 19 with the highest area under the receiver operating characteristic (ROC) curve and best balance between sensitivity and specificity. At a cut-point score of 19 (and below) the ROC for the ACT was .727, with a sensitivity of 69.2% and specificity of 76.2%. Lower cut-point scores yielded higher specificity, while higher cut-point scores yielded higher sensitivity (Table 21.3).
A follow-up longitudinal study was conducted to evaluate the reliability, validity, interpretation, and responsiveness of ACT scores in a sample of asthma patients new to the care of an asthma specialist (Schatz et al., 2006). Reliability was evaluated using internal consistency and test-retest methods. The validity of the ACT was evaluated by computing correlations between ACT scores and the specialist’s rating of asthma control, scores on the Asthma Control Questionnaire (ACQ; Juniper, O’Byrne, Guyatt, Ferrie, & King, 1999), and percent predicted FEV₁ values. The method of known-groups validity was used to evaluate the discriminant validity of the ACT. Differences in ACT scores were compared across patients who differed on the following three criterion measures: (1) specialist global rating of asthma control, (2) change in patient’s therapy, and (3) percent predicted FEV₁ values. The results are summarized in Tables 21.2 and 21.3.

### Table 21.2 Comparison of Mean (Standard Deviations) ACT Scores Across Groups Differing in Asthma Control

<table>
<thead>
<tr>
<th>Specialist Rating of Control</th>
<th>Not controlled at all (n = 2)</th>
<th>Poorly controlled (n = 28)</th>
<th>Somewhat controlled (n = 103)</th>
<th>Well controlled (n = 224)</th>
<th>Completely controlled (n = 79)</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT sum scoring</td>
<td>7.5 (0.7)</td>
<td>15.5 (4.4)</td>
<td>16.9 (4.7)</td>
<td>20.8 (3.4)</td>
<td>21.5 (3.9)</td>
<td>34.5</td>
<td>0</td>
</tr>
<tr>
<td>Change Patient’s Therapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stepped Down (n = 52)</td>
<td>20.6 (4.1)</td>
<td>20.8 (3.6)</td>
<td>16.6 (5.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Predicted FEV₁ Values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30–59% (n = 48)</td>
<td>18.3 (4.7)</td>
<td>19 (4.8)</td>
<td>19.9 (4.4)</td>
<td>20.9 (3.5)</td>
<td></td>
<td></td>
<td>4.3</td>
</tr>
<tr>
<td>60–79% (n = 113)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80–100% (n = 192)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101–140% (n = 87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 21.3 Summary of the Performance of ACT Cut-Point Scores in Screening for Uncontrolled Asthma (n = 436)

<table>
<thead>
<tr>
<th>Cut-Point Score</th>
<th>Odds Ratio</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive Predictive Value (%)</th>
<th>Negative Predictive Value (%)</th>
<th>% Correctly Classified</th>
<th>Area Under ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 10</td>
<td>7.41</td>
<td>9.0</td>
<td>98.7</td>
<td>75.0</td>
<td>71.2</td>
<td>71.3</td>
<td>0.539</td>
</tr>
<tr>
<td>≤ 11</td>
<td>7.04</td>
<td>14.3</td>
<td>97.7</td>
<td>73.1</td>
<td>72.2</td>
<td>72.3</td>
<td>0.56</td>
</tr>
<tr>
<td>≤ 12</td>
<td>9.11</td>
<td>21.8</td>
<td>97.0</td>
<td>76.3</td>
<td>73.9</td>
<td>74.1</td>
<td>0.594</td>
</tr>
<tr>
<td>≤ 13</td>
<td>7.96</td>
<td>29.3</td>
<td>95.1</td>
<td>72.2</td>
<td>75.4</td>
<td>75.0</td>
<td>0.622</td>
</tr>
<tr>
<td>≤ 14</td>
<td>7.58</td>
<td>36.1</td>
<td>93.1</td>
<td>69.6</td>
<td>76.8</td>
<td>75.7</td>
<td>0.646</td>
</tr>
<tr>
<td>≤ 15</td>
<td>9.27</td>
<td>44.4</td>
<td>92.1</td>
<td>71.1</td>
<td>79.0</td>
<td>77.5</td>
<td>0.682</td>
</tr>
<tr>
<td>≤ 16</td>
<td>9.41</td>
<td>52.6</td>
<td>89.4</td>
<td>68.6</td>
<td>81.1</td>
<td>78.2</td>
<td>0.71</td>
</tr>
<tr>
<td>≤ 17</td>
<td>7.68</td>
<td>57.9</td>
<td>84.8</td>
<td>62.6</td>
<td>82.1</td>
<td>76.6</td>
<td>0.714</td>
</tr>
<tr>
<td>≤ 18</td>
<td>6.72</td>
<td>62.4</td>
<td>80.2</td>
<td>58.0</td>
<td>82.9</td>
<td>74.8</td>
<td>0.713</td>
</tr>
<tr>
<td>≤ 19</td>
<td>7.20</td>
<td>69.2</td>
<td>76.2</td>
<td>56.1</td>
<td>84.9</td>
<td>74.1</td>
<td>0.727</td>
</tr>
<tr>
<td>≤ 20</td>
<td>6.31</td>
<td>75.9</td>
<td>66.7</td>
<td>50.0</td>
<td>86.3</td>
<td>69.5</td>
<td>0.713</td>
</tr>
<tr>
<td>≤ 21</td>
<td>5.01</td>
<td>82.7</td>
<td>51.2</td>
<td>42.6</td>
<td>87.1</td>
<td>60.7</td>
<td>0.669</td>
</tr>
<tr>
<td>≤ 22</td>
<td>4.35</td>
<td>87.9</td>
<td>37.3</td>
<td>38.1</td>
<td>87.6</td>
<td>52.8</td>
<td>0.626</td>
</tr>
<tr>
<td>≤ 23</td>
<td>5.32</td>
<td>94.0</td>
<td>25.4</td>
<td>35.6</td>
<td>90.6</td>
<td>46.3</td>
<td>0.597</td>
</tr>
<tr>
<td>≤ 24</td>
<td>3.31</td>
<td>94.7</td>
<td>15.1</td>
<td>33.0</td>
<td>87.0</td>
<td>39.6</td>
<td>0.551</td>
</tr>
<tr>
<td>Continuous</td>
<td>0.78</td>
<td>44.4</td>
<td>92.01</td>
<td>71.1</td>
<td>79.0</td>
<td>77.5</td>
<td>0.774</td>
</tr>
</tbody>
</table>

Note. Specialist global rating of not controlled at all, poorly controlled, or somewhat controlled.
rating of asthma control, (2) percent predicted FEV$_1$ values, and (3) asthma specialist’s treatment recommendation. Screening accuracy was evaluated against the specialist’s rating of asthma control. The responsiveness of ACT scores was evaluated against changes in the specialist control rating, percent predicted FEV$_1$ values, and ACQ scores over three months following the baseline visit using correlation and known-groups validation methods.

The internal consistency reliability of ACT scores was .85 at the baseline visit ($n = 313$) and .79 at the follow-up visit ($n = 248$). Among patients who showed no change (stable asthma control) in specialist rating from baseline to follow-up visit the test-retest reliability of the ACT was .77 (intra-class correlation between ACT scores at baseline and follow-up visits).

In support of convergent validity, ACT scores were found to correlate significantly with the specialist rating of asthma control ($r = .52$, $p < .001$), scores from the Asthma Control Questionnaire ($r = -.89$, $p < .001$), and percent predicted FEV$_1$ values ($r = .31$, $p < .001$). The discriminant validity of the ACT was supported by results of known-groups validity test results. Mean ACT scores differed significantly across groups of patients who differed on each clinical measure related to asthma control according to hypotheses (Table 21.4). Mean ACT scores were significantly lower among patients with poorer control as judged by the specialist than among patients with more favorable control ratings ($F = 27.5$, $p < .001$). Likewise, patients with poorer lung function (percent predicted FEV$_1$) scored significantly lower on the ACT than patients with better lung function ($F = 16.1$, $p < .001$). Finally, patients whose therapy was stepped up as a result of the visit with the specialist scored significantly lower than patients with no change in therapy or a step down in therapy ($F = 32.9$, $p < .0001$).

Results of analyses to evaluate the ability of the ACT to screen for patients with asthma control problems found a cut-point score of 19 and below to have the highest area under the ROC curve (ROC = .71) and the best balance between sensitivity (71.3%) and specificity (70.8%) (Table 21.5). This finding confirmed the previously derived cut-point score of 19 and below from the initial development and clinical validation study (Nathan et al., 2004). In addition, analyses were conducted to identify an optimal cut-point score for distinguishing between patients with “somewhat controlled” asthma and patients with “poorly controlled” asthma. Results showed that a cut-point score of 15 and below yielded the highest percentage of patients correctly classified (68.2%) and the highest area

<table>
<thead>
<tr>
<th>Table 21.4 Discriminant Validity Tests of Mean ACT Scores ($n = 301$)</th>
<th>Mean (SD) ACT Score</th>
<th>F-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specialist Assessment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not controlled at all ($n = 11$)</td>
<td>13.3 (4.7)</td>
<td>27.5*</td>
</tr>
<tr>
<td>Poorly controlled ($n = 59$)</td>
<td>14.5 (4.8)</td>
<td></td>
</tr>
<tr>
<td>Somewhat controlled ($n = 87$)</td>
<td>17.4 (4.7)</td>
<td></td>
</tr>
<tr>
<td>Well controlled ($n = 116$)</td>
<td>20.3 (3.4)</td>
<td></td>
</tr>
<tr>
<td>Completely controlled ($n = 28$)</td>
<td>21.5 (3.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Percent predicted FEV$_1$</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60% ($n = 23$)</td>
<td>14.4 (4.7)</td>
<td>16.1*</td>
</tr>
<tr>
<td>60% to 79% ($n = 62$)</td>
<td>15.7 (4.8)</td>
<td></td>
</tr>
<tr>
<td>80% to 100% ($n = 133$)</td>
<td>19.2 (4.6)</td>
<td></td>
</tr>
<tr>
<td>&gt;100% ($n = 83$)</td>
<td>19.5 (4.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Therapy Recommendation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stepped down ($n = 13$)</td>
<td>17.9 (5.6)</td>
<td>32.9*</td>
</tr>
<tr>
<td>No change ($n = 141$)</td>
<td>20.4 (3.6)</td>
<td></td>
</tr>
<tr>
<td>Stepped up ($n = 147$)</td>
<td>16.1 (5.0)</td>
<td></td>
</tr>
</tbody>
</table>

* $p < .001$. 


under the ROC curve (ROC = .68). Based on these findings along with the findings from the initial development and validation study, a cut-point score of 19 and below is recommended for screening asthma control problems in general, and a cut-point score of 15 and below is recommended for identifying poorly controlled asthma.

The responsiveness of ACT scores was demonstrated with correlations between changes in ACT scores and changes in the specialist control rating, percent predicted FEV1 values, and ACQ scores over three months. Changes in ACT scores were moderately correlated with changes in specialist rating of asthma control ($r = .44, p < .001$) and highly correlated with changes in ACQ scores ($r = -.81, p < .001$). The correlation between ACT scores and percent predicted FEV1 values was considerably lower ($r = .29, p < .001$). Mean changes in ACT scores across groups of patients differing in the level of change in specialist ratings of asthma control and clinically meaningful changes in percent predicted FEV1 values and ACQ scores provided additional evidence of the responsiveness of the ACT (Table 21.6). Mean changes in ACT scores differed significantly across groups of patients that differed in the level of change in specialist asthma control ratings ($F = 24.2, p < .001$), percent predicted FEV1 values ($F = 32.9, p < .001$), and ACQ scores ($F = 71.0, p < .001$). Consistent with hypotheses, ACT scores showed significant improvement among patient groups whose asthma control rating, percent predicted FEV1, and ACQ scores improved. Conversely, ACT scores worsened among patient groups whose asthma control rating, percent predicted FEV1, and ACQ scores worsened.

Evidence of the reliability, validity, responsiveness, and usefulness of the ACT in clinical practice is captured in nearly 350 peer-reviewed journal articles. A few of these studies are mentioned here. The reliability of the ACT has been investigated using a variety of methods including internal consistency (Cronbach’s alpha), test-retest (intra-class correlation between two assessments), and principal components analysis. With very few exceptions, the reliability of the ACT has exceeded .80 (Grammatopoulou et al., 2011; Kosinski, Kite, Yang, Rosenzweig, & Williams, 2009; Nguyen, Chavannes, Le, Lan Thi Tuyet, & Price, 2012; Palmsten, Schatz, Chan, Johnson, & Chambers; Rodrigo et al., 2008; Schatz et al., 2007; Schuler, Faller, Wittmann, & Schultz, 2015; Uysal et al., 2013; Wallenstein et al., 2007; Zhou, Ding, Lin, & Yin, 2009).

Table 21.5 Performance of the ACT at Various Cut-Points in Screening for Uncontrolled Asthma

<table>
<thead>
<tr>
<th>Cut-Point Score</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive Predictive Value (%)</th>
<th>Negative Predictive Value (%)</th>
<th>% Correctly Classified</th>
<th>Area Under ROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 10</td>
<td>17.5</td>
<td>98.6</td>
<td>93.3</td>
<td>52.4</td>
<td>56.4</td>
<td>0.581</td>
</tr>
<tr>
<td>≤ 11</td>
<td>22.5</td>
<td>98.6</td>
<td>94.7</td>
<td>53.9</td>
<td>58.9</td>
<td>0.606</td>
</tr>
<tr>
<td>≤ 12</td>
<td>25.6</td>
<td>97.9</td>
<td>93.2</td>
<td>54.8</td>
<td>60.3</td>
<td>0.618</td>
</tr>
<tr>
<td>≤ 13</td>
<td>28.8</td>
<td>95.9</td>
<td>88.5</td>
<td>55.3</td>
<td>60.9</td>
<td>0.623</td>
</tr>
<tr>
<td>≤ 14</td>
<td>35.0</td>
<td>93.9</td>
<td>86.2</td>
<td>57.0</td>
<td>63.2</td>
<td>0.644</td>
</tr>
<tr>
<td>≤ 15</td>
<td>41.9</td>
<td>90.5</td>
<td>82.7</td>
<td>58.9</td>
<td>65.2</td>
<td>0.662</td>
</tr>
<tr>
<td>≤ 16</td>
<td>50.0</td>
<td>87.1</td>
<td>80.8</td>
<td>61.5</td>
<td>67.8</td>
<td>0.685</td>
</tr>
<tr>
<td>≤ 17</td>
<td>58.8</td>
<td>80.9</td>
<td>77.1</td>
<td>64.3</td>
<td>69.4</td>
<td>0.698</td>
</tr>
<tr>
<td>≤ 18</td>
<td>62.5</td>
<td>74.8</td>
<td>73.0</td>
<td>64.7</td>
<td>68.4</td>
<td>0.687</td>
</tr>
<tr>
<td>≤ 19</td>
<td>71.3</td>
<td>70.8</td>
<td>72.6</td>
<td>69.3</td>
<td>71.0</td>
<td>0.713</td>
</tr>
<tr>
<td>≤ 20</td>
<td>78.8</td>
<td>57.1</td>
<td>66.7</td>
<td>71.2</td>
<td>68.4</td>
<td>0.679</td>
</tr>
<tr>
<td>≤ 21</td>
<td>88.1</td>
<td>45.6</td>
<td>63.8</td>
<td>77.9</td>
<td>67.8</td>
<td>0.669</td>
</tr>
<tr>
<td>≤ 22</td>
<td>92.5</td>
<td>32.0</td>
<td>59.7</td>
<td>79.7</td>
<td>63.5</td>
<td>0.622</td>
</tr>
<tr>
<td>≤ 23</td>
<td>96.3</td>
<td>19.1</td>
<td>56.4</td>
<td>82.3</td>
<td>59.3</td>
<td>0.577</td>
</tr>
<tr>
<td>≤ 24</td>
<td>98.8</td>
<td>12.9</td>
<td>55.2</td>
<td>90.5</td>
<td>57.7</td>
<td>0.558</td>
</tr>
</tbody>
</table>

ROC = Receiver operating characteristic curve.

Categorized by asthma specialist as somewhat controlled, poorly controlled, or not controlled at all.
Table 21.6 Mean Changes in ACT Scores as a Function of Changes in Specialist Ratings of Control, Changes in FEV<sub>1</sub>, and ACQ Scores

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD) Change in ACT Score</th>
<th>F-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Specialist Assessment</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worse (n = 44)</td>
<td>−0.02 (1.7)</td>
<td></td>
</tr>
<tr>
<td>Same (n = 85)</td>
<td>0.73 (2.2)</td>
<td>24.2*</td>
</tr>
<tr>
<td>Better by 1 rating level (n = 80)</td>
<td>1.88 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Better by 2 rating levels (n = 37)</td>
<td>4.8 (3.9)</td>
<td></td>
</tr>
<tr>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 10% improvement (n = 212)</td>
<td>1.17 (2.8)</td>
<td>32.9*</td>
</tr>
<tr>
<td>≥ 10% improvement (n = 34)</td>
<td>4.32 (4.0)</td>
<td></td>
</tr>
<tr>
<td><strong>ACQ score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Better (n = 93)</td>
<td>3.94 (3.4)</td>
<td>71.0*</td>
</tr>
<tr>
<td>Same (n = 129)</td>
<td>0.39 (1.6)</td>
<td></td>
</tr>
<tr>
<td>Worse (n = 24)</td>
<td>−1.08 (2.2)</td>
<td></td>
</tr>
</tbody>
</table>

* p < .001.

In support of its validity, ACT scores have been shown to correlate significantly with percent predicted FEV<sub>1</sub> (Grammatopoulou et al., 2011; Papakosta et al., 2011; Wallenstein et al., 2007), dyspnea (Grammatopoulou et al., 2011), fractional exhaled nitric oxide (measure of airway inflammation; Habib, Alzoghaibi, Abba, & Hasan, 2014; Papakosta et al., 2011), asthma-specific quality of life and asthma symptom frequency (Schatz et al., 2007), healthcare utilization (Guilbert et al., 2011), and Global Initiative for Asthma (GINA) guidelines of asthma control (Miedinger et al., 2011). Similarly, known-groups validity tests demonstrating the discriminant validity of the ACT have shown mean ACT scores to differ significantly across patient groups differing in physician rating of asthma control (Grammatopoulou et al., 2011; Okelo et al., 2013; Uysal et al., 2013), GINA asthma control classification (Nguyen et al., 2012), and percent predicted FEV<sub>1</sub> values (Zhao, Lv, Liu, Cai, & Shao, 2010) in the hypothesized direction. The predictive validity of the ACT was demonstrated in several studies. ACT scores have been shown to predict future asthma exacerbations (Ko, Fanny et al., 2012; Sato et al., 2009), future asthma-related emergency room visits/hospitalizations and oral corticosteroid/beta agonist dispensing (Schatz et al., 2007), and GINA-defined asthma control (Hasegawa et al., 2013; Korn et al., 2011; Thomas et al., 2009).

Moreover, the ACT has been shown to be responsive to changes at the initiation of therapy (Al Moamary et al., 2012), changes in clinical status (Monteiro de Aguiar et al., 2014), long-term omalizumab treatment (Dal Negro et al., 2011), changes in course of asthma (Palmsten et al., 2008), changes in symptoms and disease control (Roxo, Jaqueline Petroni Faria et al., 2010; Schatz et al., 2006; Wallenstein et al., 2007), and changes in percent predicted FEV<sub>1</sub> values (Heffler et al., 2016; Schatz et al., 2006). Lastly, a number of studies have shown the ACT to be an effective screener for patients with asthma control problems (Grammatopoulou et al., 2011; Uysal et al., 2013; Wallenstein et al., 2007; Zhou et al., 2009).

While the standard ACT was developed and validated for individuals age 12 years and older, researchers from the University of Utah tested the ACT-CR with children age 2 through 11 years using a 1-week recall period that allows for more frequent assessments (Nkoy et al., 2013). The ACT-CR was shown to be reliable across multiple assessment periods throughout the study (Cronbach’s alpha = .82–.92). The validity of ACT-CR was investigated in several ways. In support of criterion-related validity, ACT-CR scores were shown to correlate significantly with scores of the Asthma Control Questionnaire (ACQ) at various time points throughout the study (r > .64). In support of discriminant validity, ACT-CR scores were shown to differentiate between patients who differed in
oral corticosteroid use and unscheduled acute asthma-related visits to the doctor in the hypothesized manner. Lastly, ACT-CR scores were shown to be responsive to changes in asthma control as determined by changes in ACQ scores.

General Strategy for Interpretation

Studies of validity are about the meaning of scores on a particular instrument and whether or not the scores have their intended interpretations. Typically, validity is most often determined by correlations between the instrument in question and conceptually related criterion measures. In contrast to such correlational data, the data presented in this section were designed to yield interpretation guidelines for differences or changes in ACT scores. Four types of data are presented. The first involves normative data on the ACT derived from a large national health and wellness survey. The second includes statistics on the accuracy of ACT scores in screening for patients with asthma control problems. The third involves the use of the MCID in ACT scores for determining the meaningfulness of differences or changes in individual patient and group-level mean scores. Lastly, interpretation guidelines for ACT scores are presented on the basis of results of analyses linking differences in ACT scores to external criteria, such as future asthma exacerbations and excessive use of short-acting beta agonists.

In the National Health and Wellness Survey (NHWS), the ACT was administered to respondents who self-reported that they were diagnosed with asthma by a doctor. For purposes of generating normative data ACT surveys from the years 2012, 2013, and 2014 were combined to ensure sufficient sample sizes for the presentation of norms by four major regions of the country (Northeast, Midwest, South, and West). The mean and standard deviation are 19.5 and 4.7, respectively, for ACT scores in the total combined sample. The median ACT score is slightly higher (21.0) than the mean, reflecting some skewness of the score distribution in the total combined sample, with more patients scoring above the mean. Mean ACT scores were very similar across the four regions: Northeast (19.8), Midwest (19.6), South (19.2), and West (19.6), and the median scores observed in each region ranged from 20 to 21.

ACT scores can be used to distinguish between those patients whose asthma is well controlled from those whose asthma is not well controlled. Based on empirical evidence from two clinical studies involving patients who were under the routine care of an asthma specialist and patients new to care of an asthma specialist, an ACT cut-point score of 19 (scores 19 and below) was determined to be optimal for screening patients with asthma control problems (Nathan et al., 2004; Schatz et al., 2006). A cut-point score of 19 was shown to have the highest area under the ROC curve (.73) and a sensitivity and specificity of 72.0% and 74.1%, respectively, among patients under the routine care of asthma specialists. In patients new to the care of an asthma specialist, a cut-point score of 19 was shown to have the highest area under the ROC curve (.71) and a sensitivity and specificity of 71.3% and 70.8%, respectively. In addition, a cut-point score has been determined to optimally distinguish between asthma patients with somewhat controlled asthma from patients with poorly controlled asthma (Schatz et al., 2006). An ACT cut-point score of 15 and below (versus ACT scores from 16 to 19) was determined to be optimal for differentiating between patients with poorly controlled asthma and patients with somewhat controlled asthma (ROC = .68; sensitivity = 61.1% and specificity = 73.9%; Schatz et al., 2006).

MCID

A multifaceted approach based on anchor-based and distribution-based methods (Yost & Eton, 2005) was used to determine the MCID of the ACT (Schatz et al., 2009). The results of this MCID study suggested that a difference of three points in mean ACT scores between two groups of patients is clinically meaningful. Conversely, a difference of mean scores between two groups of patients of less than two points would not be clinically significant, even if statistical significance is observed. This information would be particularly important for interpreting the results of intervention studies that use the ACT
as an outcome measure. The results of this MCID study also determined that a change in the ACT score of three points for an individual patient represents a clinically meaningful change in asthma control over time. Such information is useful for conducting responder analyses where comparisons are made between two groups on the percentage of patients meeting or exceeding a clinically meaningful change in asthma control, or for simply evaluating whether an individual patient’s asthma control is meaningfully improving or deteriorating over time.

**Risk Prediction**

Another approach to evaluate the meaningfulness of differences in ACT scores is to link ACT scores at a point in time (i.e., baseline) to an outcome or event at some point later in time (predictive validity). Using logistic regression methods, the odds of future asthma exacerbations and future high short-acting beta agonist (SABA) canister dispensing as a function of differences in baseline ACT scores from a score of 20 (i.e., threshold for controlled asthma) was investigated. Asthma exacerbations were defined as asthma hospitalization, emergency room visits, or oral corticosteroid dispensing during the six months following baseline administration of the ACT. A dichotomous variable was derived from the number of asthma exacerbations (one or more exacerbations vs. no exacerbations). A dichotomous variable was also derived for SABA canister dispensing over the six months after the administration and completion of the ACT at baseline (>6 SABA canisters vs. ≤6 SABA canisters).

Baseline ACT scores were found to be significantly related ($p < .001$) to future asthma exacerbations and more than six SABA dispensings. Results of the logistic regression analysis showed that lower ACT scores were associated with an increased probability of one or more future asthma exacerbations and greater than six SABA canisters dispensed over the six months after the completion of the ACT by patients. The odds ratios of one or more asthma exacerbations associated with a 1-point, 2-point, 3-point, 4-point, and 5-point lower score from the cut-point score of 20 (controlled asthma) were found to be 1.09, 1.21, 1.33, 1.46, and 1.60, respectively. Thus, scoring 3 points lower than 20 on the ACT means that an asthmatic patient is 33% more likely to have an asthma exacerbation in the next six months. The odds of having an asthma exacerbation in the next six months increases dramatically at scores below 15. The odds ratio for greater than six SABA canister dispensings associated with a 1-point, 2-point, 3-point, 4-point, and 5-point lower score from the cut-point score of 20 (controlled asthma) was found to be 1.21, 1.46, 1.76, 2.13, and 2.57, respectively. Thus, scoring three-points lower than 20 on the ACT means that an asthma patient is 76% more likely to receive greater than six SABA canisters in the next six months. The odds of receiving greater than six SABA canisters in the next six months also increases dramatically at scores below 15.

**Applications of the ACT**

The ACT is well suited to measure the degree of asthma control in a variety of settings ranging from primary care and other clinical practices to disease management and clinical trials. Its brevity and ease of administration make it easy to integrate into clinical practice and clinical studies. The ACT can be easily integrated into ordinary clinical practice in several applications. Point-in-time assessments can help to (1) identify patients whose asthma is poorly controlled, (2) identify patients who may require a change in therapy, and (3) facilitate communication between patients and healthcare providers. With repeated use with a given patient or group of patients, the ACT also can be used to monitor changes in asthma control over time, and determine the effects of treatment strategies for asthma.

The brevity of the ACT also makes it an ideal instrument for use in monitoring the asthma control of asthma populations. Managed care organizations or others can use the ACT with supplementary measures, such as biomarker, healthcare utilization, and treatment compliance, to assess and track the burden of asthma and benefits of treatment. Moreover, managed care organizations may wish to
use the ACT to identify those individuals most likely to have asthma control problems, indicating a need for more intensive care management and intervention. Others may wish to use ACT scores to monitor groups of asthma patients over time to determine the effects of particular treatment interventions or to measure the health risk (e.g., future asthma exacerbations, excessive SABA canister dispensing) of the asthma members in the plan. Lastly, the ACT could play an integral part in public awareness campaigns and educating asthma patients on the importance of controlling their asthma.

**Pain Impact Questionnaire (PIQ)**

Pain is defined by the American Pain Society (American Pain Society, 2004) and the International Association for the Study of Pain (Merskey & Bogduk, 1994) as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage” (p. 212). Pain is one of the most common reasons for patients to seek medical attention. Chronic pain (i.e., persistent or intermittent pain lasting at least three months) is one of the most prevalent conditions in the world and has a profound impact on the sufferer’s functioning and health-related quality of life (HRQOL; International Association for the Study of Pain, Subcommittee on Taxonomy, 1986). It is estimated that 50 million people in the United States alone are affected by chronic pain (Phillips, 2000).

The most frequently used pain measures are single items focusing on current pain intensity or pain within the past 24 hours. More elaborate traditional questionnaires, such as the McGill Pain Questionnaire (MPQ; Melzack, 1975), focus on describing the pain experience. Subsequent instruments, such as the Brief Pain Inventory (Cleeland & Ryan, 1994) included items describing the impact of pain on well-being and functioning. The combination of a pain severity and a pain impact item was also used in the short Bodily Pain (BP) scale of the SF-36 Health Survey, the most popular generic health outcome measure. Similarly, recent measures of chronic pain (Amtmann et al., 2010; Anatchkova, Saris-Baglama, Kosinski, & Bjorner, 2009; Becker et al., 2007) combine the assessment of pain intensity with the assessment of pain impact.

There is a need for a simple method to quantify pain severity (i.e., the intensity of pain) and pain impact (i.e., limitations in functional status, emotional well-being) for monitoring and improving pain management. The six-item Pain Impact Questionnaire (PIQ-6) was designed to meet this need for a single, valid measure of both pain severity and pain impact that is easy for clinicians to use, and for clinicians and patients to interpret. In developing this instrument, the goal was to construct an assessment that is brief, reliable, and valid, yet comprehensive in its coverage by reflecting different aspects of pain impact. The PIQ-6 has been cross-calibrated with the BP scale of the SF-36 to allow easy comparison of results. However, the PIQ-6 covers a broader range of pain impact and is more reliable and more comprehensive in its coverage of the impact of pain on well-being, role, and social function.

Two versions of the PIQ-6 are available: the standard PIQ-6, which uses a four-week recall, and the acute version, PIQ-6A, which uses a one-week recall. The PIQ-6 and PIQ-6A are scored so that high score means a high level of pain impact. An alternative scoring is possible, where high scores indicate little pain impact. These scores have the advantage of being directly comparable to the BP scores from the SF-36 questionnaire. The two scoring options are named PIQ-R and PIQ-R Acute. PIQ-R is the reverse score of PIQ-6 and PIQ-R Acute is the reverse score of PIQ-6A.

**Summary of Development**

The PIQ-6 was developed from a large chronic pain item bank (Anatchkova et al., 2009). This bank was developed to provide a comprehensive assessment of chronic pain and the impact of chronic
pain on patients’ functioning and well-being. Based on literature review, items were selected to cover four domains of pain: intensity, frequency, experience, and impact on function and well-being. Psychometric analyses showed that items on pain experience (e.g., stabbing pain, burning pain) did not fit a unidimensional pain model and these items were subsequently excluded, leaving a final bank of 45 items on pain intensity, frequency, and impact. The item bank fit a unidimensional model and was without differential item functioning (DIF) across sex, age, marital status, race, level of education, and employment status. These items were calibrated to an IRT model to enable administration by a computerized adaptive test, the Chronic Pain-CAT (Anatchkova et al., 2009). Computerized adaptive testing (CAT) selects the most informative items for each respondent and thus provides optimal measurement precision for a given number of items. However, experience from other tests, such as the HIT-6, have shown that it is possible to use IRT to select a fixed set of items that provides measurement precision almost as good as a CAT. The development of the PIQ-6 aimed to improve the range and precision of short-form measures of chronic pain (such as the SF-36 BP scale) while maintaining a low response burden.

The PIQ-6 was developed in four steps (Becker et al., 2007): (1) initial item selection and item development based on conceptual considerations and the psychometric properties of items in the chronic pain bank, (2) psychometric evaluation of the candidate items based on new data and selection of the final items, (3) development of procedures for administration, scoring, and interpretation, and (4) psychometric evaluation.

To maintain comparability with the widely used two-item BP scale from the SF-36, these two items were selected without modification. Based on conceptual considerations and on results from the chronic item bank development, six items were developed and revised to represent the subdomains and the measurement width of the chronic pain item bank while achieving a sufficiently consistent format for a short form. Conceptually, the aim was to include items covering a broad range of pain impact including impact on emotional well-being, leisure activities, and work functioning. The chosen items demonstrated high discriminative power (in IRT terms, large “slope parameters”), and covered a wide range of pain impact (in IRT terms widespread “threshold parameters”). Thus, the items represented issues that are typically encountered at various levels of pain severity (Bjorner et al., 2003a). The items use a four-week recall period and have five to six response choices.

The eight candidate items were tested in two new samples: a general population sample and a sample of persons with chronic pain. The results were used for item selection, development of scoring procedures, interpretation guidelines, and evaluation. The general population sample ($N = 829$) was recruited aiming at a nationally representative general population sample using an online panel in 2004 (Becker et al., 2007). For collecting the study sample 3,662 members were initially invited via email. Panelists were contacted until the desired sociodemographic representation was reached (with regards to gender, age, and income), yielding a response rate of 23%.

The chronic pain patients sample was collected in collaboration with the Kaiser Permanente Care Management Institute (KP-CMI) using the Internet (22%) or interactive voice recognition (IVR) technology (78%). A random sample of 10,700 KP members from six regions was invited through a mailed letter to participate. The sample consisted of chronic pain patients who self-reported that they had been told by a doctor or health professional that they had chronic pain that required regular use of medication. A total of 306 pain patients completed all the PIQ-6 items (response rate = 53.7%).

Psychometric analyses of the eight candidate items used confirmatory factor analysis for categorical data, IRT item calibration and test of item fit, and test of differential item functioning by sociodemographic groups. The analyses supported the essential unidimensionality of the pain impact items but suggested that two items could be deleted because of redundancy with other items in the test. Thus, the final version contains six items.
Initial norming of the PIQ-6 standard version used the 1998 SF-36v2 norm sample (Ware et al., 2007). The PIQ-6 scoring was developed to achieve a mean score of 50 and a standard deviation of 10 in this sample. Subsequently, norm data for both the standard and acute version were collected as part of the 2009 SF-36v2 norm study (Maruish, 2011). In this study, the original scoring resulted in mean scores close to 50 for both the standard and acute versions. Therefore, the original scoring procedure was retained.

Available Modes of Administration and Scoring
With just six items on a one-page form, the PIQ-6 can be integrated easily into many settings, including routine primary care practice, disease management programs, public awareness campaigns, clinical trials, and quality improvement initiatives. The questionnaire is self-administered and can be completed in the practice setting, at home, or in any other location. It can be administered as a pencil-and-paper questionnaire, in a mail-back form, during telephone or face-to-face interviews, or by almost any electronic application (IVR, PDA, desktop software, EMR). Online software versions are currently available to administer the form and provide immediate scoring and a comprehensive PIQ-6 feedback report to the patient/clinician or healthcare provider. The feedback report includes individualized feedback on the test taker’s performance.

When scored, the PIQ-6 yields an overall Pain Impact Score. It is also possible to interpret the response to single items—in particular the Pain Severity Rating item. The Pain Impact Score is normed so the US adult general population in 1998 had an average score of 50 and a standard deviation of 10. Higher scores on the PIQ-6 indicate a greater degree of pain severity and pain impact on a person’s life. Scores above 50 indicate that the person’s pain is more severe and has more impact on his/her life than the average of the US general population.

The original scoring of the PIQ-6 was based on IRT. However, because the IRT score can only be estimated using a computer, a simpler sum score approach is recommended. Each response choice is assigned a weight and the total score is calculated as the sum of these weights. If necessary, the calculation can be done by hand. The weights were developed so the sum score would match the IRT score as closely as possible (Becker et al., 2007). While the sum-scoring approach and the IRT approach will give the same result on average, the scores will not be identical.

Some clinicians may find it useful to know the Pain Severity Rating of the respondent in addition to the total Pain Impact Score. The Pain Severity Rating of the PIQ-6 is simply the response to Item 1 of the PIQ-6 standard form and does not require additional scoring. Higher ratings of this item indicate a higher severity of pain.

Basic Psychometric Information
The psychometric properties of the PIQ-6 were evaluated in the general population sample and the chronic pain sample described previously. The PIQ-6 showed good internal consistency reliability (coefficient alpha = .94). Convergent validity was supported by strong correlations with single-item pain severity rating scales (Visual Analog Scale [VAS], NRS; r = .81–.84, p < .001).

Known-groups validity was substantiated by significant mean score differences between chronic pain patients (mean = 64, SD = 7) and the US general population (mean = 50, SD = 10, p < .001), and across groups differing in medical conditions. PIQ-6 also shows strong associations with other outcomes indicative of chronic pain.

General Strategies for Interpreting Test Results
When using the sum-scoring approach, the score range for the PIQ-6 score ranges from 40 (no pain) to 78 for both the acute and standard version. Higher scores indicate greater pain impact. Table 21.7
presents general guidelines for interpreting the PIQ-6 score. The empirical data from which the guidelines have been developed are discussed further in the rest of this section. The table distinguishes four levels of pain impact based on the following cut-points:

- The cut-point of 50 is the mean of the original norm sample, with scores below 50 indicating no or minimal pain impact.
- The cut-point of 58 indicates the best split between the general population and patient with chronic pain based on the optimal combination of specificity (65%) and sensitivity (78%) as evaluated by their sum. Also, scores above 58 identified persons (specificity = 87%; sensitivity = 81%) with at least moderate pain severity as indicated by a VAS pain score larger than 40. PIQ-6 scores in the 50–57 range indicate some, but not severe pain impact.
- The cut-point of 64 identifies patients with severe pain (specificity = 92%; sensitivity = 73%) as indicated by a VAS pain score larger than 70 (Anderson, Syrjala, & Cleeland, 2001).
- PIQ-6 scores in the 58–63 range suggest substantial pain impact, scores 64 or higher suggest severe pain impact. If a patient’s impact score is higher than 58, a healthcare provider should discuss the patient’s condition with him/her and determine whether intervention is warranted (Becker et al., 2007).

The Pain Severity Rating ranges from 1 to 6, with higher ratings reflecting a higher pain severity. If a patient’s severity rating is equal to or higher than 4, a healthcare provider should discuss the patient’s condition with him/her and determine whether intervention is warranted. Data presented in the following sections provide more detailed interpretation guidelines for score levels and score differences.

### Norm-Based Interpretation

Normative data makes it possible to interpret PIQ-6 scores for an individual respondent or the average for a group of respondents, by comparing them with the distribution of scores for other individuals. Scores can be understood as departures from expected or typical scores; these expected or typical scores are called norms.

Norms for the PIQ-6 came from the 2009 SF-36v2 norm study (see Maruish, 2011). This study was based on a national probability sample of US noninstitutionalized adults aged 18 years and older. An oversample of respondents aged 65 years or older was also included. A weight variable was used to...
weight observations to the distribution on age and gender in the US general population. Data was collected through the Internet during two time periods in 2009. Participants were randomized to receive different versions of questionnaires, with a total of 1,954 respondents giving complete sets (no missing) of responses to the standard form and 1,960 respondents giving complete responses to the acute form.

Data for the US adult general population (i.e., a mixed group of adults of which some had chronic illness) and for a chronic pain sample revealed that the average PIQ-6 standard score was 50.7 with a standard deviation of 9.1. The average PIQ-6 acute score was 49.6 and the standard deviation was 9.6. For patients with chronic pain, the average PIQ-6 standard score was 63.3 and the standard deviation was 7.3. Severe pain impact, according to PIQ-6 scores, is experienced by approximately 10% of the general population and about 50% of chronic pain patients. Descriptive statistics of PIQ-6 scores stratified by gender for the general population sample indicated that on average, females score slightly higher (i.e., report more) impact than males do on both the standard and acute forms. PIQ-6 scores across seven different combined-gender age groups were found to increase from age group 18–24 to age group 55–64, but then seems to stabilize or decline slightly on both the standard and acute forms.

Many individuals suffering from pain report different disease conditions either as a cause of their pain or as a comorbid condition. Table 21.8 and Table 21.9 list descriptive statistics for different disease conditions self-reported by the general population. Patients within each disease group may suffer

<table>
<thead>
<tr>
<th>Chronic Condition</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>Comorbidities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Osteoarthritis</td>
<td>275</td>
<td>60.0</td>
<td>9.2</td>
<td>53</td>
<td>60</td>
<td>67</td>
<td>61% Hypertension, 47.9% Allergic rhinitis, 34.4% Gastroesophageal reflux disease</td>
</tr>
<tr>
<td>Rheumatoid arthritis</td>
<td>159</td>
<td>59.4</td>
<td>9.8</td>
<td>51</td>
<td>59</td>
<td>66</td>
<td>66.3% Hypertension, 41.9% Allergic rhinitis, 33.1% Gastroesophageal reflux disease</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>111</td>
<td>58.5</td>
<td>10.3</td>
<td>50</td>
<td>60</td>
<td>65</td>
<td>53.6% Hypertension, 34.8% Osteoarthritis, 33.9% Allergic rhinitis</td>
</tr>
<tr>
<td>Clinical depression</td>
<td>247</td>
<td>57.5</td>
<td>10.7</td>
<td>48</td>
<td>57</td>
<td>65</td>
<td>54.3% Allergic rhinitis, 41.3% Hypertension, 37% Obesity</td>
</tr>
<tr>
<td>Sleep apnea</td>
<td>205</td>
<td>56.8</td>
<td>10.8</td>
<td>48</td>
<td>57</td>
<td>64</td>
<td>61% Hypertension, 49.5% Allergic rhinitis, 44.3% Obesity</td>
</tr>
<tr>
<td>Gastroesophageal reflux disease</td>
<td>310</td>
<td>56.5</td>
<td>10.1</td>
<td>48</td>
<td>56</td>
<td>63</td>
<td>57.1% Hypertension, 52.1% Allergic rhinitis, 30.8% Osteoarthritis</td>
</tr>
<tr>
<td>Heart disease</td>
<td>260</td>
<td>56.1</td>
<td>9.9</td>
<td>48</td>
<td>56</td>
<td>63</td>
<td>65.7% Hypertension, 43.8% Allergic rhinitis, 29.8% Gastroesophageal reflux disease</td>
</tr>
<tr>
<td>Diabetes</td>
<td>267</td>
<td>56.0</td>
<td>10.5</td>
<td>47</td>
<td>55</td>
<td>64</td>
<td>75.1% Hypertension, 44.7% Allergic rhinitis, 39.9% Obesity</td>
</tr>
</tbody>
</table>
Table 21.9  PIQ-6 (Acute Version) Scores Among Persons With Chronic Conditions

<table>
<thead>
<tr>
<th>Chronic Condition</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>Comorbidities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatoid arthritis</td>
<td>145</td>
<td>59.3</td>
<td>9.8</td>
<td>53</td>
<td>59</td>
<td>68</td>
<td>63.8% Hypertension 53% Allergic rhinitis 34.9% Gastroesophageal reflux disease</td>
</tr>
<tr>
<td>Osteoarthritis</td>
<td>286</td>
<td>59.2</td>
<td>9.5</td>
<td>51</td>
<td>60</td>
<td>67</td>
<td>62.3% Hypertension 48.1% Allergic rhinitis 34.9% Gastroesophageal reflux disease</td>
</tr>
<tr>
<td>Osteoporosis</td>
<td>151</td>
<td>58.2</td>
<td>9.7</td>
<td>51</td>
<td>60</td>
<td>67</td>
<td>58.9% Hypertension 41.8% Osteoarthritis 36.1% Allergic rhinitis 55.1% Hypertension 44.9% Obesity</td>
</tr>
<tr>
<td>Clinical depression</td>
<td>259</td>
<td>56.9</td>
<td>10.7</td>
<td>48</td>
<td>56</td>
<td>67</td>
<td>63.7% Hypertension 51.7% Allergic rhinitis 46.3% Obesity</td>
</tr>
<tr>
<td>Sleep apnea</td>
<td>196</td>
<td>56.7</td>
<td>11.3</td>
<td>45</td>
<td>58</td>
<td>67</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
from comorbid conditions (the most prevalent comorbid conditions are reported in the last column of each table). For both the standard and the acute versions, respondents self-reporting osteoarthritis, rheumatoid arthritis, or osteoporosis have the highest mean PIQ-6 scores (i.e., report the most pain impact across disease groups; see Column 3). The last column reports the most prevalent comorbid conditions within each disease group outlined in the first column of the table. For example, of 1,007 patients reporting osteoarthritis or degenerative arthritis, 68.2% also suffer from back pain, 46.2% report high blood pressure, and 28.1% report depression as comorbid conditions.

**MCID and Responder Criteria**

When interpreting differences, it is important to distinguish between differences that are of at least some importance to the patient and differences that are trivial. The term *minimal clinical important*...
score difference (MCID) usually refers to a mean difference between groups. The MCID is important, for example, when planning a clinical trial in order to determine the smallest score difference a study should have the statistical power to be able to detect. The term responder criterion refers to the minimal score change that would indicate a significant improvement or deterioration for an individual patient. A responder criterion is often larger than the MCID for the same measure, because the score for an individual patient is less precisely assessed than the mean score for a group of patients. In line with the MCID for the SF-36 BP scale (Bjorner, Kosinski, & Ware, 2003c), an MCID of three points is recommended for the PIQ-6. The responder criterion can be assessed in different ways (Maruish, 2011). A conservative estimate is provided by the reliable change index (RCI; Jacobson & Truax, 1991). Based on the estimated reliability for the PIQ-6 (.94) and the general population standard deviation (9.1), an RCI-based responder criterion is calculated to be 6.2 points.

**Interpretation Based on Association With Other Measures**

Criterion-based tests of validity are based on analyses of relationships between the measures in question and other variables referred to as “external criteria,” measured either concurrently or after some period of time. Criteria reported in this section are a Visual Analog Scale (VAS) for measuring pain severity, the patient’s self-reported work or leisure days missed due to chronic pain, and the SF-36v2 BP scale.

Table 21.10 presents descriptive statistics of PIQ-6 scores of a combined sample of chronic pain patients and the PIQ-6 development sample, stratified by pain severity as assessed by one of the most frequently used pain severity measures, a single 100mm VAS ranging from 0 (no pain) to 100 (pain as bad as it could be). The combined sample was stratified by pain severity level based on VAS scores, using established cut-points (Anderson et al., 2001). As expected, the mean PIQ-6 score was lowest among patients reporting no pain (patient with a VAS score = 0 had a mean PIQ-6 score = 41.9) and highest for patients rated as suffering from severe pain (patients with a VAS score > 71 had a mean PIQ-6 score = 66.5). The mean PIQ-6 score among patients suffering from mild pain (VAS = 1 to 40) was 51.7; for patients suffering from moderate pain (VAS = 41 to 70) the mean was 60.6.

Table 21.11 presents the percentage of employed chronic pain patients who reported to have missed at least one or more work days, and the percentage of chronic pain patients who were unemployed and reported to have missed at least one or more leisure days within the past four weeks. As expected, the percentage of chronic pain patients who missed at least one day of work or leisure activities was higher with each increasing level of PIQ-6 scores. Thus, if a respondent is employed and has a PIQ-6 score of 62 (“Substantial” pain impact), the likelihood that he reports at least one work day missed within the last four weeks is about 34%. Or, if a respondent is currently unemployed and

<table>
<thead>
<tr>
<th>VAS Score/Pain Severity</th>
<th>PIQ-6 Standard and Acute</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1–40</td>
</tr>
<tr>
<td>No Pain</td>
<td>Mild</td>
</tr>
<tr>
<td>Mean</td>
<td>41.9</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.2</td>
</tr>
<tr>
<td>25th percentile</td>
<td>40</td>
</tr>
<tr>
<td>50th percentile</td>
<td>40</td>
</tr>
<tr>
<td>75th percentile</td>
<td>40</td>
</tr>
<tr>
<td>Total (N)</td>
<td>47</td>
</tr>
</tbody>
</table>
A patient (illustrated by the data) has a PIQ-6 score of 62 ("Substantial" pain impact), the likelihood that he missed at least one day of leisure activities within the last four weeks is about 46%.

Because the two items constituting the SF-36 BP scale are included in the PIQ-6, the BP scale can always be calculated from PIQ-6 item responses. However, researchers may want to take advantage of the broader content coverage and measurement precision of the PIQ-6, but may also be interested in comparing their PIQ-6 scores with results from earlier studies that have used the SF-36 BP scale. For this reason, analyses were performed to calculate the expected BP score for each level of the Pain Impact Score (see Bjorner et al., 2003c for description of the approach). Figure 21.4 shows the expected BP norm-based score for each level of the PIQ-6 score. The graph shows that the relation between the PIQ-6 and the BP score is approximately linear in the Pain Impact Score range of 50–75 (approximately equal to a BP score range of 25–52). In this range, a 1-point increase in PIQ-6 score is approximately equal to a 1-point decrease in BP score.

### SUMMARY

The inclusion of patient-reported measures of disease impact in the routine clinical care of chronically ill patients has the potential to add valuable information about the disease and its treatment. It
also promotes effective self-management in which patients become more active participants in their own care. These tools yield accurate and time-sensitive information about patients’ symptoms as well as their impact on functioning and well-being, often expanding the nature of patient–clinician communication. This chapter has discussed the concepts and measurement ideas underlying measures of disease impact. It presented four examples of disease impact measures: the HIT-6, the ACT and AIS, and the PIQ-6. All of these measures are brief and have strong measurement properties, and the results are readily interpretable by both patients and providers.

Technology solutions for the collection, scoring, and interpretation of patient-reported data and integration into electronic health record systems are emerging. However, patients, providers, and healthcare systems—including primary care practices—still face logistical challenges for full integration of these data into the clinic workflow.

A disease impact measure intended for use in a clinical setting needs to be easy to administer, score, and interpret. In addition to being reliable and valid, the tool needs to contribute new information to the clinical encounter and enhance the provider’s ability to make fully informed treatment recommendations. With routine use, disease impact assessments will yield information that enriches the patient–provider interaction, optimizing treatment decisions and improving patient outcomes.

ACKNOWLEDGMENTS

The authors would like to thank Jillian Silva-Jones for project management and editorial support. Also, we thank Scott Cohoon for document formatting, proofreading, and quality control. The development of this chapter was supported by the Patient Insights division of Optum.

COPYRIGHT INFORMATION

The Headache Impact Test (HIT-6), the Asthma Control Test (ACT), and the Pain Impact Questionnaire (PIQ-6) are trademarks of Optum. Permission to reproduce and to use these tools and the associated trademark(s) documented in this manual is routinely granted royalty-free to individuals and organizations that collect their own data for purposes of scholarly research. Permissions for both scholarly and commercial use can be obtained by completing a Survey License Form. All other uses, commercial and noncommercial, may require payment of an annual license fee. Completion of the Survey License Form will result in the quotation of any user fees and, upon user request and approval by Optum, the issuance of a license and invoice. Any organization or individual wishing to reproduce the survey documented herein and/or any associated intellectual property (e.g., the trademarks, scoring algorithms, interpretation guidelines, and/or normative data documented in this manual) for any purpose must register or obtain a license from Optum. For information about registering or obtaining a license, go to https://campaign.optum.com/optum-outcomes/survey-request-form.html

REFERENCES


Bjorner, J. B., Kosinski, M., & Ware, John E. Jr. (2003c). Using item response theory to calibrate the Headache Impact Test (HIT) to the metric of traditional headache scales. *Quality of Life research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 12*(8), 981–1002.


Assessment of Disease Impact

Stewart, W.
Thomas, M., Kay, S., Pike, J., Williams, A., Rosenzweig, J.
Schatz, M., Kosinski, M., Yarlas, A.
Rodrigo, G.
Phillips, D.
Peters, S.
Ware, J. E., Jr., Bjorner, J. B., & Kosinski, M. (2000). Practical implications of item response theory and computerized adaptive testing: A brief summary of ongoing studies of widely used headache impact scales. Medical Care, 38(9 Suppl.), I73–82.


It is now well established that psychiatric disorders represent a highly prevalent phenomenon in society. Results from the National Institute of Mental Health (NIMH) Epidemiologic Catchment Area (ECA) study, involving nearly 20,000 community individuals, comprehensively emphasized that fact (Myers et al., 1984; Regier et al., 1988; Robins et al., 1981). This means that techniques to assess and quantify the psychological distress that is an inherent aspect of these states have assumed greater importance in helping to accomplish the screening, diagnosis, and treatment evaluation of these conditions.

The overall rate for affective disorders in these studies was 5.1%, while that for anxiety disorders was 7.3% (Regier et al., 1988). Six-month prevalence estimates for affective disorders ranged from 4.6% to 6.5% across the five ECA sites (Meyers et al., 1984). Comparable 6-month estimates for anxiety disorders demonstrated rates for panic disorder ranging from 0.6% to 1.0%, while the prevalence of agoraphobia varied from 2.5% to 5.8% across the various centers (Weissman & Merikangas, 1986). These rates do not include states of psychological distress that, for a variety of reasons, did not qualify for formal diagnoses in the DSM system.

In medical populations, prevalence rates of psychiatric disorder/distress were even more elevated. This is particularly true of anxiety and depressive disorders, which account by far for the majority of psychiatric diagnoses assigned to medical patients (Barrett, Oxman, & Gerber, 1988; Derogatis et al., 1983; Von Korff, Dworkin, & Krueger, 1988). In reviews of psychiatric prevalence in medical populations, Barrett et al. observed overall prevalence rates of 25%–30% while Derogatis and Wise reported prevalence estimates that varied between 22% and 33%.

Focusing on anxiety states, Kedward and Cooper (1966) observed a prevalence rate of 27% in an earlier study of a London general practice, while Schulberg and his colleagues (1985) reported a combined rate of 8.5% for phobic and panic disorders among American primary care patients. Wise and Taylor (1990) concluded that 5%–20% of medical inpatients are distressed by the symptoms of anxiety, while 6% receive formal anxiety diagnoses.

These data support the contention that psychiatric disorders are a persistent and demonstrable set of problems that affect substantial numbers of the population, and reflect vulnerabilities that are particularly intensified in the context of the additional stresses of a comorbid medical illness. This is because as noted by Derogatis and Wise (1989), “symptoms of anxiety and depression can be independent concomitants of, reactions to, substitutes for, or integral aspects of, an entire spectrum of medical disorders” (p. 81).

Particularly in serious chronic medical conditions, where the patient’s coping capacity and psychological integrity are severely tested, symptoms of anxiety and depression sometimes become indistinguishable features of the principal medical illness. These psychological distress states, although they do not formally qualify as diagnostic entities, nonetheless are associated with a substantial degree of discomfort and significantly reduced quality of life. A study lending persuasive support to this observation was reported by Johnson, Weissman, and Klerman (1992), who found that levels of service burden and health impairment were as high or higher for individuals with subdiagnostic
levels of depressive symptoms as they were for individuals receiving a diagnosis of major depression or dysthymic disorder. When such states coexist with a medical disorder, they frequently serve to further undermine the patient’s sense of well-being and can in some instances undermine adherance to medical treatment regimens.

The detection and characterization of psychiatric conditions and distress states comorbid with primary medical disorders is extremely important, because there is persuasive evidence that unidentified, and therefore untreated, psychiatric problems consistently have nonsalutary effects on the course and outcomes of medical conditions. Derogatis and DellaPietra (1994) reviewed the array of problems associated with the identification of anxiety and depressive disorders in primary care settings, and numerous authors (e.g., Bridges & Goldgr, 1984; Katon et al., 1990) have emphasized that principal among these is the high rate of somatic presentations among these patients. In addition to increases in morbidity and mortality associated with untreated comorbid conditions (Kamerow, Pincus & MacDonald, 1986; Regier et al., 1988), increased levels of healthcare utilization have also been well documented (Katon et al., 1990). In addition, high levels of psychological distress among medical patients have been associated with increased medical costs (Allison et al., 1995) and increased length of hospital stay (Saravay, Pollack, Steinberg, Weinschel, & Habert, 1996). Because primary care physicians are functioning very much in the role of “gatekeepers” to effective mental health treatment, identification and measurement of these comorbid conditions assumes increased consequence.

THE SCL-90-R AND RELATED SCALES

SCL-90-R

The SCL-90-R (Derogatis, 1977, 1994) is a 90-item self-report symptom inventory. It represents an evolution directly from the Hopkins Symptom Checklist (HSCL; Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974a, 1974b), which has roots in a number of earlier tests, such as the Cornell Medical Index (Wider, 1948).

A prototype version of the SCL-90-R was first described in 1973 (Derogatis, Lipman, & Covi, 1973), and the final version of the instrument was completed 2 years later (Derogatis, 1975). The inventory measures psychological symptoms and distress in terms of nine primary symptom dimensions and three global indices. The primary symptom dimensions of the SCL-90-R are Somatization (SOM), Obsessive-Compulsive (O-C), Interpersonal Sensitivity (INT), Depression (DEP), Anxiety (ANX), Hostility (HOS), Phobic Anxiety (PHOB), Paranoid Ideation (PAR), and Psychoticism (PSY). The global measures were designed to provide summary measures of overall distress status, each from a slightly different perspective. These measures are termed the Global Severity Index (GSI), Positive Symptom Distress Index (PSDI), and Positive Symptom Total (PST).

The SCL-90-R and its companion instruments in the series were developed to be utilized with an extensive range of respondents. The inventory may be validly employed with community respondents, a broad spectrum of medical outpatients and inpatients, and the large majority of patients with psychiatric disorders. The SCL-90-R is currently available in over 30 languages, including Dutch, English, French, German, Italian, Russian, Spanish, and many others. Online scoring, administration, and interpretation programs are also available for the SCL-90-R.

Brief Symptom Inventory

The Brief Symptom Inventory (BSI; Derogatis, 1993; Derogatis & Melisaratos, 1983; Derogatis & Spencer, 1982) comprises 53 items and represents the brief form of the SCL-90-R. It was also completed
in 1975, and reflects psychological distress/disorder in terms of the same nine symptom dimensions and three global indices as its longer counterpart. The BSI was designed specifically for measurement situations in which time constraints will not allow at least 15 minutes, the time typically required to complete the SCL-90-R. Scores on the SCL-90-R and the BSI are highly correlated, however, and very often the brief version of the test is preferred, even in the absence of time constraints. As with the SCL-90-R, the three global indices, nine principal symptom dimensions, and 53 individual items reflect the three basic levels of clinical interpretation of the test.

**BSI-18**

More recently (Derogatis, 2000), a new instrument, the BSI-18, was added to this integrated series of measures. As its name implies, the BSI-18 is a brief, 18-item version of the BSI. The instrument is designed to serve primarily as a screening measure for psychological distress and disorder in medical and community populations, and secondarily as an outcomes measure. Unlike its predecessors, however, the BSI-18 does not measure the standard nine symptom dimensions characteristic of the series, but rather measures only three. Somatization (SOM), Depression (DEP), and Anxiety (ANX) are the three primary symptom dimensions of the BSI-18.

Reduction in the number of BSI-18 symptom dimensions was in part a decision based on time constraints inherent in the screening paradigm. The selection of these particular three dimensions, however, was a reflection of a consensus that close to 80% of psychiatric disorders occurring in community and medical populations are anxiety or depressive disorders (Derogatis & Wise, 1989; Katon & Sullivan, 1990). In addition, a large proportion of these cases present with a profusion of somatic symptoms and manifestations that can act to confound the diagnostic process (Katon, 1987; Kirmayer, Robbins, Dworkin, & Yaffe, 1993). By focusing assessment on these three prominent facets of psychological dis-integration, it was felt that the BSI-18 would be highly sensitive to the most relevant and common indicators of psychological distress and disorder.

**Companion Scales**

The SCL-90-R represents a well-validated test instrument that is broadly applied and highly utilized in its own right; however, it also serves as the centerpiece of a series of matched, multimodality tests. A major advantage of multimodality approaches resides in the fact that they enable assessment of clinical status through both self-report and expert clinical judgment, using comparable measuring instruments in each modality. This goal was achieved in the current series through the development of several matched “companion” clinical rating scales to the SCL-90-R/BSI.

The Derogatis Psychiatric Rating Scale (DPRS) is a multidimensional clinical rating scale designed to be the clinician’s version of the SCL-90-R/BSI. The first nine dimensions of the DPRS match the nine symptom constructs of the self-report instruments. Eight additional dimensions, potentially important to clinical interpretation, but not easily amenable to reliable self-report, also make up the scale. A brief form of the DPRS (termed the Brief Derogatis Psychiatric Rating Scale [BDPRS]) is also available; it consists of only the nine matching SCL-90-R/BSI symptom constructs.

The SCL-90 Analogue Scale is a second companion scale to the SCL-90-R/BSI. It is designed for health professionals (e.g., physicians, nurses, social workers, lay interviewers) who have not received extensive training in psychopathology and psychiatric nosology. It is a graphic or analogue scale that represents the nine primary symptom dimensions of the SCL-90-R along 100 mm lines, extending from “not at all” at the minimum distress point to “extremely” at the maximum. Any of the three companion clinical observer’s scales may be used in conjunction with either the SCL-90-R or the BSI.
In judging the clinical meaning of psychological test scores, just as in evaluating the clinical implications of a blood chemistry panel, reliable, meaningful standards must be available for contrast and comparison. No matter how well designed and validated a test is, without such standards (referred to as norms in psychological measurement) the test will have little utility in evaluating the individual patient’s clinical status. Norms provide an interpretive point of reference; they define the patient’s status on the characteristic(s) being measured relative to a representative sample of like individuals. If norms are well constructed, they help the clinician to make a meaningful and informed judgment of the patient’s status.

Details of norm construction are somewhat esoteric and arcane; however, there are several key properties of quality norms to look for. First, the norm must be based on a representative sample of individuals from the population to which comparisons are intended. If the comparison to be made is with “normal,” then a representative cohort of community individuals free of discernable psychiatric disorder should constitute the normative group. If, on the other hand, a clinician wishes to compare the psychological distress profile of a Stage II breast cancer patient to like patients, then the normative cohort should include respondents diagnosed with a similar condition. Second, the norms should be actuarial in nature, enabling the clinician to attach some sense of clinical meaning to the patient’s scores, usually in terms of percentile or probability equivalents. Third, the level of generalizability of the norm must be clearly stated. It is well established that distributions of many important biological and psychological characteristics are altered considerably as a function of parameters such as gender, age, and health status. Particularly in the area of psychological distress, normative inferences should be based on gender-specific norms, with the added realization that the presence of an active medical condition almost invariably inflates distress levels significantly.

Currently there are four formal norms for the SCL-90-R and BSI: (1) psychiatric outpatients, (2) community nonpatients, (3) psychiatric inpatients, and (4) community adolescents (Derogatis, 1994). All norms for the SCL-90-R/BSI are actuarial in nature and gender-keyed. Gender-keying represents an important normative refinement when attributes involving emotional expression or psychological distress are being assessed because of well-established gender differences in reporting emotional distress.

The psychiatric outpatient norms for the SCL-90-R are based on 1,002 heterogeneous outpatients who presented for treatment at the outpatient psychiatry departments of four major teaching hospitals located in the East and Midwest. The same sample was utilized in creating the Outpatient norm for the BSI. The community nonpatient norms were established on a cohort of 973 individuals who represent a stratified random sample from a diversely populated county in a major Eastern state. Again, the Community norm for the BSI was also developed based on this sample. The psychiatric inpatient norms for both the SCL-90-R and BSI are based on a sample of 423 individuals who constituted a heterogeneous group of patients from the psychiatric inpatient services of three major Eastern hospitals. The adolescent community norms for the SCL-90-R are based upon 806 adolescents who were enrolled in two geographically distinct Midwestern high schools. Adolescent community norms for the BSI were developed from a sample of 2,408 adolescents, ranging in age from 13 to 19, who attended six different schools in two distinct states. Detailed demography for all published norms for the SCL-90-R may be found in the SCL-90-R: Administration, Scoring & Procedures Manual (Derogatis, 1977, 1994). Similar data for the BSI is provided in Brief Symptom Inventory (BSI): Administration, Scoring & Procedures Manual (Derogatis, 1992, 1993).
Norms for the BSI-18 were derived from populations distinct from those that served as the normative base for the SCL-90-R and the BSI. Community norms for the BSI-18 were developed from a community sample of 1,136 individuals who were employees across all levels of a large national corporation, and Medical (oncology) norms were derived from a sample of 1,543 patients who presented at a large Eastern urban cancer center, with a broad range of diagnoses.

RELIABILITY AND VALIDITY

One must consider two important psychometric constructs when evaluating the suitability, appropriateness, and/or integrity of any psychological test, rating scale, or related measure or procedure for a particular purpose. One is its reliability and the other is its validity.

Reliability

Reliability essentially pertains to the consistency or replicability with which an instrument measures the characteristic(s) under observation. It is the converse of measurement error, and represents the proportion of variation in any measurement that is due to systematic variation of the attribute under study (e.g., depression, hardiness, impulsivity) as opposed to variance due to random or systematic error. Two formal types of reliability estimates are available for the symptom dimensions of the SCL-90-R: (1) internal consistency and (2) test-retest. The former serves to reflect the homogeneity of the item sets developed to represent each symptom construct; test-retest reliability is much more of a measure of temporal stability, or score consistency across time.

SCL-90-R Reliability

Internal consistency coefficients for the nine dimensions of the SCL-90-R were calculated from the data of 209 “symptomatic volunteers” (Derogatis, Rickels, & Rock, 1976) in the form of coefficients alpha ($\alpha$). Coefficient alpha treats within-form correlations among the items as analogous to correlations between alternate forms, and makes the assumption that the average correlation among actual items is equivalent to the correlation among items in the hypothetical alternate form (Nunnally, 1970). Coefficients in this assessment were quite satisfactory, ranging from a low of .77 for PSY to a high of .90 for DEP. Internal consistency coefficients for the SCL-90-R were also developed more recently by Horowitz, Rosenberg, Baer, Ureno, and Villasenor (1988) based upon 103 outpatients presenting for psychotherapy. Coefficients alpha in that study ranged from a low of .84 for I-S to a high of .90 for DEP (see Table 22.1).

The test-retest coefficients presented in Table 22.1 were developed from a sample of 94 heterogeneous psychiatric outpatients who presented for evaluation and treatment at the psychiatric outpatient department of a major Eastern teaching hospital. One week elapsed between testings, and as is clear from the sizes of the coefficients, the SCL-90-R possesses very acceptable test-retest reliability. Coefficients ranged from a low of .78 on HOS to a high of .90 on the PHOB dimension. All other stability coefficients fell in the mid .80s. In addition to these estimates of temporal stability, Horowitz et al. (1988) also evaluated the test-retest reliability of the SCL-90-R in their sample of 103 psychiatric outpatients. Even across 10 weeks, coefficients were well within the acceptable range, with the coefficient for the GSI reported as .84, and subscale coefficients ranging from a low of .70 for O-C to a high of .83 for PAR.

BSI Reliability

Internal consistency reliability coefficients for the BSI were established based on a sample of 719 psychiatric outpatients, using Cronbach’s coefficient alpha. The alpha coefficients for the nine
dimensions of the BSI ranged from a low of .71 on the PSY dimension to a high of .85 on DEP. Independent investigators have reported internal consistency coefficients in a comparable range for the BSI (Aroian & Patsdaughter, 1989; Croog et al., 1986).

As indicated previously, test-retest reliability is an indicator of the consistency of measurement across time. If untreated, psychological distress or psychopathology tends to endure for moderate to substantial periods of time; therefore, a test designed to measure symptomatic distress should register high test-retest coefficients over a span of 2 weeks. To address this issue, a sample of 60 nonpatient individuals were tested across a 2-week interval. Coefficients ranged from a low of .68 for SOM to a high of .91 for PHOB. The GSI also revealed an excellent stability coefficient of .90, providing assurance that the BSI represents consistent measurement across time. Internal consistency and test-retest reliability coefficients for the nine primary symptom dimensions and three global indices of the BSI are represented in Table 22.2.

BSI-18 Reliability

Because it is a more recent instrument, a comparable body of reliability studies including test-retest evaluations has not yet been accomplished. However, internal consistency coefficients (i.e., coefficients $\alpha$) have been calculated (Derogatis, 2000) based on the 1,134 individuals who constitute the community normative sample and were found to be very acceptable: SOM = .74, DEP = .84, ANX = .79, GSI (Total) = .89. The coefficient for the GSI was also confirmed by Silver, Holman, McIntosh, Poulin, and Gil-Rivas (2002), who reported an $\alpha$ coefficient of .93 in a large study of psychological reactions to the 9/11 attacks.

In an interesting recent study, Meijer, Vries, van Brugen, and Demence (2011) completed Mokken analyses and parametric item response theory analyses on three large samples of outpatients ($N = 487$), students ($N = 286$), and prisoners ($N = 207$), with the BSI-18 evaluating the factor structure. They report that the BSI-18 formed a strong Mokken scale for outpatients and prisoners, with only the ANX and DEP scales forming a medium Mokken scale for the student sample. They concluded that their analysis argues for a unidimensional structure for the instrument, although at the second order the obvious unidimensional construct being measured is psychological distress.
Table 22.2 Internal Consistency and Test-Retest Reliability Coefficients for the Nine Primary Symptom Dimensions and Three Global Indices of the BSI

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Number of Items</th>
<th>Internal Consistency (α) (N = 719)</th>
<th>Test-Retest (Âtt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Somatization (SOM)</td>
<td>7</td>
<td>.80</td>
<td>.68</td>
</tr>
<tr>
<td>II. Obsessive-Compulsive (O-C)</td>
<td>6</td>
<td>.83</td>
<td>.85</td>
</tr>
<tr>
<td>III. Interpersonal Sensitivity (I-S)</td>
<td>4</td>
<td>.74</td>
<td>.85</td>
</tr>
<tr>
<td>IV. Depression (DEP)</td>
<td>6</td>
<td>.85</td>
<td>.84</td>
</tr>
<tr>
<td>V. Anxiety (ANX)</td>
<td>6</td>
<td>.81</td>
<td>.79</td>
</tr>
<tr>
<td>VI. Hostility (HOS)</td>
<td>5</td>
<td>.78</td>
<td>.81</td>
</tr>
<tr>
<td>VII. Phobic Anxiety (PHOB)</td>
<td>5</td>
<td>.77</td>
<td>.91</td>
</tr>
<tr>
<td>VIII. Paranoid Ideation (PAR)</td>
<td>5</td>
<td>.77</td>
<td>.79</td>
</tr>
<tr>
<td>IX. Psychoticism (PSY)</td>
<td>5</td>
<td>.71</td>
<td>.78</td>
</tr>
</tbody>
</table>

Global Indices
- Global Severity Index (GSI) — .90
- Positive Symptom Distress Index (PSDI) — .87
- Positive Symptom Total (PST) — .80

Alternate-forms reliability is a third form of reliability that is typically illustrated in correlation between score distributions from two different forms of a test. While there is no pure alternate form of the BSI, the SCL-90-R is a test that measures identical symptom constructs. To evaluate the level of agreement between the two test forms, correlations were calculated based upon a sample of 565 psychiatric outpatients. Coefficients across the nine primary symptom dimensions are given in Table 22.3. The data demonstrate very high correlations between the BSI and the SCL-90-R on all nine symptom dimensions. At least for psychiatric populations, the two tests show high agreement on all nine of the symptom constructs.

Table 22.3 Correlations Between Symptom Dimensions of the SCL-90-R and the BSI Based on 565 Psychiatric Outpatients

<table>
<thead>
<tr>
<th>SOM</th>
<th>O-C</th>
<th>I-S</th>
<th>DEP</th>
<th>ANX</th>
<th>HOS</th>
<th>PHOB</th>
<th>PAR</th>
<th>PSY</th>
</tr>
</thead>
<tbody>
<tr>
<td>.96</td>
<td>.96</td>
<td>.94</td>
<td>.95</td>
<td>.95</td>
<td>.99</td>
<td>.97</td>
<td>.98</td>
<td>.92</td>
</tr>
</tbody>
</table>

Validity

Two major issues should be understood concerning the validation of psychological test instruments: the first issue involves the specificity of validity, and the second has to do with the programmatic nature of the validation process. The former refers to the fact that in order for the question “Is this test valid?” to have any scientific meaning, the conditional modifier “For what purpose?” must be introduced. Psychological tests are not valid in general; like all other scientific measuring instruments, they are valid for certain specific measurement purposes and invalid for most others.

The second issue reflects the fact that psychometric theorists have increasingly stressed construct validity as the principal criterion for the validation of psychological tests and the assignment of meaning to them (Messick, 1975, 1981). The validation process, when accomplished successfully,
involves an extensive program of experiments and analyses that are highly analogous to the steps necessary to prove a scientific theory. Data from predictive, content, convergent-discriminant, and other types of validation experiments serve to contribute to the ultimate validation of the tests. The process of establishing the validity of a test is represented by an integrated series of studies that function to extend and redefine the limits of generalizability of the test as a definition of the construct (e.g., depression) being measured.

SCL-90-R Validity
Convergent-discriminant validation is a basic form of validity that is designed to demonstrate that the measure of interest correlates substantially with distinct measures of the same construct, and shows little or no correlation with measures of dissimilar constructs. Derogatis, Rickels, and Rock (1976) demonstrated convergent-discriminant validity for the SCL-90-R in a study contrasting its dimensions with those of the Minnesota Multiphasic Personality Inventory (MMPI). In addition to the standard MMPI clinical scales, the MMPI was also scored for the Wiggins (1969) content scales, and Tryon’s (1966) cluster scales. Results illustrated that SCL-90-R dimensions had their highest correlations with analogous MMPI constructs, in every case except O-C, for which there is no directly comparable MMPI scale. Boleloucky and Horvath (1974) reported a comparable study comparing SCL-90-R dimensions to the dimensions of the Middlesex Hospital Questionnaire ( MHQ). In their study, there was good convergence between like scales on the majority of test dimensions, with good discrimination between dissimilar scales as well.

More recently, Koeter (1992) evaluated the convergent-discriminant validity of the ANX and DEP dimensions of the SCL-90-R in comparison with the General Health Questionnaire (GHQ), and concluded that both instruments showed good convergent and discriminant validity. Similarly, Wiznitzer et al. (1992) utilized receiver operating characteristic (ROC) analysis to contrast the SCL-90-R with the Young Adult Self-Report (YASR) and the GHQ-28. The SCL-90-R and the YASR performed at equivalent levels in this population, with both outperforming the GHQ-28. Choquette (1994) contrasted the DEP dimension of the SCL-90-R with the Beck Depression Inventory (BDI) and Diagnostic Interview Schedule (DIS) criteria in identifying clinical depression in alcoholic patients, and concluded that the SCL-90-R and the BDI performed comparably, a finding similar to that of Moffett and Radenhausen (1983) in a comparable population.

Approaching validation from another perspective, Derogatis and Cleary (1977) cast the theorized dimensional structure of the SCL-90-R into a binary “hypothesis matrix” (i.e., each item was assigned a “1” for the factor it loaded on and a “0” for all others). Subsequently, data from the SCL-90-Rs of 1,002 psychiatric outpatients were factor analyzed and the solution was rotated toward the target matrix via the Procrustes method (Hurley and Cattell, 1962). Rotations were also accomplished via normalized varimax procedures (Kaiser, 1958). Comparisons of both solutions matched the hypothesized dimensional structure of the SCL-90-R cleanly, with only the PSY dimension showing some scatter.

A rigorous and systematic series of validation experiments reflecting elements of concurrent, criterion oriented, and construct validity for the SCL-90-R were reported by the British investigators Peveler and Fairburn (1990). They compared and correlated scores from the SCL-90-R with those from the Present State Examination (PSE; Wing, Cooper, & Sartorious, 1974), a clinician-administered, detailed, structured interview. Two distinct samples were utilized in the study: a sample of diabetics \( n = 102 \), representing a chronic medical disease group, and a cohort of bulimics \( n = 71 \), exemplifying patients with high levels of “neurotic” symptoms. Three distinct validation experiments comprised the study. In the first investigation, the case-finding power of the SCL-90-R was evaluated via ROC analysis and logistic regression analysis. In this experiment, the proficiency of the SCL-90-R to detect PSE-defined psychiatric “caseness” was evaluated. The instrument performed efficiently in
each instance, with areas under the curve (AUC) of .90 + .03 in both cases. In the diabetic sample the optimum sensitivity was 88% with a specificity of 80%, while with the bulimic sample, sensitivity was 76% with a specificity of 92%. Logistic regression analysis relating the GSI from the SCL-90-R to the probability of being a PSE-defined case also characterized the instrument favorably. Sensitivity among diabetics was 72%, and specificity was 87%. In the bulimic sample, values were 77% and 91%, respectively.

Peveler and Fairburn (1990) also evaluated the validity of the global indices of the SCL-90-R as accurate measures of general severity of psychopathology by correlating them with global indices from the PSE. Across both samples, all coefficients were statistically significant, and ranged from approximately .60 to .82. In addition, the validities of the SCL-90-R subscales were tested by evaluating their capacities to predict the presence of PSE syndromes through discriminant function analysis. Appropriate subscales were revealed in 12 of 14 instances in the diabetic sample, and 11 of 14 cases in the bulimic cohort.

Finally, another concurrent validation study with the DEP subscale of the SCL-90-R by correlating it with two independent depression inventories, the BDI and the Asberg Rating Scale Peveler and Fairburn (1990). Correlations were .80 and .81, respectively.

As has been noted previously, the type of validation of most interest to clinicians and researchers is the more tangible, pragmatic form, that is, “predictive,” or more generally, “criterion-oriented” validity. Current estimates, based primarily upon SCL-90-R: A Bibliography of Research Reports, 1975–1990 (Derogatis, 1990), SCL-90-R: Symptom Checklist-90-R Bibliography (National Computer Systems, 1996), and Brief Symptom Inventory (BSI) Bibliography (National Computer Systems, 1996), suggest that there are now well over 2,000 published reports pertaining to SCL-90-R criterion-oriented validation, and approximately half that many associated with the BSI. Therapeutic intervention studies evaluating treatments as diverse as meditation (Carrington et al., 1980), multicenter psychotherapy protocols (Shapiro & Firth, 1987), and numerous psychotropic drug trials (Ballenger et al., 1988; Noyes et al., 1984) attest to the instrument’s sensitivity to treatment-induced change. Characteristic SCL-90-R profiles for most major diagnostic groups have been established, including those for anxiety (Cameron, Thyer, Nesse, & Curtis, 1986), depression (Prusoff, Weissman, Klerman, & Rounsaville, 1980), panic disorder (Buller, Maier, & Benkert, 1986), and sexual dysfunctions (Derogatis, Meyer, & King, 1981). Such profiles have also been developed for recently delineated compound nosologic subtypes, for example, comorbid panic/depression (Wetzler, Kahn, Cahn, van Praag, & Asnis, 1990) and substance abuse (Steer, Platt, Ranieri, & Metzger, 1989). In addition to many studies of this nature, the SCL-90-R and BSI have been utilized as distress measures with most major medical illness groups (e.g., cancer, cardiovascular, diabetes, renal diseases).

**BSI Validity**

A comprehensive review of criterion-oriented validity studies involving the BSI was recently made available by Derogatis (1993). Approximately 120 research reports on the BSI were reviewed involving an extensive range of substantive areas. In addition, Derogatis and Derogatis (1996) also published a comprehensive review of research with both the SCL-90-R and the BSI. These studies collectively demonstrate the BSI to be broadly sensitive to the manifestations of psychological distress and interventions designed to ameliorate it across a broad range of conditions. As an illustration of the BSI’s general sensitivity to psychological distress status, several of the more interesting of these studies are briefly reviewed next.

Evidence for the BSI’s sensitivity in a screening paradigm is provided by a report that contrasted several methods for the psychosocial screening of newly diagnosed cancer patients (Zabora, Smith-Wilson, Fetting, & Enterline, 1990). These investigators reported an 84% “hit rate” for the BSI in
identifying patients who were determined by independent criteria to be suffering from clinical levels of distress, both at time of initial diagnosis and at 1-year follow-up. Additionally, a comparative cost benefits analysis resulted in a strong recommendation for the BSI.

Gift (1991) also reported on the sensitivity of BSI subscales, in this instance to differential respiratory status in a sample of adult asthmatics. In an attempt to determine the underlying causes of episodes of dyspnea (difficulty breathing) in these patients, she utilized the BSI and measured airway obstruction and oxygen saturation during periods of high and low dyspnea. Significant elevations were noted on the ANX, DEP, SOM, and HOS scales during periods of high dyspnea.

Very recently, Meshberg-Cohen, Siviks, and McMahon (2014) reported on a very unique study testifying to the sensitivity of the BSI. These authors conducted a randomized clinical trial contrasting expressive writing as a therapeutic intervention compared to control writing conditions in a sample of residential drug-dependent women. The women wrote for 20 minutes per day; at assessments at 2 weeks and 1 month follow-up the women in the expressive writing group showed a significant reduction in their BSI ANX scores.

Thompson, Gallagher, and Breckenridge (1987) demonstrated high sensitivity for the BSI in another study of treatment-induced change. These investigators compared the relative efficacy of three distinct psychotherapies in applications with depressed elderly patients. Although no substantial differences were observed between treatments, the BSI showed significant reductions in psychological distress for all three interventions across time, a finding that supported an alternate hypothesis.

Finally, in a study reported by Chiles, Benjamin, and Cahn (1990), the BSI was utilized with a random sample of 802 members of the Washington State Bar Association to contrast the psychological distress levels of smokers versus nonsmokers. Results showed that among male members of the bar, almost all BSI subtests revealed smokers to be significantly more highly distressed than nonsmokers. SOM, ANX, and DEP made the greatest contribution to discrimination, with the highly distressed group also showing significantly greater alcohol use. No comparable differences were observed among females, revealing a gender by smoking status interaction.

BSI-18 Validity

Obviously, there are many fewer validation reports for the BSI-18 because the test has become available only more recently. Even so, a confirmatory factor analysis was reported by Derogatis (2000), in which an attempt was made to reproduce the hypothesized (three-)dimensional structure of the instrument using the community normative sample of 1,134. Two dimensions, SOM and DEP, were replicated almost perfectly; however, the third hypothesized (Anxiety) dimension split into two subcomponents, a traditional general Anxiety dimension and a small (two items) Panic factor. Even with the division of the Anxiety dimension, the empirical structure of the BSI-18 matched the hypothesized structure of the instrument quite well.

More recently, Wiesner et al. (2010) studied the factorial invariance of the hypothesized structure of the BSI-18 across race and ethnicity in a large sample of women (N = 4,711). They found a great deal of consistency in the factor structure; however, there were significant differences in mean scores on the three dimensions between white and black samples. Petkus and his colleagues (2010, p. 582) also did a confirmatory factor analysis on the factor structure of the BSI-18 in a sample of homebound older adults. Their results showed that the “theorized three-factor structure of the BSI-18 fit the data well.”

A confirmatory factor structure based on 1,543 cancer patients was also reported by Zabora et al. (2001). This report also featured a study in which the BSI-18 was utilized as a screen for psychiatric caseness in this sample, using previously developed cutoff criteria on the full BSI as the standard. Sensitivity was reported as .91, with specificity as .96.
INTERPRETATION OF THE SCL-90-R AND THE BSI

The SCL-90-R and the BSI were designed to be interpreted in terms of three distinct but related classes of information: global scores, dimension scores, and individual test items. This strategy also holds for the BSI-18; however, because the item set is much smaller with the BSI-18 there is simply less to be learned from the item level. The optimal interpretation of the test protocol is dependent upon integration of information from all three source levels.

A significant advantage associated with all three instruments concerns the fact that test scores are reported in terms of standardized area T-scores. Scores of this type possess considerable advantages because they are based on a normalizing area transformation (as opposed to a linear transformation) of the raw score distribution. This feature gives the clinician the capacity to make statements concerning the respondent’s proportional status relative to the norms, and thereby to place him/her in an accurate normative (percentile) position. As an example, regardless of the specific score under consideration, an area T-score of 60 will always assign the respondent to the 84th percentile of the referent norm. Similarly, an area T-score of 70 will place the individual in the 98th percentile. This feature enables the clinician to not only make accurate comparisons between the patient’s status and various standards of interest, but also enables meaningful comparisons within individual profiles (e.g., comparison of levels of depression versus anxiety). As a result, a more meaningful interpretation of potential therapeutic change is possible.

Global Scores

The GSI represents the most sensitive single quantitative indicator concerning the respondent’s overall psychological distress status on the SCL-90-R/BSI/BSI-18 series of tests. It reflects information on both the number of symptoms of distress the individual is enduring, and the intensity level of his/her distress. By comparison, the PSDI is designed to be more of a “pure” intensity measure, adjusted for numbers of symptoms. The PSDI can also prove useful in communicating about the respondent’s distress posture, that is, whether he/she is apt to be an “augmenter,” typically exaggerating distress, or a “minimizer,” more likely to be stoic and understated. The PST reveals the number of symptoms the respondent has endorsed to any degree. It contributes to interpretation by conveying the “breadth” or array of symptoms that the individual is currently experiencing. Although there are no formal “validity” scales on the SCL-90-R/BSI, the PST can serve as a coarse indicator of whether or not the respondent is attempting to consciously misrepresent his/her status. In the case of symptom suppression, PST scores (i.e., the number of symptoms endorsed, regardless of the reported level of distress) of 3 or less for adult normal females and 2 or less for adult normal males on the SCL-90-R are extremely uncommon, and should be viewed with some misgivings. On the question of score augmentation, PST scores greater than 70 for females and greater than 65 for males are rarely observed as valid scores outside of psychiatric inpatient populations. Analogous scores are available for the BSI, printed in its standard administration manual (Derogatis, 1993). Although crude indicators, these values can be useful in identifying individuals in the community population with extreme response styles.

Dimension Scores and the SCL-90-R/BSI Profile

A major advantage of the SCL-90-R/BSI resides in the fact that they are designed to provide multidimensional symptom profiles, which contain a considerable amount of information concerning the nature and syndromal pattern of the respondent’s symptomatic distress. Multidimensional measurement
significantly enhances the utility of clinical assessment in comparison to unidimensional measurement by providing a syndromal context within which to understand specific dimensional psychopathology. In conjunction with global scores and data on specific symptoms, multidimensional measurement enhances the development of an integrated picture of the respondent’s current clinical status and level of psychological integration.

**Individual Symptoms**

The third element in the interpretive paradigm for clinical evaluation with these instruments involves the use of the discrete items or symptoms. Not only do we make reference here to the items comprising the nine primary symptom dimensions, but also the additional or “configural” items of the test. For example, an elevation on the DEP dimension scale plus a substantial score on suicidal ideation should be interpreted differently and with more immediate alarm than an equivalent DEP score in the absence of evidence of suicidal ideation. In such an instance, suicidal ideation would be treated as a “symptom of note,” the presence of which should clearly alter the clinical decision process. As another example, clinical levels of depression combined with early morning awakening, loss of interest, and high levels of guilt may signal the emergence of a major affective disorder. The same DEP score with a dissimilar pattern of accompanying symptoms might be interpreted as suggesting a less dramatic level of morbidity.

The configural items are not pure reflections of any one specific dimensional construct; they are designed to aid in making accurate predictions concerning the respondent’s clinical status. They represent clinically significant symptoms that are not unique to any of the SCL-90-R/BSI primary symptom dimensions. As examples, sleep and appetite disturbances represent potentially significant clinical manifestations. They do not occur solely in the context of a specific syndrome, but their presence in a particular individual can be a significant aid in clinical decision-making.

**Caseness Criteria for the SCL-90-R/BSI**

When the SCL-90-R, BSI, or any other psychological inventory or rating scale is utilized in a screening paradigm, an operational definition of “caseness” must be established. The caseness criterion essentially refers to the numerical value, that is, a “cutoff” score, on a test indicator, at or above which the respondent is considered to be a “positive” or a case. The caseness criterion is a probabilistic value, chosen to maximize valid case identification (e.g., sensitivity and specificity), and minimize errors (i.e., false positives and false negatives). In psychiatric screening, it is difficult to develop a definitive caseness criterion value for a particular test because other important parameters (e.g., gender, age, prevalence of the condition in the population being screened) can significantly affect the validity of any criterion value. Nevertheless, it is possible to establish a common criterion for caseness that has demonstrated generalizability across a range of populations and has proven useful in general screening situations.

A general caseness criterion value is given here for the SCL-90-R and the BSI. It is not possible in the context of this chapter to provide complete supporting data for the general caseness criterion given here; however, this criterion has shown effectiveness in accurately discriminating individuals who make up the normative community nonpatient cohort from those who make up the psychiatric outpatient sample. Further, in a multicenter epidemiologic study of the prevalence of psychiatric disorder in newly admitted cancer patients, the *predictive value of a positive* was 86% using this criterion (Derogatis et al., 1983). According to this definition a “case” is defined by:

\[ \text{Positive } D_x = (T_{GSI} \geq T_{63}) \text{ or } T_{2D} \geq T_{63} \]
This definition should be read as, if the respondent has a GSI $T$-score (using the community nonpatient norms) greater than or equal to a $T$-score of 63, or any two primary dimension scores are greater than or equal to a $T$-score of 63, then the individual shall be considered at high risk for a psychiatric diagnosis and therefore, a case.

**THE SCL-90-R AND BSI IN DEFINED CLINICAL POPULATIONS**

The studies cited in this section represent a small proportion of the research done with the SCL-90-R/BSI in medical cohorts, and indicate the potential value of brief measures of psychological status for evaluating patients and treatment outcomes in various psychiatric and medical populations.

**Anxiety, Depressive, and Somatoform Disorders**

There is compelling evidence (Derogatis & DellaPietra, 1994; Derogatis & Wise, 1989) that anxiety and depressive disorders account for approximately three-quarters of the psychiatric conditions seen in the healthcare system. Their diagnoses are frequently confounded by the fact that numerous somatic symptoms are often an integral part of the clinical presentation (Kirmayer, Robbins, Dworkin, & Jaffe, 1993; Simon & Von Korff, 1991). Some authorities have indicated they believe depression is the most prevalent problem in primary care (Katon & Sullivan, 1990), while Murray & Lopez (1997) suggest that by 2020 depression will be the second leading cause of disability in the United States after heart disease. The comorbidity of these conditions in primary care populations is extremely high, such that Lowe and his colleagues (2008) observed a 75% comorbidity with Anxiety or Somatization disorders in primary care patients diagnosed with depression. Recognition of the scope of the problem led the US Preventive Services Task Force to recommend that every primary care provider screen all adult patients for depression (Pignone et al., 2002).

Historically, the SCL-90-R has been utilized often in clinical research focused on depression. Weissman et al. (1977) used the instrument to characterize primary versus secondary depressions, as well as to describe and characterize depressions in five distinct populations (Weissman, Sholomsk, Pottenger, Prusoff, & Locke, 1977). Meanwhile, Wetzler et al. (1990) have profiled differences between depressed and panic patients on the SCL-90-R, and Stewart, Quitkin, Terman, and Terman (1990) contrasted atypical depressions with seasonal affective disorders using the instrument. In addition, Bryer, Borelli, Matthews, and Kornetsky (1983) used the SCL-90-R in a depressed sample to predict suppressors versus nonsuppressors on the dexamethasone suppression test (DST). Employing discriminant function analysis, these investigators were able to correctly predict DST status in 73% of cases.

In a demonstration of differential sensitivity for the SCL-90-R with high relevance for treatment, Rosenberg, Bech, Mellergard, and Ottoson (1991) used the instrument to discriminate various categories of panic patients with and without comorbid clinical depression. The SCL-90-R showed significant differences between patients diagnosed as having concomitant major depression, minor depression, and absence of mood disorder based on the Hamilton Rating Scale for Depression. In addition, the SCL-90-R effectively discriminated between diagnostic categories of current major depressive episode, other mood disorder, and no mood disorder based on the Standardized Clinical Interview for Diagnosis of DSM-III Disorders (SCID). The authors concluded that the data support a common diathesis for panic and mood disorders, with more severe cases of the condition being characterized by symptoms of both anxiety and depression. Vollrath, Koch, and Angst (1990) have also reported on comparisons of patients with panic disorder versus those with panic and comorbid depression with the SCL-90-R. These investigators found that the PHOB dimension, and to a lesser degree the anxiety dimension,
effectively discriminated these subgroups, with the panic/depression group revealing greater general severity and a more specific nosology.

Along similar lines, Andreescu et al. (2009) used the BSI to establish three subtypes of late-life depressed patients with concomitant anxiety. They found the “chronic worry” subtype to be less responsive to maintenance treatment with antidepressants in both time to response and time to recurrence measures. Gavin, Lindhorst and Lohr (2011) also used the BSI in a 17-year prospective study of depression in adolescent girls who were pregnant and unmarried. The principal findings were that those girls who had high depression scores antenatally tended to continue with depression throughout five subsequent time periods of the study, with prevalence increasing from 19.8% to 35.2%.

In the case of anxiety disorders, Cameron, Thyer, Nesse, and Curtis (1986) used the SCL-90-R to profile patients with distinct DSM-III anxiety disorders. This same research group (Cameron & Hudson, 1986) employed the instrument in an engaging study to evaluate the influence of exercise on severity of anxiety in patients diagnosed with anxiety disorders. Thirty-one percent of patients with panic attacks were exercise sensitive, compared to only 7% of other patients. The SCL-90-R ANX and PHOB subscales were particularly effective in making this discrimination. Ae Lee and Cameron (1986) evaluated the relationship between Type A behavior, symptom distress patterns, and family history of coronary heart disease among males and females with anxiety disorders. Significant correlations between SCL-90-R ANX and HOS scores and Jenkins Activity Scale (JAS) Type A scores were observed among males, but not among female patients.

For well over a decade, Katon and Roy-Byrne (1991) and other investigators argued for the existence of a mixed anxiety-depression syndrome, an aspect of phenomenology finally recognized in DSM-IV (American Psychiatric Association, 1994). They cite strong evidence to substantiate the existence of this syndrome as a distinct diagnostic entity, with studies involving the SCL-90-R contributing substantial empirical confirmation. Individuals afflicted with the condition are found to have a high incidence of medically unexplained problems, and be proportionally greater utilizers of healthcare systems. They also appear to be at increased risk for more severe anxiety and mood disorders. Similarly, Clark and Watson (1991) have developed a tripartite model of anxiety and depression. Based on a meta-analysis of psychometric data, they argue that at the clinical level, anxiety and depressive phenomena may be explained by a general distress factor and two specific factors of anxiety and depression. The authors mobilize an impressive body of data to support their theory, in particular noting that this pattern was very explicit in numerous studies with the SCL-90-R.

Katon et al. (1990), focusing on the prognostic value of somatic symptoms, used the SCL-90-R to provide an operational definition of “high-distressed-high-utilizers” within two large primary care practices. The high-distress group was further divided into four subgroups on the basis of numbers of unexplained somatic symptoms. The investigators observed linear increases in SCL-90-R dimension scores of SOM, DEP, and ANX, as well as independent diagnoses of psychiatric disorder, as they moved progressively through the somatic symptom subgroups from “low” to “high.” Kellner, Hernandez, and Pathak (1992) also reported an interesting study with the SCL-90-R and somaticizing patients. These researchers related distinct dimensions of the SCL-90-R to different aspects of hypochondriasis. In their study they observed high scores on the SCL-90-R SOM and ANX dimensions to be predictive of hypochondriacal fears and beliefs, while elevations on DEP were not. Further, they observed that fear of disease correlated most highly with the SCL-90-R ANX score, but that the false conviction of having a disease was more highly correlated with scores on SOM.
**Stress Conditions**

Some theorists view the construct of stress as little more than a chronic variant of anxiety with more explicit environmental linkages. Variations in the construct range from states of mild dysphoria arising from problems of daily living to a formal diagnostic entity, posttraumatic stress disorder (PTSD). Formal nosologic status was conferred on PTSD in DSM-III (American Psychiatric Association, 1980).

Addressing the more dramatic end of the stress spectrum, Horowitz, Wilner, Kaltreider, and Alvarez (1980) used the SCL-90-R to help distinguish PTSD from other anxiety-based disorders. Davidson, Kudler, Saunders, and Smith (1991) also used the SCL-90-R to profile the symptom patterns and severity of PTSD in groups of World War II versus Vietnam veterans. Vietnam vets exhibited more severe PTSD symptom and revealed higher distress scores on a number of SCL-90-R subscales. More recently, Weathers and his colleagues (Weathers et al., 1996) derived what they termed a “war-zone PTSD scale (WZ-PTSD)” from the SCL-90-R. They report the subscale to have good reliability and discriminative validity, with diagnostic utility (for PTSD) superior to that of a number of dedicated PTSD scales. A similar subscale specific for “crime-related PTSD” was reported by Saunders, Arata, & Kilpatrick (1990). This scale was highly effective, demonstrating 89% correct assignment in a discriminant function analysis that used the Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981) as an external criterion.

In addition to war, natural and man-made disasters have high potential for trauma, and can be extremely stressful to those who experience them. Winje (1996) reported on a longitudinal study with the SCL-90-R of the parents of children and spouses who were involved in a fatal school bus accident. The course and duration of posttraumatic symptoms were assessed at 1, 3, and 5 years after the accident. Analyses were done in terms of loss status and prior exposure to trauma. Significant proportions of the sample evidenced high levels of symptomatic distress (50%, 39%, and 42%, respectively) throughout the follow-up period. Individuals who suffered loss were not significantly more distressed than those who did not; however, individuals who had suffered previous trauma revealed a significantly smaller reduction in symptoms over time than did those free of prior traumatic experiences. Green, Grace, Lindy, Titchner, and Lindy (1983) also utilized the SCL-90-R to document residual levels of stress and functional impairment after another man-made disaster, the Beverly Hills Supper Club Fire. In a fashion similar to Winje (1996), they observed significant levels of residual symptomatology and distress.

**Suicidal Patients**

A prominent issue in treatment planning concerns the reliable early identification of the potentially suicidal patient. Several studies have addressed this question using the SCL-90-R/BSI. Bulik, Carpenter, Kupfer, and Frank (1990) contrasted 67 patients suffering from recurrent major depression and a history of attempted suicide with 163 recurrent depressives without a history of suicidal behavior. Four subscales (SOM, INT, PAR, and PSY), as well as the global scores, significantly discriminated attempters from nonattempters. Logistic regression analysis with these and other measures enabled 77% correct prediction of cases.

There is increasing evidence (Coryell, 1988) that panic disorder has associated with it an increased risk for suicide, just as a diagnosis of depression. In an analogous evaluation of panic patients who did and did not attempt suicide, Noyes et al. (1991) reported findings similar to those of Bulik and her colleagues (1990). Seven of the nine primary symptom dimensions of the SCL-90-R and the GSI successfully discriminated suicide attempters from those who did not make attempts. Just as Bulik et al.
(1990), these investigators found patients who made suicide attempts had greater severity of distress in general, with particular elevations on measures of inferiority feelings and self-deprecation.

Anxiety disorders more broadly defined continue to be identified as promoting increased risk for suicidal behavior in primary care populations. Bomyea and her associates (2013) have recently published a report that supports this hypothesis based on a large sample of primary care individuals ($N = 1,620$) referred as part of a Coordinated Anxiety Learning and Management (CALM) study. These investigators report that suicidal ideation and previous suicidal behavior were relatively commonplace in this population of primary care patients with anxiety disorders. The presence of an anxiety disorder, particularly with a comorbid mood disorder, physical impairment, or low social support led to a greater endorsement of suicidal ideation and behavior.

**Alcohol and Substance Abuse**

Because of the relatively high prevalence of alcohol and substance abuse disorders, and their prominent comorbidity with other psychiatric disorders, evidence of the utility of the SCL-90-R/BSI with these classes of patients is very important. A highly relevant study for this population was recently reported by Wang et al. (2010), who confirmed the factorial structure of the BSI-18 and its factorial invariance based on distinct drug-using populations from three different states. At the primary dimension level they confirmed the three-factor structure of the BSI-18 and further identified a single second order global factor—global distress. Invariance of the factor structure is an essential characteristic of any instrument that is intended to be used as an outcomes measure in research focused on assessing treatment program effectiveness.

In terms of sensitivity to symptoms of psychological distress in this population, Johnson, Brems, Mills, and Fischer (2007) observed a sample of 815 individuals receiving alcohol detoxification with the BSI. Results showed that symptomatic distress levels were very high, being manifest in close to 80% of the sample. The authors state that distress levels in this group most closely matched the levels of the psychiatric outpatient norm for the BSI.

Turning to drug dependency, a number of researchers have published studies with the SCL-90-R that have high relevance for treatment design. Carey, Carey, and Meisler (1991) demonstrated the dual impact of comorbid conditions in a study contrasting a heterogeneous sample of psychiatric patients who also abused drugs with a matched sample of psychiatric outpatients with no history of drug abuse. The sample with additional drug abuse had significantly higher symptom distress scores on six of nine subscales and all three globals of the SCL-90-R. Following on the work of Rounsaville et al. (1983), which showed a sensitivity of 89% for the SCL-90-R in detecting psychopathology among heroin addicts, Steer, Platt, Hendriks, and Metzger (1989) used modal profile analysis with Dutch and American cohorts of heroin addicts to identify three distinct subtypes based on the SCL-90-R: anxious-depressed, hostile, and paranoid. In addition to the observation that the paranoid subtype was much more likely to also use marijuana, the authors discussed a number of distinct treatment planning options that could hinge upon the availability of this information.

The same group of investigators (Steer, Platt, Ranieri, & Metzger, 1989) conducted a similar analysis of SCL-90-R data from 458 methadone patients. They observed the same three modal subtypes, and in addition, defined a fourth somatizing subtype. The potential utility and impact on treatment planning of subtype membership in this group of chemical abusers was also discussed. In demonstrating its sensitivity to differential levels of psychopathology in the patient with substance abuse, Kleinman et al. (1990) administered the SCL-90-R to three distinct groups of cocaine abusers: (1) those free of any additional DSM-III-R diagnosis, (2) those with an additional DSM-III-R Axis II (personality disorder) diagnosis, and (3) those with an additional DSM-III-R Axis I (clinical) diagnosis. Mean GSI
scores for the three groups were 0.53, 0.65, and 0.87, respectively, illustrating high levels of discriminative sensitivity.

Johnson, Brems, and Fisher (1996) also compared psychopathology levels of drug abusers not receiving treatment with those in treatment. Using data from Mercier et al.’s (1992) treatment sample, SCL-90-R scores were significantly higher for all subscales than the nontreatment sample except for the PSDI index. As predicted, drug abusers in treatment were found to be more symptomatic than those not in treatment except on the hostility and paranoid ideation scales where the nontreatment group exhibited higher levels. Approximately 60% of male nontreatment abusers and 47% of female nontreatment abusers obtained GSI scores equal to or greater than the cutoff for caseness warranting a dual diagnosis. The implications of this study are that the presence of a comorbid condition is associated with a greater likelihood that drug abusers will seek treatment.

### Sexual Victimization

Sexual abuse and victimization is a trauma and source of distress that can convey a long-standing residual emotional vulnerability. Both physical and sexual abuse, particularly during childhood, are extremely traumatic experiences that can have dramatic psychological sequelae. As examples, Kelly (1990) reported a study dealing with the stress engendered in the parents of children who have been abused. She contrasted SCL-90-R symptom profiles of parents of children who were sexually abused, a second group whose children were ritually abused in the context of cult worship, and the parents of nonabused controls. Results showed both groups of parents of abused children displayed substantially elevated profiles, with the parents of the ritually abused children being significantly more distressed than the parents of the other abuse group.

Coffey, Leitenberg, Henning, Turner, and Bennett (1996) investigated 192 women with a history of childhood sexual abuse and whether their methods of coping with victimization resulted in healthy psychological adjustment. Women who had been sexually abused revealed a higher GSI score on the BSI than women in the nonabused group, and a greater proportion of their BSI subscale scores fell in the clinical range. These findings tend to support the idea that women with a history of sexual abuse tend to experience greater difficulties with psychological adjustment in general. In terms of coping strategies, most victims of sexual abuse utilized methods of disengagement, a strategy that contributed to higher degrees of psychological distress. The authors suggest that it is important to appreciate how specific methods of coping with sexual abuse, especially disengagement, can be ultimately counterproductive, and lead to greater distress and poorer adjustment.

Using the BSI, Bennett and Hughes (1996) attempted to develop a distress norm for a college population by examining a cohort of female college students who were victims of sexual abuse. Results of their evaluation showed that abuse victims had substantially elevated BSI scores and increased adjustment problems compared to individuals absent an abuse history. Further, they found that college females who had suffered sexual abuse demonstrated BSI symptom profiles essentially equivalent to individuals undergoing psychological treatment.

Chronic pain is also a condition that has been found to be associated with earlier sexual abuse. Toomey, Seville, Mann, Abashian, and Grant (1995) evaluated a heterogeneous group of chronic pain patients and observed that those patients with a history of sexual abuse scored higher overall on the SCL-90-R than nonabused patients. The authors advanced the hypothesis that the abuse experience may sensitize individuals, resulting in a greater frequency of site-specific pain syndromes, and greater psychological distress. Similar findings were reported by Walker et al. (1995), who also found that female patients with chronic pelvic pain evidenced significantly higher symptomatic distress levels compared to patients without pain. In fact, the mean score for chronic pelvic pain sufferers fell in the
60th percentile of the psychiatric outpatient norm on the majority of SCL-90-R subscales. The pain group also was found to have a history of diagnosable psychiatric disorders, especially major depression, as well as somatization disorder, drug abuse, phobia, and sexual dysfunction. They also revealed a significantly greater incidence of sexual abuse as compared to a nonpain (tubal ligation) group.

Recently, Arcelus and his associates (Arcelus, Claes, Witcomb, Marshall and Bourman, 2016) used the SCL-90-R to document distress levels in a large sample of transgender youth with a focus on describing the subgroup that participated in nonsuicidal self-injury (NSSI). They observed higher distress levels on the SCL-90-R with significantly poorer self-esteem in those individuals who reported higher levels of NSSI.

Medical Populations

From their inception, the SCL-90-R and the BSI were designed for applications in primary care and specialized medical populations. In terms of screening, these populations almost certainly contain the highest prevalences of occult psychiatric disorder (Derogatis & DellaPietra, 1994). Snyder, Lynch, Derogatis, and Gruss (1980) reported an early study with the SCL-90-R in a family practice setting. Their research showed that those patients who had significant communications problems with their physicians also demonstrated significantly higher symptom profiles on the SCL-90-R. Weidner, Connor, Hollis, and Connor (1992) used the SCL-90-R to show significant decreases in DEP and HOS scores were associated with reductions in serum cholesterol over the course of a 5-year dietary intervention program. Working with diabetics, Irvine, Cox, and Gonder-Fredrick (1992) observed that worry over hypoglycemia, and behaviors focused on avoiding this condition, were clearly correlated with elevations on multiple SCL-90-R dimension scores.

The instruments in this series have also been utilized extensively in oncology. Early in its development Craig and Abeloff (1974) utilized the SCL-90-R to demonstrate clinical levels of psychological distress in cancer patients, and Abeloff and Derogatis (1977) used the scale to describe the specific psychological symptom picture of breast cancer patients. Derogatis, Abeloff, and Melisaratos (1979) employed the SCL-90-R to show that length of survival with metastatic breast disease was distinctly related to coping style, a finding also reported by Rogentine et al. (1979) with a malignant melanoma sample.

The BSI has also been used extensively with oncology populations. For example, Baider, Peretz, and Kaplan DeNour (1992) evaluated a heterogeneous group of cancer patients who had completed treatment, some of whom were also Holocaust survivors. Consistent with other research demonstrating the vulnerability associated with previous trauma, the Holocaust survivors revealed significantly greater distress. Gilbar (1991) also used the BSI to compare a heterogeneous group of cancer patients who completed their chemotherapy regimen to a group who terminated therapy prior to completion. Among other findings, the patients who dropped out of chemotherapy scored significantly higher on hostility and a number of other BSI scales. Further on this topic, Gotay and Stern (1995) have provided a very useful review of SCL-90-R/BSI studies in oncology.

More recently, Vodermaier, Linden, and Siu (2009) did an extensive assessment of instruments designed to screen for emotional distress in cancer patients based on 2,747 publications. Concerning the BSI-18, this review concluded that

the BSI-18 demonstrated excellent reliability and validity in a large mixed sample of cancer patients with a sensitivity of .91 and specificity of .93 respectively, and in adult survivors of childhood cancer with a sensitivity and specificity of .97 and .85, respectively. Internal consistency was high for the
anxiety and depression subscales. Results of a factor analysis confirmed the scale’s three factor structure (i.e., depression, anxiety and somatization).

(p. 1481)

Omidi and Zargar (2015) also reported an innovative trial of a mindfulness intervention with tension headache sufferers, using the BSI as a primary outcomes measure, and demonstrated high sensitivity to therapeutic change. Similarly, Gluck and Maercker (2011) used the BSI as an outcomes measure in testing the feasibility of a web-based mindfulness intervention, and Kulz et al. (2014) used the BSI-18 in an intriguing randomized clinical trial of mindfulness-based cognitive therapy with obsessive-compulsive disorder.

An interesting cardiology application of one of the scales was recently published by Ruz et al. (2010). These investigators administered the BSI ANX subscale to patients who had recently been admitted to hospital for an acute myocardial infarction. They established high reliability for the scale and convergent validity with the State-Trait Anxiety Inventory (STAI). Patients with high anxiety scores had higher rates of complications, and in a logistic regression, the BSI was an independent predictor of complication rate.

Also in the area of cardiology, Bringager, Friis, Arnesen, and Dammen (2008) reported a very interesting prospective study. Focused on the observation that up to 60% of patients referred to cardiology/emergency departments suffering with chest pain are ultimately diagnosed with panic disorder, these investigators followed patients who initially presented with chest pain for close to a decade. At the end of 9 years, they found that individuals with SOM scores of 69 or greater on the SCL-90-R at baseline demonstrated a fivefold increase in the rate of enduring panic disorder 9 years later.

Several studies have employed the BSI to investigate psychological factors associated with HIV infection. Kennedy, Skurnick, Foley, and Louria (1995) examined psychological distress among heterosexual couples with at least one partner HIV positive. Contrary to prediction, family support was not found to play a role in emotional distress. Gender was the only variable found to significantly affect psychological well-being such that females had higher elevations on all BSI® subscales than males. This was true for both HIV positive and HIV negative females with HIV positive male partners. It is assumed from these findings that women in a relationship affected by HIV have greater difficulties coping than men. Hopefully, awareness of these psychological vulnerabilities will influence clinicians to institute specific treatment for HIV patients and their partners.

Research on HIV positive and HIV negative homosexual men with a diagnosable personality disorder was conducted by Johnson, Williams, Rabkin, Goetz, and Remien (1995). HIV positive men with personality disorders indicated significantly more psychological distress on BSI DEP and ANX scales and the GSI than HIV negative men and men without a personality disorder. Furthermore, one-third of those HIV positive men with personality disorders (N=21) also had a comorbid Axis I disorder. It was concluded that the presence of both HIV and a personality disorder may enhance vulnerability to concurrent Axis I clinical disorders, particularly anxiety and depression.

HEALTH SYSTEMS PLANNING

The SCL-90-R and BSI have also been utilized effectively in treatment planning studies with a health-care systems orientation. Katon et al. (1990) used the SCL-90-R to define “highly distressed” patients among 767 high healthcare utilizers in a large HMO. Fifty-one percent of the sample fit their criterion. Not only did these patients make disproportionate use of healthcare facilities, they also revealed
a high prevalence of chronic medical problems, experienced significant limitation of activities associated with their illnesses, and had substantially elevated prevalence of major depressive disorder, dysthymia, and anxiety disorders. From a somewhat analogous perspective, Drossman et al. (1991) evaluated the nature of healthcare behavior in a sample of almost 1,000 patients with inflammatory bowel disease. In this study, the SCL-90-R was found to have significant predictive value in a regression model predicting the number of physician visits during the previous six months.

Perhaps the most dramatic study of this type was a 6-month follow-up study reported by Allison et al. (1995) with a sample of 381 cardiac rehabilitation patients, referred for a variety of cardiovascular disorders and/or procedures. Using the SCL-90-R, these investigators partitioned their cohort into “high psychological distress” versus low distress groups. Comparisons across the 6-month interval revealed that the high-distress group had significantly higher rates of cardiac rehospitalization and recurrent cardiac events compared to the low distress group. More striking, however, was the fact that the mean rehospitalization costs for the high-distress patients was almost five times the mean cost of the low distress group (i.e., $9,504 versus $2,146). The authors concluded that psychological distress has an obvious adverse impact on coronary patients, and systematic assessment programs should be instituted to accomplish successful identification and appropriate treatment of these patients.

In a somewhat unique investigation focused on the infrequently assessed variable of patient weight and its effects on personal distress, Martinez and her colleagues (2014) evaluated the relationship between body mass index (BMI) and psychological distress in a sample of 563 residents of a Spanish town (Malaga). They classified individuals into four ordinal categories on the basis of their BMI (underweight, normal, overweight, and obese), and subsequently assessed psychological distress using a Spanish version of the SCL-90-R. They observed a systematic U-shaped relationship ($p < .001$ for quadratic trend) between the respondents’ weight category and levels of psychological distress. The U-shaped relationship was present on each of the instrument’s nine primary symptom dimensions, as well as the global GSI score. Essentially, the results showed underweight and obese patients with the highest distress scores, while normal weight patients had the lowest. There were essentially no differences between males and females in this regard.

THE SCL-90-R AND THE BSI AS TREATMENT OUTCOME MEASURES

The ideal outcomes instrument will be highly sensitive to a broad range of treatment interventions and will demonstrate sensitivity to change along the entire spectrum of psychological dysregulation, from mild disaffection and dysphoria in community populations to dramatic psychopathology in institutionalized individuals. Limitations to sensitivity, either qualitative or quantitative, along the distress continuum can seriously constrain the usefulness of an outcomes measure. Similarly, desirable outcomes instruments are sensitive to changes induced by a wide variety of therapeutic interventions, and are not limited to narrowly registering the effects of very specific therapeutic modalities. The sections that follow will endeavor to demonstrate the extremely broad sensitivity of the SCL-90-R/BSI, both to the broad continuum of psychological disregulation and to the effects of an extensive spectrum of traditional and nontraditional therapeutic interventions.

Considerations for Outcomes Measures

For a psychological test to achieve optimal utility as an outcomes measure, it should possess the capability of documenting the test respondent’s status in meaningful clinical terms. Test scores, in and of
themselves, are insufficient to communicate real-world status because the constructs that psychological tests serve to operationalize (e.g., depression, anxiety, quality of life) are intangible. This means that good psychological outcomes measures should have representative, well-developed norms, to enable the interpretation of a patient’s score or change of status in meaningful terms. Well-constructed norms are designed to communicate the probabilistic expectation of a particular test score in the referent population of interest (e.g., community adults, psychiatric outpatients), and help establish the phenomenologic meaning of the patient’s current status and any changes that have taken place.

An important refinement of any such library of norms, at least concerning psychopathology, is that it be gender-keyed. It is well established (although often overlooked) that men and women are distinct in their reports of emotional distress and psychological symptoms, with women being much more likely to acknowledge emotional distress. Norms that fail to take these powerful effects into account can result in seriously distorted interpretations.

Another important aspect of valid outcomes measurement concerns the distinction between statistically and clinically significant change. It has been apparent for some time now (Garfield, 1981; Jacobson, 1988; Jacobson, Follette, & Ravenstorf, 1986) that significant differences defined on a purely statistical basis are not synonymous with clinically meaningful differences. This situation has been accepted in large measure because no realistic solution to the problem has been offered. However, some advances have been made in this regard.

Jacobson and Truax (1991) have proposed a dual-criterion method for determining the clinical significance of therapeutically induced change. Optimal application of their technique requires that norms be available for both “normal” or community individuals, and the “clinical” group (e.g., psychiatric outpatients, inpatients) under evaluation. The dual criterion for clinically significant change requires that (1) the patient return to normal functioning, and (2) a reliable magnitude of change take place. Meeting the former criterion relies on establishing a cutoff score for discriminating “healthy” from “dysfunctional” status, a value determined from the test’s normative distributions. A reliable change index (RCI) is calculated based on the standard error of the difference between an individual’s pre- and posttreatment scores. Only if an individual’s posttreatment score crosses the cutoff into the functional distribution and exceeds the magnitude of the reliable change index is the change considered clinically significant. The SCL-90-R and BSI are among the very few instruments currently available with carefully constructed, community and clinical, gender-keyed norms, based on accurate area T-scores. With norms available for community adults, community adolescents, and inpatient and outpatient psychiatric patients, the SCL-90-R/BSI are among the few psychiatric outcomes measures that enable calculation of clinically significant change across a broad spectrum of clinical populations. Both Jacobson and Truax (1991) and Lambert (1994) provide more detailed discussions of this methodology.

A study utilizing the Jacobson and Truax (1991) methodology was recently reported with the BSI in the context of a trial focused on the inpatient treatment of anorexia nervosa. Schlegl, Quadflieg, Lowe, Cuntz, and Voderholzer (2014) reported on a study involving 435 inpatient anorexics with a mean treatment duration of approximately 90 days. Multiple outcomes measures were used; however, the BSI showed 34.1% clinically significantly improved, 32.05% reliably improved, and 34.5% unchanged. This approach to establishing the clinical significance of treatment interventions provides much more useful information than simply whether or not the treatment effect was statistically significant.

Outcomes in Clinical Drug Trials

Pharmacotherapeutic drugs represent one of the cornerstones of the modern treatment of psychiatric disorders, both as sole interventions and in conjunction with psychotherapeutic approaches.
The SCL-90-R and BSI have proven their value as primary outcomes measures in drug trials for over 30 years, and have accrued substantial utility and validity in this capacity. For example, Ravaris, Robinson, Ives, Alexander, and Bartlett (1980) used the SCL-90-R in the first definitive double-blind controlled trial comparing a monoamine oxidase inhibitor (MAOI) with a tricyclic antidepressant (TCA). These investigators compared the tricyclic antidepressant amitriptyline to the MAO inhibitor phenelzine in the treatment of 105 depressed outpatients. Results demonstrated both drugs to have significant efficacy beyond placebo in reducing symptomatic distress over the 6 weeks of the trial, and in drug–drug comparisons phenelzine proved significantly better than amitriptyline in reducing anxiety. The study showed by comparisons with community norms for the SCL-90-R that although distress was significantly reduced at the end of 6 weeks, it remained elevated above normal levels. A number of additional trials using the SCL-90-R have also demonstrated the efficacy of phenelzine. Soloff et al. (1993) compared phenelzine with haloperidol and placebo in a randomized, double-blind trial with a sample of hospitalized borderline personality disorders. Phenelzine was found to be superior to haloperidol in the treatment of these patients on multiple SCL-90-R dimension and global scores. McGrath, Stewart, and Nunes (1993) also utilized the SCL-90-R in a comparison of phenelzine and imipramine in an intervention with treatment-refractory depressed outpatients. Analysis of symptomatic response showed that 67% of those patients who were refractory to imipramine showed clinical improvement on phenelzine.

The SCL-90-R also served as one of the principal outcome measures in a large multicenter trial (Ballenger et al., 1988) evaluating the efficacy of alprazolam in the treatment of agoraphobia and panic disorder. In this study, the scale demonstrated substantial efficacy for alprazolam compared to placebo. Also, Woodman and Noyes (1994) used the BSI to evaluate the efficacy of divalproex sodium in the treatment of panic disorder. All patients were moderately to markedly improved, with a large majority showing sustained improvement at 6 months follow-up. In a strong demonstration of the SCL-90-R’s sensitivity to drug effects, Noyes et al. (1984) reported a double-blind crossover comparison of diazepam (Valium) versus the beta-blocker propranolol (Inderal) in the treatment of panic-driven agoraphobia. SCL-90-R measures of ANX, PHOB, and the GSI showed very significant efficacy for diazepam against propranolol in this study that had no placebo group.

**Psychotherapy Trial Outcomes**

Psychotherapeutic efficacy, in both absolute and relative terms, is an issue of major interest in contemporary healthcare. Does psychotherapy work? Does one psychotherapy work better than another, and if so, for whom? Does the incremental benefit of adding psychotherapy to a drug treatment regimen justify the additional costs? Is psychotherapy more effective than drugs for some disorders, and if so, which? These are all important questions with high relevance for today’s healthcare. In order to obtain answers to these questions, numerous outcomes studies have been conducted assessing the efficacy and effectiveness of psychotherapy. Many of them have utilized the SCL-90-R/BSI as primary outcomes measures in their evaluations.

The meta-analysis of brief dynamic psychotherapy (BDP) studies reported by Crits-Christoph (1992) represents a convincing demonstration of the sensitivity of the SCL-90-R to psychotherapy outcomes. Aggregating over almost one dozen studies, this analysis highlighted sensitivity to BDP efficacy in comparison to waiting list control where effects were large ($d = 0.82$), revealed small effects in comparison to nonpsychiatric interventions ($d = 0.20$), and shared equivalent effects when compared to alternative psychotherapies ($d = 0.05$). These results are consistent with those of earlier trials. For example, Horowitz, Marmar, Weiss, Dewitt, and Rosenbaum (1984) studied the efficacy of BDP
with bereaved individuals. They found SCL-90-R ANX and DEP subscales and global measures were highly sensitive to treatment-induced improvement. They further noted that magnitude of distress reduction was significantly correlated with baseline distress levels.

The SCL-90-R has demonstrated sensitivity to the therapeutic impact of less typical methods of intervention. Bohachick (1984) reported significant reductions in distress among a cohort of hypertensives exposed to the addition of a progressive relaxation paradigm to their standard exercise regimen compared to exercise only. Carrington et al. (1980) compared two distinct meditation techniques to progressive relaxation and waiting list control in a sample of 154 self-defined high-stress individuals. Evaluations at the end of 6 months on the SCL-90-R revealed the two meditation techniques to be significantly better than progressive relaxation at reducing symptomatic distress. In one of the more unusual therapeutic outcome studies in the literature, Griffith, Mahy, and Young (1986) reported significant reductions in symptomatic distress as a result of participation in the West Indian ritual of Spiritual Baptist "mourning." With the exception of the Somatization dimension, all SCL-90-R measures showed significant efficacy for the solitary contemplative experience in reducing psychological symptomatic distress.

The SCL-90-R has also been utilized to evaluate alternatives to dynamic psychotherapies. Fairburn et al. (1991) contrasted two variations of cognitive behavioral therapy (CBT) and interpersonal therapy in a sample of bulimic patients. Although all three interventions showed efficacy on the SCL-90-R from admission to treatment termination, none of the interventions showed treatment superiority. Beck and colleagues (Beck, Stanley, Baldwin, Deagle, & Averill, 1994) used the SCL-90-R to establish the relative efficacy of CBT versus relaxation training and a minimum contact condition for the treatment of panic disorder in a small group format. At the end of 10 weeks, the PHOB dimension revealed significant reductions in the CBT group compared to the other two interventions. Shear, Pilkonis, Cloitre, & Leon (1994) also compared CBT with nonprescriptive reflective treatment over 12 treatment sessions in a sample of patients with panic disorder. SCL-90-R profiles at discharge and at 6-month follow-up showed the two interventions to be equally effective at reducing symptomatic distress, a finding shared with other outcomes measures. At follow-up, however, the SCL-90-R demonstrated continued improvement for the nonprescriptive reflective treatment, a finding not shown by other measures.

THE SCL-90 ANALOGUE AND DEROGATIS PSYCHIATRIC RATING SCALE

A specific advantage associated with the SCL-90-R/BSI concerns the fact that valid, matched clinical rating scales exist that may be used in conjunction with the self-report measures. If clinicians’ judgments about the patient’s psychological status are important to the project of interest, the same symptom constructs may be measured from both patient and clinician perspectives. Differences in perceptions can be accurately evaluated by comparing clinician judgments with patient self-ratings. Comparisons can be greatly facilitated by converting both sets of measurements to respective standardized scores, thereby enabling comparisons in a common metric.

As mentioned previously, the SCL-90 Analogue is a clinical observer’s rating scale designed specifically for the health professional without detailed training in psychopathology or mental health. The SCL-90 Analogue is brief and uncomplicated, usually requiring less than five minutes to complete. In addition to representations for the nine SCL-90-R symptom dimensions, the rating scale also contains an analogue global distress scale.
An example of the use of the SCL-90 Analogue scale is provided by a study done by Derogatis, Abeloff, and McBeth (1976) with a small sample of cancer patients. Shortly after admission, patients completed an SCL-90-R. Subsequently, the primary treating oncologist filled out an SCL-90 Analogue Scale on the patient based upon a clinical interview. Raw scores were converted to area T-scores for each patient on each measure, and doctor–patient difference scores (TΔ) were calculated. Results showed that as physicians’ ratings of global psychological distress rose, they tended to judge the patient to be increasingly distressed on interpersonal sensitivity and anxiety dimensions, but viewed much less distress arising from depression than did the patient. Analyses also demonstrated that the highest subscale correlations with the physicians’ independent global ratings of patient psychological distress were on ANX (r = .50) and HOS (r = .48). Correlations between the physicians’ global distress ratings and the patients’ self-rated global scores showed only the correlation with the PSDI (r = .43) to be significant. This result indicated that oncologists were basing their judgments much more on selective indicators of distress rather than numbers of manifest symptoms.

In another study with cancer patients, Schleifer et al. (1991) used the SCL-90 Analogue to evaluate factors that affect oncologists’ adherence to chemotherapy protocols. The sample consisted of 107 breast cancer patients who were followed for 26 weeks of treatment. Fifty-two percent of patients experienced an unjustified regimen modification. Physician perception of psychological distress was not a significant factor in modifying prescription in the majority of protocols; however, on the vincristine protocol, the global severity score and a number of SCL-90 Analogue subscale scores were significantly related to nonadherence. Steer and Hasset (1982) also used the SCL-90 Analogue to identify the differential weights assigned various dimensions of psychopathology in arriving at staff judgments of global severity of illness. Over 1,000 mental health patients were contrasted with 809 substance abuse clients. They found that I-S and PSY were the best predictors of global severity ratings in mental health patients, while ANX and PAR scores predicted best among ratings of substance abusers.

The Derogatis Psychiatric Rating Scale (DPRS) has also been utilized in a variety of interesting studies. Winokur, Guthrie, Rickels, and Nael (1982) used the DPRS as a validating instrument for patients’ self-ratings of psychological distress on the SCL-90-R. Approximately 60 nonpsychiatric medical patients from two settings participated in the trial. Two psychiatrists who were completely unaware of each other’s or patients’ self-report completed all DPRS ratings. Psychiatrist-patient correlations were generally high, with DEP (r = .63), ANX (r = .63), and PHOB (r = .72) showing the highest agreement. The authors report sensitivities for the SCL-90-R depression scale of .91 and .89 in the two groups of patients, with specificities of .78 and .85, respectively. Perconte and Griger (1991) used both the DPRS and the SCL-90-R to discriminate differential treatment responders among Vietnam veterans suffering from PTSD. Although the investigators did not report on levels of agreement between the two instruments, both were highly successful in discriminating successful, unchanged, and relapsing patients. Similarly, Fricchione et al. (1992) used the DPRS and the SCL-90-R to evaluate high versus low deniers among patients with end-stage renal disease. DPRS subscales of interpersonal sensitivity, anxiety, and sleep disturbance were significantly elevated among the low deniers, as were numerous SCL-90-R scales.

**CONCLUSION**

The SCL-90-R, BSI, and BSI-18, along with their matching clinical rating scales, represent a unique set of brief, multidimensional measurement instruments for the assessment of psychological symptoms and psychological distress. As such, they offer the psychologist several options for screening and/or
assessing patients for psychological disorders in busy primary care settings. Their successful use in hundreds of published outcomes research and clinical studies, across an extremely broad spectrum of applications, provides convincing confirmation of their reliability, validity, and utility. Sensitivity to pharmacologic, psychotherapeutic, and other treatment interventions, as well as sensitivity to clinically meaningful changes in psychopathology and psychological distress states, provides a compelling endorsement for these test instruments as broadly effective measures of clinical status and change. An additional advantage of these instruments is that the scales are available in more than two dozen languages and have been extensively utilized worldwide.

REFERENCES


There has been a growing interest in psychological assessment in medical settings, which have increasingly adopted a biopsychosocial perspective (Fava & Sonino, 2007). This evolution was likely prompted by increased recognition of the role of psychosocial factors in medical practice. Although a surgical procedure can be deemed successful from a medical perspective, patients can nonetheless evidence suboptimal outcomes or even worsening of functioning as a result (Block & Sarwer, 2013). In many circumstances, these negative outcomes may at least partially be accounted for by psychopathology or behavioral nonadherence (Block & Sarwer, 2013). In other instances, a patient may exaggerate or even feign symptoms in an effort to obtain disability payments or workers' compensation (McDermott & Feldman, 2007), or attention associated with the sick role, as in the case of factitious disorder (APA, 2013). Conversely, some patients minimize symptoms in an effort to appear to be well-adjusted during surgical evaluations (Ambwani et al., 2013).

Psychological assessment allows clinicians to gather a large amount of information efficiently, and provides means for synthesizing this information, thus maximizing the use of professional time. An integrative assessment, using objective psychometric testing in conjunction with a clinical interview and medical chart review, can offer more precise, empirically validated information about the patient than a clinical interview alone (Grove & Meehl, 1996). Testing can provide clinically relevant information on risk factors that the patient may be sensitive about disclosing during an interview; it can also aid in differential diagnoses and provide information on the extent to which the patient may be over- or under-reporting symptoms. In sum, psychological assessment that includes an objective, broadband instrument can play an important role in medical settings by informing multidisciplinary teams about how psychopathology symptoms, behavioral tendencies, and personality characteristics contribute to patients' medical conditions, and it can aid providers in treatment planning.

The Minnesota Multiphasic Personality Inventory (MMPI) instruments have long been used in medical settings (Arbisi & Seime, 2006). In fact, the original MMPI was developed and initially used for assessing both psychiatric and general medical patients (McKinley & Hathaway, 1943). However, the earlier versions of the test took over 90 minutes to administer, had many different interpretative strategies that were all not well validated, and included scales that were not psychometrically sound. This chapter introduces the rationale for, and development of, the newest version of the MMPI, the MMPI-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011; Tellegen & Ben-Porath, 2008/2011) and its use in medical settings. The test offers a number of advantages when used in primary care settings. First, the MMPI-2-RF provides a broadband assessment of constructs (discussed later) relevant to this area of practice (e.g., depression, anxiety, thought disorder, behavioral disinhibition, substance use, and family problems), with a relatively brief administration time (25–35
minutes via computer administration; Ben-Porath & Tellegen, 2008/2011). As detailed later, its use rests on sound conceptual foundations, and there is strong empirical evidence to support its use in medical settings. Use of the MMPI-2-RF can aid or challenge diagnostic assessments; provide clinical information on risk factors that patients may be sensitive to disclosing directly to their provider early on, when rapport may not yet have been established; and provide information about possible symptom over- or under-reporting symptoms.

In this chapter, a brief overview of the rationale for, and development of, the MMPI-2-RF scales will be presented, followed by a brief discussion of their psychometric properties. Next, administration, scoring, and interpretation will be discussed. Lastly, the literature on using the test in medical settings is reviewed followed by a case example.

THE MMPI AND MMPI-2 IN MEDICAL SETTINGS

The MMPI instruments have been among the most widely used psychological assessment measures for over seven decades. Hathaway and McKinley (1943) designed and published the first iteration of the MMPI to provide an assessment of targeted disorders, diagnosed based on the descriptive nosological system put forth by Kraepelin (1921). The test was initially developed for use in medical settings to help distinguish patients with genuine medical complaints from those whose problems had a psychological underpinnings, such as anxiety-related symptoms (McKinley & Hathaway, 1943). Both the clinical samples and the nonclinical sample used to develop the scales of MMPI were collected at the University of Minnesota Hospital. For many years, the MMPI was one of the most frequently used psychological tests in medical settings (cf. Piotrowski & Lubin, 1990).

Despite its widespread use, criticisms of the MMPI and calls for revisions of its scales began to surface in the 1960s and the 1970s. Hathaway (1960) discussed how he had hoped the MMPI would serve as a means to develop improved instruments that would be widely adopted. In the next decade, Hathaway (1972a) began to become less optimistic about the test, stating that if no further development of its scales were to occur, he feared that it would turn from being “a hopeful innovation to an aged obstacle” (p. xiv). In a separate paper, Hathaway (1972b) stated that the lack of progress in refining the Clinical Scales of the MMPI was likely due to the absence of an alternative diagnostic system at the time. Butcher (1972) edited a book based on presentations at the Fifth Annual Symposium on Recent Developments in the Use of the MMPI, which focused on the prospect of revising the measure and how to go about doing so. Loevinger (1972) suggested that the scales of the test should measure quantitative traits congruent with personality constructs. Moreover, she was concerned about a general factor variance that not only saturated the MMPI Clinical Scales scores, but other self-report instruments as well. Norman (1972) identified the Clinical Scales’ heterogeneity and excessively high intercorrelations as the most serious problem with the MMPI. Meehl (1972) commented that the role of theory in test construction and interpretation should be a focus of any MMPI revision.

In 1982, a committee assembled by the MMPI Publisher, the University of Minnesota Press, launched the “MMPI Restandardization Project” to develop the second iteration of the inventory. The committee’s goals were to maintain continuity with the scales of the original MMPI, revise test items (and eliminate items that were not working properly), and update the test norms. However, in order to maintain continuity and familiarity with its previous version, the developers of the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) did not address many of the fundamental criticisms of the original MMPI. Nonetheless, the MMPI-2 was accepted and widely used by many
practicing clinicians. It maintained its status as one of the most widely used psychological assessment instruments, including in medical settings. In fact, 90% of clinical health psychologists reported using the MMPI-2 in their practice (Piotrowski & Lubin, 1990) and it was the most frequently used measure by neuropsychologists (Camara, Nathan, & Puente, 2000). The most common uses of the MMPI-2 in medical settings included screening for psychopathology and substance abuse problems, psychological effects of medical conditions, responses to medical treatment, and predicting future pain symptoms (Graham, 2011).

DEVELOPMENT OF MMPI-2-RF

Despite widespread use of the MMPI-2 in behavioral medicine settings, research on the instrument itself suggested structural problems, most notably with the MMPI-2 Clinical Scales (cf. Jackson, 1971; Loevinger, 1972; Meehl, 1972; Norman, 1972). It was long recognized that the Clinical Scales of the instrument were not psychometrically optimal. The scales contained heterogeneous item content (which can produce ambiguous scale scores), subtle items (which contribute random variance), and item overlap, which artificially inflated scale intercorrelations. In fact, many of the correlations between the Clinical Scales of the MMPI-2 were inconsistent with known comorbidity rates (Tellegen et al., 2003). To deal with this challenge, interpretative strategies were developed, such as reliance on codetypes, subscales, and a host of supplementary scales. However, these strategies had their own substantial psychometric shortcomings (Ben-Porath, 2012). Introduction of a set of Restructured Clinical (RC) Scales for the MMPI-2 (Tellegen et al., 2003) represented the first step toward remedying the psychometric challenges of the Clinical Scales. Recognizing and addressing the role of Demoralization in MMPI-based testing was a necessary first step in this process.

The Role of Demoralization

Based on his own and others’ work related to the assessment of mood and anxiety, Tellegen (1985) indicated that many self-report measures used in clinical psychology (including the MMPI) contained larger than usual intercorrelations because their scales, in addition to a targeted construct, also assessed a general factor he termed demoralization. Demoralization is a result of “persistent failure to cope with internally or externally induced stress” (Frank, 1974, p. 271) and is defined as nonspecific distress that often takes the form of feeling sad, unhappy, and overwhelmed. Findings from factor analyses of mood terms conducted by Watson and Tellegen (1985) generated the conceptual model depicted in Figure 23.1. Two orthogonal dimensions termed positive and negative affect (later termed positive and negative activation and abbreviated PA and NA, respectively) were identified. On the low end of positive activation were descriptors related to anhedonia, low energy, and a loss of interest (closely associated with major depressive disorder). On the high end of negative activation were descriptors related to anxiety-related psychopathology. This model also included a dimension incorporating shared hedonic features of low PA and high NA, labeled happiness–unhappiness, where on the unhappiness end, descriptors related to demoralization were observed. In the psychotherapy literature, demoralization is hypothesized to be common across all forms of psychopathology and is often what initially motivates individuals to seek treatment (Frank, 1974). It is also a construct observed in behavioral medicine settings as a factor that motivates individuals to seek medical care (Fava et al., 1995; Mangelli et al., 2005).
Restructuring the Clinical Scales

Restructuring the Clinical Scales involved a multistep process. Identifying and removing demoralization was the first step. Using four different databases (for replication purposes), factor analyses of the Clinical Scales guided development of a demoralization measure. The second step in this process was to identify major distinctive core components of each of the Clinical Scales. For example, a factor analysis of Clinical Scale 2 resulted in a two-factor solution: demoralization and lack of positive emotions. In some instances, more than two factors were identified. Demoralization items were removed from the Clinical Scales and were placed on a separate scale labeled Demoralization (RCd). The third step identified items that emerged from Step 2, which were internally coherent and mutually distinct. Items with high correlations with the remaining items identified as reflecting the major distinctive core component of a clinical scale were retained, and any item that met this criterion across multiple scales was deleted. This resulted in 11 nonoverlapping “seed” scales. A twelfth seed scale that represented demoralization was added to ensure demoralization variance would not be reintroduced in the next and final step. The fourth step aimed to derive the final set of RC scales. It involved using the rest of the MMPI-2 item pool to enhance the seed scales and finalize construction on the Restructured Clinical (RC) Scales (Tellegen et al., 2003). The resultant nine RC Scales (one measuring Demoralization and the others assessing a major distinctive core component of the eight original Clinical Scales) were added to the MMPI-2 in 2003 to help clarify Clinical Scale interpretation. Since their addition, a substantial body of empirical research has documented their improved psychometric properties in comparison with the original Clinical Scales (Ben-Porath, 2012).

Figure 23.1 Watson and Tellegen’s (1985) Consensual Model of Mood and Anxiety

Note. Figure adapted from Watson and Tellegen (1985); bolded indicators are associated with mood and anxiety disorders.
Development of Additional Scales

The rationale, methods, and empirical support of the RC scales stimulated further exploration of the MMPI-2 item pool. Contemporary models of psychopathology suggested that it was best conceptualized in a hierarchical fashion (Caspi et al., 2014; Kotov, Chang et al., 2011; Krueger & Markon, 2006). Building on this research, factor analyses of the RC scales demonstrated that the constructs assessed by these measures conform to three broadband dimensions assessed by the MMPI-2-RF Higher-Order Scales: Emotional/Internalizing Dysfunction, Thought Dysfunction, and Behavioral/Externalizing Dysfunction. A total of 23 more narrowly focused Specific Problems Scales and two Interest Scales were also developed to assess facets of the Clinical Scales not represented by the RC Scales, more narrowly focused facets of the RC Scales themselves, or MMPI-2 content domain not represented by the RC Scales. The Personality Psychopathology—Five (PSY-5) scales—measures of dimensional personality disorder traits—were also revised for the MMPI-2-RF (Harkness & McNulty, 1994). Finally, seven validity scales from the MMPI-2 were revised for the MMPI-2-RF and two new validity scales that assess somatic over-reporting (Wygant, Ben-Porath, & Arbisi, 2004) and cognitive over-reporting (Gervais, Ben-Porath, Wygant, & Green, 2007) were added. A summary and description of the 51 scales that make up the MMPI-2-RF (Ben-Porath & Tellegen, 2008/2011, Tellegen & Ben-Porath, 2008/2011) are provided in Table 23.1.

Table 23.1 Scales of the MMPI-2-RF

<table>
<thead>
<tr>
<th>MMPI-2-RF Scale</th>
<th>Place Within the Hierarchy</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Validity Scales:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable Response Inconsistency (VRIN-r)</td>
<td>Non-content-Based Validity Scale</td>
<td>Detects random responding</td>
</tr>
<tr>
<td>True Response Inconsistency (TRIN-r)</td>
<td>Non-content-Based Validity Scale</td>
<td>Detects fixed responding (both true and false)</td>
</tr>
<tr>
<td>Infrequent Responses (F-r)</td>
<td>Over-Reporting Validity Scale</td>
<td>Detects infrequent responses</td>
</tr>
<tr>
<td>Infrequent Psychopathology Responses (Fp-r)</td>
<td>Over-Reporting Validity Scale</td>
<td>Detects infrequent psychopathology responses</td>
</tr>
<tr>
<td>Infrequent Somatic Responses (Fs)</td>
<td>Over-Reporting Validity Scale</td>
<td>Detects infrequent somatic responses</td>
</tr>
<tr>
<td>Symptom Validity (FBS-r)</td>
<td>Over-Reporting Validity Scale</td>
<td>Detects infrequent somatic/cognitive responses</td>
</tr>
<tr>
<td>Response Bias Scale (RBS)</td>
<td>Over-Reporting Validity Scale</td>
<td>Detects infrequent somatic/cognitive responses</td>
</tr>
<tr>
<td>Uncommon Virtues (L-r)</td>
<td>Under-Reporting Validity Scale</td>
<td>Detects self-virtuous presentations</td>
</tr>
<tr>
<td>Adjustment Validity (K-r)</td>
<td>Under-Reporting Validity Scale</td>
<td>Detects above average level of adjustment</td>
</tr>
<tr>
<td><strong>Emotional/Internalizing Domain:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional/Internalizing Dysfunction (EID)</td>
<td>Higher-Order Scale</td>
<td>Broad range of emotional symptoms and difficulties</td>
</tr>
<tr>
<td>Demoralization (RCd)</td>
<td>Restructured Clinical Scale</td>
<td>Unhappiness, sadness, dissatisfied with current life circumstances</td>
</tr>
<tr>
<td>Suicidal/Death Ideation (SUI)</td>
<td>Specific Problems Scale</td>
<td>Current/past suicidal ideation/behaviors</td>
</tr>
<tr>
<td>Helplessness/Hopelessness (HLP)</td>
<td>Specific Problems Scale</td>
<td>Hopelessness and pessimism</td>
</tr>
<tr>
<td>Self-Doubt (SFD)</td>
<td>Specific Problems Scale</td>
<td>Self-doubt, lacking confidence, feeling useless</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>MMPI-2-RF Scale</th>
<th>Place Within the Hierarchy</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inefficacy (NFC)</td>
<td>Specific Problems Scale</td>
<td>Passive, indecisive, inefficacious</td>
</tr>
<tr>
<td>Low Positive Emotions (RC2)</td>
<td>Restructured Clinical Scale</td>
<td>Anhedonia, lack of interest, displays vegetative symptoms of depression</td>
</tr>
<tr>
<td>Introversion/Low Positive Emotionality—Revised (INTR-r)</td>
<td>PSY-5 Scale</td>
<td>Lacks positive emotional experiences, avoids social situations</td>
</tr>
<tr>
<td>Dysfunctional Negative Emotions (RC7)</td>
<td>Restructured Clinical Scales</td>
<td>Anxiety, anger, fear</td>
</tr>
<tr>
<td>Stress/Worry (STW)</td>
<td>Specific Problems Scale</td>
<td>Stress-reactive, worry-prone, obsessive rumination</td>
</tr>
<tr>
<td>Anxiety (AXY)</td>
<td>Specific Problems Scale</td>
<td>Feeling anxious, having nightmares</td>
</tr>
<tr>
<td>Anger Proneness (ANP)</td>
<td>Specific Problems Scale</td>
<td>Anger, irritability, impatient</td>
</tr>
<tr>
<td>Behavior-Restricting Fears (BRF)</td>
<td>Specific Problems Scale</td>
<td>Fears that restrict normal activities</td>
</tr>
<tr>
<td>Multiple Specific Fears (MSF)</td>
<td>Specific Problems Scale</td>
<td>Specific fears of certain animals and acts of nature</td>
</tr>
<tr>
<td>Negative Emotionality/Neuroticism—Revised (NEGE-r)</td>
<td>PSY-5</td>
<td>Self-critical, guilt-prone, anxious</td>
</tr>
<tr>
<td><strong>Thought Dysfunction Domain:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thought Dysfunction (THD)</td>
<td>Higher-Order Scale</td>
<td>Symptoms and difficulties associated with disordered thinking</td>
</tr>
<tr>
<td>Psychoticism-Revised (PSYC-r)</td>
<td>PSY-5 Scale</td>
<td>Unrealistic thinking, impaired reality testing</td>
</tr>
<tr>
<td>Ideas of Persecution (RC6)</td>
<td>Restructured Clinical Scale</td>
<td>Belief that others are seeking to harm them, paranoid delusions</td>
</tr>
<tr>
<td>Aberrant Experiences (RC8)</td>
<td>Restructured Clinical Scale</td>
<td>Unusual thought and perceptual processes</td>
</tr>
<tr>
<td><strong>Behavioral Externalizing Domain:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioral Externalizing Dysfunction (BXD)</td>
<td>Higher-Order Scale</td>
<td>Behaviors and difficulties associated with undercontrolled behaviors</td>
</tr>
<tr>
<td>Antisocial Behaviors (RC4)</td>
<td>Restructured Clinical Scale</td>
<td>Fails to conform to social norms, impulsive, acts out when bored</td>
</tr>
<tr>
<td>Disconstraint-Revised (DISC-r)</td>
<td>PSY-5 Scale</td>
<td>Impulsivity, acting out</td>
</tr>
<tr>
<td>Juvenile Conduct Problems (JCP)</td>
<td>Specific Problems Scale</td>
<td>History of problematic behaviors at school</td>
</tr>
<tr>
<td>Substance Abuse (SUB)</td>
<td>Specific Problems Scale</td>
<td>Significant past and current substance abuse</td>
</tr>
<tr>
<td>Hypomanic Activation (RC9)</td>
<td>Restructured Clinical Scale</td>
<td>Sensation-seeking, risk-taking, excitability</td>
</tr>
<tr>
<td>Aggressiveness-Revised (AGGR-r)</td>
<td>PSY-5 Scale</td>
<td>Interpersonal aggression and assertiveness</td>
</tr>
<tr>
<td>Aggression (AGG)</td>
<td>Specific Problems Scale</td>
<td>Physically aggressive, violent behavior, losing control</td>
</tr>
<tr>
<td>Activation (ACT)</td>
<td>Specific Problems Scale</td>
<td>Heightened excitation and energy levels</td>
</tr>
<tr>
<td><strong>MMPI-2-RF Scale</strong></td>
<td><strong>Place Within the Hierarchy</strong></td>
<td><strong>Content</strong></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Somatic/Cognitive Domain:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic Complaints (RC1)</td>
<td>Restructured Clinical Scale</td>
<td>Multiple somatic problems that may have psychological components underlying them</td>
</tr>
<tr>
<td>Malaise (MLS)</td>
<td>Specific Problems Scale</td>
<td>Poor health and feeling weak/tired</td>
</tr>
<tr>
<td>Gastrointestinal Complaints (GIC)</td>
<td>Specific Problems Scale</td>
<td>Poor appetite, nausea, vomiting</td>
</tr>
<tr>
<td>Head Pain Complaints (HPC)</td>
<td>Specific Problems Scale</td>
<td>Recurrent headaches, neck pain, and developing head pain when upset</td>
</tr>
<tr>
<td>Neurological Complaints (NUC)</td>
<td>Specific Problems Scale</td>
<td>Dizziness, loss of balance, numbness, weakness, paralysis</td>
</tr>
<tr>
<td>Cognitive Complaints (COG)</td>
<td>Specific Problems Scale</td>
<td>Subjective reports of memory/concentration problems, intellectual difficulties, and confusion</td>
</tr>
<tr>
<td><strong>Interpersonal Domain:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cynicism (RC3)</td>
<td>Restructured Clinical Scale</td>
<td>Distrusts others, believes others look out only for their own interests</td>
</tr>
<tr>
<td>Family Problems (FML)</td>
<td>Specific Problems Scale</td>
<td>Conflictual family relationships, lack of support of their families Unassertive and submissive</td>
</tr>
<tr>
<td>Interpersonal Passivity (IPP)</td>
<td>Specific Problems Scale</td>
<td>Does not enjoy social events, avoids social situations</td>
</tr>
<tr>
<td>Social Avoidance (SAV)</td>
<td>Specific Problems Scale</td>
<td>Shy, easily embarrassed, uncomfortable around others</td>
</tr>
<tr>
<td>Shyness (SHY)</td>
<td>Specific Problems Scale</td>
<td>Dislikes being around others, prefers to be alone</td>
</tr>
<tr>
<td>Disaffiliativeness (DSF)</td>
<td>Specific Problems Scale</td>
<td></td>
</tr>
<tr>
<td><strong>Interest Domain:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aesthetic-Literary Interests (AES)</td>
<td>Interest Scale</td>
<td>Interests associated with writing, music, and theater</td>
</tr>
<tr>
<td>Mechanical-Physical Interests (MEC)</td>
<td>Interest Scale</td>
<td>Interests associated with fixing/building things, sports, outdoor activities</td>
</tr>
</tbody>
</table>

**THE MMPI-2-RF**

The 338-item MMPI-2-RF is intended to provide a thorough and efficient assessment of clinically relevant factors in well under an hour. As described earlier, the test is a broadband instrument that assesses protocol validity across nine sets of Validity Scales as well as psychopathology/personality across 42 Substantive Scales (Table 23.1). The Substantive Scales are theoretically grounded. For example, the Emotional/Internalizing Dysfunction domain scales are theoretically and empirically derived measures in accordance with Watson and Tellegen’s (1985) model of mood and anxiety. Sellbom, Ben-Porath, and Bagby (2008) demonstrated that a higher-order Fear and Distress model of
internalizing psychopathology (Watson, 2005) can be elaborated by using the RC scales to differentiate between Demoralization (RCd), Low Positive Emotions (RC2), and Dysfunctional Negative Emotions (RC7). Demoralization (RCd) was the best predictor of distress disorders whereas Dysfunctional Negative Emotions (RC7) was the best predictors of fear disorders. The researchers also supported that Low Positive Emotions (RC2) was a unique predictor of major depressive disorder and social phobia. The scales of the MMPI-2-RF and their relevance in medical settings are briefly reviewed next.

Validity Scales

An important component of the MMPI-2-RF is its nine validity scales that assess for protocol validity (i.e., extent to which scores on the test are valid and interpretable). A protocol can be invalid due to non-content-based or content-based response patterns. Non-content-based responding occurs when an individual’s responses are not based on the content on the items, such as nonresponding, random responding, or fixed responding. The MMPI-2-RF contains one indicator (Cannot Say) and two scales that assess for non-content-based responding [Variable Response Inconsistency (VRIN-r) & Fixed Response Inconsistency (TRIN-r)]. Content-based invalid responding, on the other hand, refers to a pattern in which an individual’s responses are invalid due to over-reporting or under-reporting. There are five validity scales that assess over-reporting [Infrequent Responses (F-r), Infrequent Psychopathology Responses (Fp-r), Infrequent Somatic Responses (Fs), Symptom Validity (FBS-r), & Response Bias (RBS)] and two validity scales that assess under-reporting [Uncommon Virtues (L-r) & Adjustment Validity (K-r)]. A list of all validity scales and a description of what they assess for are listed in Table 23.1.

A patient in medical settings can produce an elevation on validity scales for a number of reasons. For instance, a candidate for bariatric surgery may under-report symptoms of psychopathology in an attempt to appear well-adjusted and be cleared for surgery. Marek, Tarescavage, Ben-Porath, Ashton, Rish, and Heinberg (2015) reported that up to 20.68% of bariatric surgery candidates evidenced some form of under-reporting on the MMPI-2-RF prior to surgery. On the other hand, a patient who suffered a minor head injury at work may over-report somatic and cognitive symptoms to aid their attempts to obtain workers’ compensation or disability benefits.

The MMPI-2-RF validity scales have been shown to have utility in detecting these over-reporting response patterns. Youngjohn, Wershba, Stevenson, Sturgeon, and Thomas (2011) reported that Symptom Validity (FBS-r) scale differentiated traumatic brain injury (TBI) litigants who passed performance validity tests from those who failed. Harp, Jasinski, Shandera-Ochsner, Mason, and Berry (2011) found that Infrequent Psychopathology Responses (Fp-r) scores best detected malingered Attention Deficit Hyperactivity Disorder (ADHD) by comparing college students who were instructed to feign with ones who actually had ADHD. In a sample of chronic pain patients, Rogers, Gillard, Berry, and Granacher (2011) found that scores on Infrequent Responses (F-r) and Infrequent Psychopathology Responses (Fp-r) differentiated patients who were deemed feigning psychopathology whereas scores on the Symptom Validity (FBS-r) and Response Bias (RBS) scale scores were associated with scoring below chance on cognitive effort tests. In a sample of compensation-seeking evaluations for pain- and neurocognitive-related disorders, Wygant et al. (2011) found that the Infrequent Responses (F-r) and Response Bias (RBS) scales yielded the largest effect sizes in differentiating probable/definite malingerers from incentive-seeking-only individuals. Finally, Gervais and colleagues (2007) found that scores on Response Bias (RBS) were associated with failure on the Word Memory Test (Green, 2004) and the Medical Symptom Validity Test (Green, 2004). In summary, a number of studies have demonstrated that the Validity Scales of the MMPI-2-RF effectively identify feigning or minimizing of symptoms in medical settings.
Substantive Scales

As previously noted, the 42 substantive scales of the test are organized hierarchically across five domains: Emotional/Internalizing Dysfunction, Thought Dysfunction, Behavioral/Externalizing Dysfunction, Somatic/Cognitive, and Interpersonal Problems. The test also includes two interest scales that may aid in helping clients find activities they enjoy doing during treatment planning (e.g., identifying an enjoyable activity when considering behavioral activation for the treatment of depressive disorders). A list of scale names and descriptions can be found in Table 23.1. The MMPI-2-RF manual also lists diagnostic (and treatment) considerations for most of the substantive scales of the inventory. MMPI-2-RF Manual for Administration, Scoring, and Interpretation (Ben-Porath & Tellegen, 2008/2011) scale scores can be used for differential diagnostic and risk assessment purposes to help guide patient care, including when a higher level of care or evaluation for the need of a psychotropic medication should be considered. This section briefly reviews the Substantive Scales by domain and broadly discusses the constructs they assess across multiple settings, including behavioral medicine settings.

Emotional/Internalizing Dysfunction

Beginning with the Internalizing domain, the higher-order scale Emotional/Internalizing Dysfunction (EID) assesses a broad range of emotional symptoms and difficulties. The next level of the hierarchy are the RC scales, Demoralization (RCd), Low Positive Emotions (RC2), and Dysfunctional Negative Emotions (RC7) that assess constructs congruent with Tellegen’s (1985) structure of mood and anxiety (Figure 23.1). The next level of the hierarchy contains the Internalizing Specific Problems Scales and the PSY-5 scales. The RCd scale has four associated Internalizing Specific Problems Scales [Suicidal/Death Ideation (SUI), Hopelessness/Helplessness (HLP), Self-Doubt (SFD), and Inefficacy (NFC)]. There are five Internalizing Specific Problems Scales associated with Dysfunctional Negative Emotions (RC7) [Stress/Worry (STW), Anxiety (AXY), Anger-Proneness (ANP), Behavior-Restriction Fear (BRF), and Multiple Specific Fears (MSF)] as well as one of the PSY-5 scales, Negative Emotionality/Neuroticism-Revised (NEGE-r). RC2 has one lower order scale, Introversion/Low Positive Emotionality (INTR-r) of the PSY-5.

The Emotional/Internalizing Dysfunction (EID) domain scales are useful for parsing out the patients’ specific problems, and they can provide further support for differential diagnosis and treatment planning. For example, elevated scores on AXY indicate that the test taker is likely constantly anxious, frightened by something nearly every day, and has frequent nightmares (Ben-Porath & Tellegen, 2008/2011). Indeed, scores on Anxiety AXY are predictive of all posttraumatic stress disorder symptom clusters (Sellbom, Lee, Ben-Porath, Arbisi, & Gervais, 2012) and identify patients who should be evaluated for the need of anxiolytic medications and have treatment of anxiety as a target for intervention. High scores on RC2 are associated with anhedonia-related symptoms of major depression (thus warranting an evaluation for an antidepressant medication) whereas high scores on RCd point to more global distress-oriented disorders and symptoms (Sellbom et al., 2008). Scores on the SUI scale have been found to incrementally predict future suicidal behaviors within one year of the evaluation after controlling for information obtained during the clinical interview and scores on RCd and RC2 in other settings (Glassmire, Tarescavage, Burchett, Martinez, & Gomez, 2015). Relatedly, the higher the T-score on SUI, the higher the relative risk is of suicidal behavior (Gottfried, Bodell, Carbonell, & Joiner, 2014). Indeed, the EID scales yield good validity and utility in assessing for mood and anxiety-related constructs and disorders in various medical settings (cf. Block, Ben-Porath, & Marek, 2013; Marek, Ben-Porath, Windover et al., 2013; Tarescavage et al., 2013).
Behavioral/Externalizing Dysfunction
The Externalizing domain contains one higher-order scale, Behavioral/Externalizing Dysfunction (BXD), which assesses a broad range of externalizing symptoms and behaviors including impulsivity, hypomanic activation, and antisocial behaviors. RC Scales Antisocial Behaviors (RC4) and Hypomanic Activation (RC9) can aid in differential interpretation of BXD. RC4 has two Externalizing Specific Problems Scales associated with it [Juvenile Conduct Problems (JCP) and Substance Abuse (SUB)] whereas Hypomanic Activation (RC9) has two lower order scales [Aggression (AGG) & Activation (ACT)] as well as two PSY-5 scales [Aggressiveness—Revised (AGGR-r) & Disconstraint (DISC-r)] associated with it to better aid in interpretation.

In terms of differentiating major depressive disorder from bipolar disorder, elevations on RC9 and its Specific Problems scales, most notably on ACT—a scale assessing episodes of heightened excitation and energy levels—is the best predictor (Watson, Quilty, & Bagby, 2011). Scores on RC4 and its Specific Problems scales are associated with social nonconformity, alcohol and substance abuse, and high disinhibition across multiple settings (see Ben-Porath, 2012 for a review of the literature), including behavioral medical settings (cf. Block, Ohnmeiss, Ben-Porath, & Burchett, 2011; Marek, Ben-Porath, Windover et al., 2013; Tarescavage et al., 2013, Wygant et al., 2007).

Thought Dysfunction
The higher-order scale (Thought Dysfunction [THD]) assesses a broad range of symptoms associated with disordered thinking, whereas its associated RC scales, Ideas of Persecution (RC6) and Aberrant Experiences (RC8), aid in differential interpretation. In addition, the Psychoticism (PSYC-r) scale of the PSY-5 is another lower order scale within the thought dysfunction domain. Associations between RC6 and ideas of reference, paranoia, interpersonal mistrust, alienation, and blame externalization have been evidenced in clinical (cf. Arbisi, Sellbom, & Ben-Porath, 2008; Handel & Archer, 2008) and nonclinical (e.g., Sellbom & Ben-Porath, 2005) samples. Scores on the RC8 scale are associated with presence of hallucinations, nonparanoid delusions, unusual thought content, conceptual disorganization, and dissociative experiences (cf. Arbisi, Sellbom, & Ben-Porath, 2008; Handel & Archer, 2008). Although the literature is not as extensive in regards to the utility of the Thought Dysfunction (THD) scales in medical settings, Pona and colleagues (2016) found that higher presurgical scores on RC6 predicted body image dissatisfaction as soon as three months following bariatric surgery.

Somatic/Cognitive
The Somatic/Cognitive domain of the test contains one RC scale, Somatic Complaints (RC1), and five Somatic/Cognitive Specific Problems Scales [Malaise (MLS), Gastrointestinal Complaints (GIC), Head Pain Complaints (HPC), Neurological Complaints (NUC), and Cognitive Complaints (COG)]. Many of these scales reflect various forms, indexes, and manifestations of somatoform psychopathology. For example, Aragona, Tarsitani, De Nitto, and Inghilleri (2008) reported that a combination of items on RC1 and MLS differentiated patients with pain disorder from patients experiencing pain from an identified somatic origin. Locke and colleagues (2010) found that RC1 differentiated patients with epileptic seizures from those with nonepileptic seizures with a sensitivity of .76 and a specificity of .60. In addition, RC1 also contributed to this differentiation above and beyond demographic information and patients’ medical history. Marek, Ben-Porath, Merrell, Ashton, and Heinberg (2014) found that patients who report a number of medical complaints as soon as one month after bariatric surgery, including dehydration, excessive vomiting, and pain, also scored higher on RC1, HPC, and NUC at the time of their presurgical evaluation. Tellegen and Ben-Porath (2008/2011) also report correlations across various settings that indicate that higher scores on the Somatic/Cognitive Substantive
Scales are associated with preoccupation with poor health, complaining of sleep disturbances, fatigue and low energy, and higher scores on somatization measures.

**Interpersonal**
The Interpersonal scales of the MMPI-2-RF are composed of one RC scale, Cynicism (RC3), and five Interpersonal Specific Problems Scales [Family Problems (FML), Interpersonal Passivity (IPP), Social Avoidance (SAV), Shyness (SHY), & Disaffiliativeness (DSF)]. High scores on RC3 are associated with hostility, anger, low trust, and negative beliefs about others (Burchett & Ben-Porath, 2010; Forbey & Ben-Porath, 2007; Tellegen/Ben-Porath, 2008/2011). The construct of cynicism measured by the RC3 scale of the MMPI-2-RF is rooted in behavioral medicine. Cynicism and its relation to coronary heart disease (CHD) has been studied for decades (cf. Almada et al., 1991; Barefoot et al., 1989; Boyle et al., 2004; Costa et al., 1986), largely related to the MMPI-2 Cook-Medley Hostility scale (ACM), which contained 10 out of the 12 items on the RC3 scale of the MMPI-2-RF. Barefoot, Larsen, von der Leith, and Schroll (1995) administered the MMPI in a sample of Danish men and women. After 27 years, these researchers found that higher scores on ACM were associated with an increased risk for myocardial infarction, even after controlling for other medical risk factors. In another study, Boyle and colleagues (2004) found that ACM scores predicted CHD mortality after controlling for disease severity. Although these studies have yet to be replicated with RC3, past research implies utility of using RC3 in medical settings.

Scores on FML and IPP describe familial discord and submissiveness, respectively. SAV, SHY, and DSF scales describe various types of social isolation and avoidant behaviors. In regard to their utility in behavioral medicine settings, scores on FML and SAV have been predictive of suboptimal spine surgery and spinal cord stimulator outcomes (Block, Marek, Ben-Porath, & Kukal. 2015; Marek, Block, & Ben-Porath, 2015) as well as psychological distress 3 months after having bariatric surgery (Marek, Ben-Porath, Merrell, Ashton, & Heinberg, 2014).

**Summary**
Scores on the MMPI-2-RF Substantive Scales are associated with a number of psychiatric diagnoses and behaviors relevant to primary care and behavioral medicine settings. It is important to note that scale scores should not be used as sole indicators of a diagnosis (or high-risk behavior). Rather, test findings should be integrated with other sources including other test scores, interview data, and collateral information to make a diagnosis or decide to advance a patient to a higher level of care. For example, a moderate T-score just above 65 (i.e., > 65T) on RC6 identifies a patient who believes that others seek to harm him or her. Whether this reflects psychopathology will depend, in part, on their actual life circumstances, which may include actual persecution. Integrating psychological test results with additional information obtained via a clinical interview, medical charts, and other measures is considered standard practice.

**Psychometrics**
The MMPI-2-RF Technical Manual (Tellegen & Ben-Porath, 2008/2011) provides comprehensive information about scale score reliability and validity across settings. Specifically, reliability estimates and associated standard errors of measurement for the 51 scales are reported for both the normative and clinical samples. Appendix A of the manual lists over 50,000 correlations between MMPI-2-RF substantive scale scores and external criteria for various samples, including medical and medico–legal ones. These data indicate that MMPI-2-RF scores have adequate reliability and they provide evidence of convergent and discriminant validity across settings. These findings generalize to medical samples beyond those included in the Technical Manual. For example, McCord and Drerup (2011) compared MMPI-2 Clinical and RC Scale scores in a mixed, outpatient neuropsychological and chronic pain
sample. Whereas both sets of scales achieved 85% classification convergence with DSM-IV-TR diagnoses, the result was achieved by using just two of the nine RC Scales versus using seven of the eight Clinical Scales, indicating that RC scale scores have improved discriminant validity when compared with the Clinical Scales. Wygant et al. (2007) compared the Clinical and RC scale scores in a bariatric surgery–seeking sample. They found that the RC scales were more internally consistent and evidenced better convergent and discriminant properties in predicting relevant behavioral and psychological factors relevant to preoperative bariatric psychological evaluations. Tarescavage, Wygant, Boutacoff, and Ben-Porath (2013) and Marek et al. (2013) extended these findings in a bariatric surgery sample using all of the MMPI-2-RF scales and drew similar conclusions. Improved discriminant validity findings (for the latter) when comparing the Clinical and RC scales were also reported for a sample of patients with seizure disorders (Bowden, White, Simpson, & Ben-Porath, 2014).

Some clinicians have been reluctant to use the MMPI-2 with older adults and patients with dementia, owing to the number of items that need to be completed. Carone and Ben-Porath (2014) published a case study of a 65-year-old woman with severe memory impairments and deficits in executive functioning due to dementia. She was referred for a neuropsychological evaluation when she scored a 21/30 on the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005). The MMPI-2-RF was administered immediately after she took cognitive tests as part of her neuropsychological evaluation. Carone and Ben-Porath (2014) reported:

Cognitive testing showed a significant decline in general intellectual functioning, poor learning of new information despite repetition rapid forgetting of information with no information recalled on delayed recall memory tasks, severely impaired recognition memory with poor encoding of information, and significant executive dysfunction including perseverations and intrusions.

(p. 1022)

On the MMPI-2-RF, the patient produced a valid and interpretable protocol. The patient’s only elevation was on the HPC scale, which was consistent with her injuries from a motor vehicle accident 18 months prior to the evaluation. In summary, dementia did not preclude reliable responding on the MMPI-2-RF. Studies also indicate that MMPI-2-RF scale scores do not evidence differential validity/test bias when used in bariatric surgery settings (Marek, Ben-Porath, Sellbom, McNulty, & Heinberg, 2015). Specifically, no to very small effect sizes were evidenced in differential validity analyses with external criteria. In sum, research has demonstrated good psychometric properties of MMPI-2-RF scores in a variety of medical settings.

ADMINISTRATION, SCORING, AND INTERPRETATION OF THE MMPI-2-RF

This next section briefly reviews standard procedures for administration and scoring of the MMPI-2-RF. An interpretation strategy and the importance of considering comparison groups when interpreting MMPI-2-RF protocols are also discussed.

Administration

The MMPI-2-RF should be administered by a qualified test user or a technician trained and working under the supervision of a qualified test user. The test can be administered via pencil and
paper, on a computer, or via a tablet. It should be administered in a quiet, comfortable environment. The test taker should be able to see, read, and comprehend the test items. The average reading difficulty for the MMPI-2-RF varies based on the method used to derive reading level. The Flesch-Kincaid index derived from Microsoft Word using all 338 items of the MMPI-2-RF yields a score of 4.5 (indicating a 4th to 5th grade reading level). Other methods suggest the reading level is somewhere between the 4th and 7th grade reading level (Dahlstrom et al., 1994; Schinka & Borum, 1993). An item-by-item analysis indicates that some items require a higher reading level. Therefore, an overall value for reading level does not guarantee that a respondent will be able to read every item of the test. If an individual falls below the requisite reading level, an audio version of the test is available. Regardless of administration type, a set of standard instructions is included. The test taker should also be supervised to both protect the integrity of the test and the test scores. Pencil-and-paper administration time is approximately 35–50 minutes whereas computerized administration with software licensed by the publisher typically requires of 25–35 minutes (Ben-Porath & Tellegen, 2011).

**Scoring**

The MMPI-2-RF can be scored by hand, mailed to the publisher for scoring, or scored by computer using the licensed software. Computerized scoring is faster and more efficient than pencil-and-paper administration and/or hand scoring. If administered via pencil-and-paper format, answer sheets can be scanned and scored using the licensed software. Computerized scoring offers a number of advantages over hand scoring. It can produce a Score Report or Interpretative Report. Both reports provide raw scores and T-scores for all 51 scales of the instrument, the percentage of items scorable on each scale, the number of items the individual left unscorable (and which scales those items were associated with), and item responses from a subset of scales that may require immediate clinical attention (e.g., Suicidal Death/Ideation Scale). The software can also plot the test taker’s scores against results produced by available comparison groups. For example, the scores of a test taker evaluated for bariatric surgery can be compared with those of other bariatric surgery candidates to provide an indication of how the individual compares with those assessed under similar circumstances who have taken the MMPI-2-RF. A clinically focused Interpretive Report generates a narrative interpretation of the test results (Ben-Porath & Tellegen, 2008). The report is fully annotated in that the source (i.e., the scale score[s] generating) of every statement provided is an indication of whether the statement is based on empirical correlates or the content of items on a scale, or is an inference of the report authors. Citations to published literature are provided for statements identified as being based on empirical correlates.

**Normative Data/Comparison Groups**

Computer-generated Score and Interpretative Reports can include comparison group data reported in the Technical Manual (Tellegen & Ben-Porath, 2008/2011). These data are currently available for a number samples representing settings and types of assessments in which the MMPI-2-RF is used. These include mental health, medical, forensic, and personnel screening settings. In terms of medical settings, the MMPI-2-RF Technical Manual (Tellegen & Ben-Porath, 2008/2011) includes comparison group data for spine surgery and bariatric surgery candidates as well as patients entering a chronic pain treatment program. Demographic characteristics of these comparison groups are reported in Table 23.2.
When to Use the MMPI-2-RF

The MMPI-2-RF provides a comprehensive assessment of a broad range of psychological constructs relevant to psychodiagnosis and treatment planning. Consequently, it can be administered as a single measure or as part of a battery of tests. The instrument can assist in assessing for risk that may be difficult to ascertain in a clinical interview. Such risk factors may include thought disorders, suicidal ideation, impulsivity, and interpersonal problems. Some patients may feel more comfortable disclosing sensitive information via the MMPI-2-RF than they would through a clinical interview when rapport is first being established. Because it includes comprehensive, well-validated validity scales, the MMPI-2-RF can also be used if it is suspected that the patient is feigning his/her symptoms or is engaging in under-reporting. Forbey, Lee, Ben-Porath, Arbisi, and Gartland (2013) report data showing that if a test taker evidences an over- or under-reporting response style on the MMPI-2-RF, they engage in a similar response style across other psychological instruments. Crighton, Tarescavage, Gervais, and Ben-Porath (2015) showed that over-reporting detected with the MMPI-2-RF generalizes to other measures administered in a medico-legal setting.

Interpretation Strategy

The MMPI-2-RF Manual for Administration, Scoring, and Interpretation provides standard interpretive guidelines intended to facilitate inter-interpreter reliability. The guidelines include an overall interpretative strategy and scale-by-scale interpretation statements. Test interpreters should first examine the number of unscorable (owing to being unanswered or answered both “true” and “false”) items. If any are found, the percent scorable responses should be examined for each scale. If a scale has less than 90% scorable responses, a nonelevated score on that scale is uninterpretable; however, elevated scores can be interpreted with the understanding that they may underestimate the magnitude or extent of the problems assessed by that scale.

The next step is to assess scores on the remaining validity indicators. Inconsistent responding should be ruled out next. Moderate elevations on VRIN-r (70T-79T) suggest there is some evidence of variable response inconsistency and that scores on the remaining validity and substantive scales should be interpreted with caution. A score of 80T or higher indicates that the protocol should not be
interpreted. Possible reasons for elevations on VRIN-r include an uncooperative test-taking approach, intentional random responding, reading or language difficulties, cognitive impairment, or errors in recording responses.

Moderate elevations on TRIN-r (70T-79T) indicate some evidence of fixed, content-inconsistent responding (e.g., most items keyed in the true or false directions). A score of 80T or higher indicates that the rest of the protocol should not be interpreted. TRIN-r elevations are indicative of a noncooperative test-taking approach. Moderate elevations on the over-reporting validity scales raise the possibility of exaggeration or over-reporting of symptoms, but do not invalidate the protocol. For example, if a patient scored 85T on the Fs scale, this may reflect the presence of significant or multiple medical conditions or exaggeration, provided that inconsistent responding had been ruled out. In such a case, the somatic and cognitive scales of the test should be interpreted with some caution as their scores are likely inflated. On the other hand, if a patient scores above 100T on the Infrequent Somatic Response Scale (Fs), their reported symptoms are uncommon even in individuals with severe medical problems who report credible symptoms.

A score of 120T on F-r or >100 on F-p indicates that scores on the substantive scales should not be interpreted. On the one hand, elevated scores on the under-reporting scales are associated with deflated substantive scale scores. If a patient engages in under-reporting her/his substantive scale scores are likely to be underestimates of problems and should be interpreted with caution. The absence of substantive scale score elevations should not be interpreted as an indication of the absence of the problem(s) it is designed to assist under these circumstances.

Interpretation of the substantive scale scores should only be done after validity scale interpretation is completed and validity scores do not indicate that the protocol is invalid. As mentioned earlier, the substantive scale scores of the MMPI-2-RF are organized in a hierarchical structure across five domains: Emotional/Internalizing Dysfunction, Thought Dysfunction, Behavioral/Externalizing Dysfunction, Somatic/Cognitive Complaints, and Interpersonal. The interpretive guidelines indicate that substantive scale interpretation begins when a scale score is at 65T and above, although it is important to recognize that this is simply a heuristic guideline. Interpretation begins with the highest Higher-Order Scale score and its associated RC Specific Problems scales and PSY-5 Scales scores. For example, if a patient scored 80T on the EID scales and 65T on the BXD scale, interpretation would begin with the EID scale score. From there, the interpreter would examine scores on the associated RC scales, including RCd, RC2, and RC7 and determine whether any are at or above 65T. Again, if multiple scales are above 65T, the clinician would interpret the highest first, and then move to its associated Specific Problems Scales. Once that is complete, the clinician would turn to the next highest RC scale within the hierarchy and its associated Specific Problems Scales, and so forth. Once all scales within EID are interpreted, the clinician moves on to the next highest Higher-Order Scale elevation. Once the elevated Higher-Order Scales are interpreted, any remaining RC scale elevations (and their associated SP scale scores) would be turned to next, followed by any remaining SP Scales. The Interpersonal scales and RC3 would be interpreted next, followed by the Interest Scales. Table 23.1 lists all MMPI-2-RF scales organized in a hierarchical fashion. (Note that Table 23.1 also presents MMPI-2-RF data associated with the case study presented near the end of this chapter.)

USING THE MMPI-2-RF IN BEHAVIORAL MEDICINE SETTINGS

This section begins with some general considerations for using the MMPI-2-RF in primary care and behavioral medicine settings and then turns to detailed discussion of some specific applications. When discussing specific applications, research demonstrating the utility of using the MMPI-2-RF in
various medical settings will also be reviewed. Lastly, an example of interpreting a MMPI-2-RF protocol administered as part of a bariatric surgery evaluation is presented.

**General Considerations**

As reviewed earlier, MMPI-2-RF scale scores have been found to be reliable and valid in a variety of medical settings (e.g., presurgical, chronic pain, epilepsy monitoring). Also discussed earlier is the availability of standard interpretive guidelines, which enhance cross-interpreter reliability. The absence of such guidelines for the MMPI-2 resulted in sometimes quite variable interpretations of the same set of test scores. The standard interpretive guidelines include diagnostic considerations to aid with rule-outs or provide additional diagnostic information that may not otherwise have been disclosed during an intake or interview. Treatment considerations, highlighting potential targets for intervention, needs for follow-up evaluation, and potential process issues are also included in these guidelines. The MMPI-2-RF Validity Scales provide information on potential threats to protocol validity, including the possibility of over-reporting in the context of external or internal incentives to do so. From a practical perspective, the inventory can be completed in under an hour and administered via tablets, allowing for bedside or exam room administration (waiting room administration is not recommended owing to possible distractions).

**Presurgical Psychological Evaluations**

Presurgical psychological evaluations are widely implemented across various medical settings, including bariatric, spine, organ transplant, and reconstructive surgery settings, among others. These assessments are aimed to identify untreated psychopathology, behavioral problems including substance use, and whether patients can appropriately adhere to the strict postsurgical guidelines required for successful outcomes. Presurgical evaluations allow the psychologist to assess whether a patient can benefit from pre- or postsurgical interventions in an effort to help outcomes. Considerable evidence has already accumulated concerning the utility of the MMPI-2-RF in bariatric surgery, spine surgery, and pain control procedures. The following section reviews empirical findings of using the MMPI-2-RF as part of the presurgical psychological evaluations of bariatric, spine, and spinal cord stimulator patients.

**Bariatric Surgery Settings**

Psychopathology is more prevalent in bariatric surgery clinics than it is observed in the general population (Kalarchian et al., 2007; Mitchell et al., 2012). Despite the high prevalence of psychopathology in this setting, psychopathology’s associations with diminished outcomes in this population have yielded inconsistent results. The inconsistent results are likely explained by a number of methodological factors and how psychopathology is conceptualized and measured (Marek, Ben-Porath, & Heinberg, 2016). Nonetheless, three bariatric surgery societies recommend that psychologists use objective psychometric testing as part of their evaluation (Mechanick et al., 2013). About a decade ago, the MMPI-2 was one of the most widely used psychological instruments in bariatric surgery settings (Fabricatore, Crerand, Wadden, Sarwer, & Krasucki, 2006; Walfish, Vance, & Fabricatore, 2007). Although up-to-date survey statistics are not available, Marek, Heinberg, Lavery, Rish, and Ashton (2016) reviewed recent literature on a number of psychological assessment instruments currently being used in bariatric surgery settings. The authors concluded that MMPI-2-RF test scores had the strongest psychometric properties in this setting, and that use of the test was supported by the most comprehensive literature.
MMPI-2-RF scores have been reported to have good reliability in bariatric surgery candidates (Tarescavage, Wygant et al., 2013). Using a sample of 732 bariatric surgery patients, Tarescavage, Wygant, and colleagues (2013) reported good internal consistency coefficients for a majority of MMPI-2-RF scales. Adequate internal consistencies were observed on relatively short SP scales (e.g., SUI contains only five items). Tarescavage, Wygant, and colleagues (2013) also reported convergent and discriminant validity with external criteria obtained via a record review. By and large, good validity coefficients were observed when correlating MMPI-2-RF scale scores with external criteria. Using a larger sample of 1,025 bariatric surgery patients from a different clinic, Marek, Ben-Porath, Windover, and colleagues (2013) examined associations between the MMPI-2-RF substantive scale scores and data obtained from patients’ medical records and psychological diagnoses obtained through a semi-structured clinical interview. The authors reported good convergent and discriminant validity coefficients. In addition, the authors reported means and standard deviations of MMPI-2-RF scores in this sample and concluded that test scores were similar to those of the bariatric surgery comparison group outlined in the MMPI-2-RF Technical Manual (Tellegen & Ben-Porath, 2008/2011).

Incremental validity of the MMPI-2-RF scores in predicting adherence to postsurgical appointments and short-term weight loss outcomes has also been evidenced (Marek, Tarescavage et al., 2015). In a study of 498 patients who underwent a Roux-en-Y Gastric Bypass, higher scores on a number of the BXD scales (after controlling for demographic factors and presurgical body mass index [BMI]) predicted suboptimal weight loss outcomes and appointment adherence 1 year after bariatric surgery. Further evidence of the MMPI-2-RF scale scores’ predictive validity were evidenced in terms of predicting adjustment-related difficulties as soon as one and three months after surgery (Marek, Ben-Porath, Merrell, Ashton, & Heinberg, 2014). In a sample of 859 bariatric surgery patients, individuals who scored higher on the somatic/cognitive measures of the test prior to surgery reported more medical complications one and three months after surgery. Patients who scored higher on Emotional/Internalizing Dysfunction (EID) scales of the MMPI-2-RF tended to report more psychological distress one and three months after surgery. Patients began to engage in maladaptive eating behaviors as soon as three months following surgery, and higher presurgical scores on the internalizing and externalizing scales of the MMPI-2-RF evidenced poorer eating behaviors after surgery.

Although the MMPI-2-RF does not assess for eating-related pathology directly (e.g., binge eating, purging), the test scores assess constructs related to eating disorders (Goodpaster et al., 2016; Marek, Ben-Porath, Ashton, & Heinberg, 2014a, 2014b). Marek, Ben-Porath, Ashton, and Heinberg (2014a, 2014b) reported that higher scores on internalizing and externalizing MMPI-2-RF scales were associated with increased risk for binge eating disorder, regardless of whether DSM-IV-TR (APA, 2000) or DSM-5 (APA, 2013) criteria were used. Goodpaster and colleagues (2016) reported that higher internalizing scores were predictive of patients who engaged in graze eating behaviors (i.e., picking and nibbling at food over a large period of time) prior to surgery.

As discussed earlier, general MMPI-2-RF interpretive guidelines identify a T-score of 65 as the cutoff for clinically significant elevation on the substantive scales of the inventory. However, interpretation of lower scores has received empirical support in the literature on use of the MMPI-2-RF in assessment of bariatric surgery candidates. Tarescavage et al. (2013) examined the clinical utility of both the SUI and SUB scales in two bariatric surgery samples from different clinics. A relatively larger proportion of patients responded to an SUI item that concerns thinking about death, which may in part reflect the extent to which their medical comorbidities are affecting these individuals’ cognitive thought process. By removing this item, the SUI scale better classified patients who had a history of suicide attempts. The SUB items are transparent. Approximately 30% of the bariatric surgery population engage in an under-reporting response style to some degree (Ambwani et al., 2013; Marek, Tarescavage et al., 2015). If the SUB scale were interpreted beginning at 55T, more patients
would be correctly classified as having a substance use problem without losing specificity. Similarly, Marek, Ben-Porath et al. (2014b) report that elevations of 55T or higher on some of the BXD scales yield higher odds ratios in regards to meeting criteria of binge eating disorder. At higher cutoffs, some of the Behavioral/Externalizing scales are indicative of social nonconformity; however, scores at the cutoffs reported on earlier are not necessarily indicative of social nonconformity, but are related to undercontrolled behaviors (e.g., loss of control over eating).

**Spine Surgery/Spinal Cord Stimulator Evaluations**

The MMPI-2 has been used widely for spine surgery and spinal cord stimulator evaluations. A primary reason for this is the number of psychological comorbidities found in this population, and their documented associations with poor surgical outcomes (Block, 2013; Block et al., 2013). Most notably, presurgical scores on Clinical Scales 1 (Hysteria), 2 (Depression), 3 (Hysteria), and 7 (Psychoasthenia) of the MMPI-2 have been associated with both pain sensitivity and diminished outcomes in these populations (Block, Gatchel, Deardorff, & Guyer, 2003; Fordyce, Bigos, Batti’e, & Fisher, 1992; Herron, Turner, Ersek, & Weiner, 1992). An algorithm used clinically to predict diminished outcomes and predict whether pre- or postsurgical psychological intervention is warranted was developed using scale score cutoffs on the MMPI-2 Clinical Scales in combination with other psychological and medical risk factors (Block et al., 2003; Block, Ohnmeiss, Guyer, Rashbaum, & Hochschuler, 2001).

More recent research has demonstrated clinical utility for the MMPI-2-RF in spine surgery and spinal cord stimulator settings. Comparison group data are published in the MMPI-2-RF Technical Manual (Tellegen & Ben-Porath, 2008/2011) and in a more recent study using a much larger sample (Block et al., 2013). Comparison group data are combined for both populations owing to both the similarities and lack of differences in scale scores that they produce. Block et al. (2013) also reported presurgical associations between self-report measures of mood, anxiety, functional disability, pain, and impairment (among others) and the MMPI-2-RF Substantive Scales in samples of 1,341 patients seeking spine surgery or a spinal cord stimulator. In summary, good convergent and discriminant validity estimates were obtained linking MMPI-2-RF and the other self-report measures. For example, RC1 was positively associated with measures of pain, functional disability, and measures of anxiety. Correlations between RC2 and measures of depression were more strongly associated than with measures of anxiety (Block et al., 2013). The MMPI-2-based algorithm just mentioned was revised to incorporate MMPI-2-RF scores instead (Block, 2013; Block, Ohnmeiss, Ben-Porath, & Burchett, 2011). The MMPI-2-RF algorithm worked as well as the MMPI-2 based algorithm, as the study demonstrated that patients who were deemed “fair” or “fair to poor” candidates for surgery reported suboptimal outcomes around 144 days postsurgery. Using hierarchical regression analyses controlling for presurgical functioning and patient expectations in a sample of 269 spine surgery patients, scores on both the RC1 and the MLS scales of the MMPI-2-RF are associated with presurgical functional disability and pain sensitivity, and they predict poorer functioning and higher pain ratings about five months after spine surgery (Marek, Block, & Ben-Porath, 2015). Using a similar analytic design, Block, Marek, Ben-Porath, and Kukal (2015) found that RCd, SFD, FML, and NEGE-r scores were associated with postsurgical functional disability and negative affect in both spine surgery and spinal cord stimulator patients (Block et al., 2015; Marek, Block et al., 2015). In summary, MMPI-2-RF Substantive Scales incrementally predict poorer outcomes after controlling for baseline measures, surgical expectations, and workers’ compensation status (Block, Marek, Ben-Porath, & Ohnmeiss, 2014; Marek, Block et al., 2015).

**Other Medical Populations**

Empirical evidence supporting use of the MMPI-2-RF in the assessment of psychogenic nonepileptic seizures and chronic pain patients has also been accumulated. In the past, the MMPI-2 was been used
often in these settings, although it became less frequently used due to the psychometric limitations of the Clinical Scales. This next section briefly reviews the history of using the MMPI-2 in these settings and presents studies supporting the validity and utility of the MMPI-2-RF in assessments of patients with suspected psychogenic nonepileptic seizures and in chronic pain settings.

Seizure Disorders

Individuals with seizure disorders are at a higher risk for emotional, thought, and behavioral disorders as compared to the general population, and up to 48% of patients with seizure disorders meet criteria for a psychological disorder (Devinsky, 2003; Gaitatzis, Trimble, & Sander, 2004). Although neuropsychological assessment of patients with seizure disorders are recommended to aid in treatment planning (Engel, 2013), no formal guidelines on psychological evaluations are suggested. Moreover, around 20%–30% of patients in epilepsy centers are diagnosed with psychogenic nonepileptic seizures (Benbadis, 2005). A patient presenting with possible psychogenic nonepileptic seizures may appear to be having epileptic seizures, but no abnormal electrical discharges are evidenced during electroencephalograph (EEG) scans. In essence, psychogenic nonepileptic seizures are a type of conversion disorder (APA, 2013).

As with the other medical populations discussed in this chapter, the MMPI instruments have been widely used with epilepsy patients (Russell, Coady, & Chaytor, 2009). Scales 1, 2, 3, and 8 of both the MMPI and MMPI-2 demonstrated good predictive utility in differentiating epileptic from nonepileptic seizure patients (cf. Crager et al., 2003; Wilkus, Dodrill, Thompson, 1984). It is not uncommon to evidence a “Conversion V” profile (Graham, 2011) in this population, such that nonepileptic seizure patients would produce higher elevations on Scale 1 and Scale 3 relative to Scale 2 (Russell, Coady, & Chaytor, 2009). Moreover, if the MMPI-2 was used a screening instrument in this population, long-term medical costs were often reduced (Crager, Berry, Fakhoury, Cibula, & Schmidt, 2002).

A number of studies have supported using MMPI-2-RF scales with this population. Bowden et al. (2014) found that the RC scales produced results indicating improved validity in comparison with the MMPI-2 Clinical Scales in a sample of 150 patients diagnosed with seizure disorders. Scores on MMPI-2 and MMPI-2-RF yielded a statistically significant level of agreement (68% to 84%) with clinically derived diagnoses of seizures. In instances of disagreement, MMPI-2 Clinical Scales were elevated whereas the RC Scales were not. Regression analyses indicated that psychological distress, subtle items, social desirability, and other demographic factors contributed to MMPI-2 Clinical Scale elevations in instances of disagreement. In a study comparing scale score differences between patients diagnosed with epileptic versus nonepileptic seizures, Locke et al. (2010) reported that nonepileptic seizure patients scored higher on the Fs, FBS-r, and RC1 and its associated Specific Problems Scales and scored lower on RC3 when compared with patients who were diagnosed with genuine epileptic seizures. These scores suggest patients with nonepileptic seizures tend to over-report somatic complaints and have a psychological component associated with their somatic symptoms. Low RC3 scores reflect naïveté and indicate that patients describe others as being well-intentioned and trustworthy while disavowing cynical beliefs.

Two scales to possibly complement RC1, the Psychogenic Nonepileptic Seizures Physical Complaints (PNES-pc) Scale and the Psychogenic Nonepileptic Seizures Attitudes (PNES-a) Scale, were developed to provide additional information to aid differential diagnosis in this population (Locke & Thomas, 2011). The scales were created using a combination of empirically based techniques (i.e., factor analyses, item response theory analyses) and asking experts to select items that best capture the aforementioned constructs. PNES-pc and PNES-a accounted for 33% of the variance in the diagnosis of psychogenic nonepileptic seizure and correctly classified 73% of patients, whereas the RC scale accounted for 20% of the variance in the diagnosis and correctly classified 67% of patients. Although the PNES-pc and PNES-a scales outperformed RC1 in Locke & Thomas’s (2011) study, the scales need to be cross-validated prior to being recommended for clinical use in this setting.
In terms of clinical utility, Myers, Fleming, Lancman, Perrine, and Lancman (2013) reported that high scores on RC1, RC2, and RC3 were predictive of patients with psychogenic nonepileptic seizures who used more emotional-focused coping strategies (e.g., self-blame and angry outbursts) when stressed. Patients who score lower on these scales (most notably RC2) used more task-oriented or avoidance strategies to reduce stress. RC3 scores also predict alexithymia in psychogenic nonepileptic seizure patients (Myers, Matzner, Lancman, Perrine, & Lancman, 2013). Higher scores on a number of the EID scales are associated with poorer quality of life in patients diagnosed with nonepileptic seizures (Myers, Lancman, Laban-Grant, Matzner, & Lancman, 2012). Elevations on RCd differentiated patients who met criteria for posttraumatic stress disorder in a sample of nonepileptic seizure patients (Myers, Perrine, Lancman, Fleming, & Lancman, 2013).

**Chronic Low Back Pain**

Psychopathology is substantially prevalent in chronic low back pain patients (Coste, Paolaggi, & Spira, 1992). Indeed, up to 41% of patients who have chronic low back pain also meet criteria for diagnosable psychological disorder (Coste et al., 1992). This population is at risk for mood, anxiety, and substance use disorders (Von Korff et al., 2005). In accordance with the biopsychosocial model of pain (Gatchel, McGearry, McGearry, & Lippe, 2014; Gatchel, Peng, Peters, Fuchs, & Turk, 2007), psychological factors interact with patients’ perspective of pain and also predict less favorable medical outcomes in this population (Hoogendoorn, van Poppel, Bongers, Koes, & Bouter, 2000). Both the American College of Physicians and the American Pain Society recommend that an assessment of psychosocial factors be integrated with patient care (Chou et al., 2007).

The MMPI and MMPI-2 were the most frequently used instruments in this population until the mid- to late 1990s (Piotrowski & Lubin, 1990; Piotrowski, 1998). However, use of the MMPI instruments in this setting has become sparse due to psychometric, practical, and theoretical limitations of the MMPI and MMPI-2 (Bradley, 1995; Keefe, Lefebvre, & Beaupre, 1995; Main & Spanswick, 1995; Sanders, 1995; Turk & Fernandez, 1995). In 1995, *Pain Forum* published debates on the utility of using the MMPI and MMPI-2 in chronic pain settings. Main and Spanswick (1995) criticized the inventory for its psychometric shortcomings. Bradley (1995) published a rebuttal that reviewed three studies of how MMPI instruments have been able to predict outcomes, medication use, disability, and other factors in chronic pain settings and can be usefully integrated into the biopsychosocial model of pain. Turk and Fernandez (1995) argued that the Clinical Scales were derived from an outdated psychopathology model and that the MMPI Restandardization Committee “forwent the opportunity to renew the nature of the MMPI and resulted in such obsolete remnants being retained” (p. 456). Keefe, Lefebvre, and Beaupre (1995) offered a more balanced view, suggesting that although advanced, quantitative procedures (just as longitudinal structural equation modeling) indeed should be more frequently utilized in chronic pain outcome studies.

Although the MMPI-2 is no longer being widely used in chronic pain settings, recent research with the MMPI-2-RF in these settings holds promise. As described earlier, McCord and Drerup (2011) demonstrated that the RC scales yielded substantially better discriminant validity in identifying depressed patients who also presented with chronic pain than did the MMPI-2 Clinical Scales. Tarescavage et al. (2015) published reliability, validity, and normative data in a sample of 811 chronic low back pain patients entering multidisciplinary treatment. Overall, internal consistency and mean inter-item correlation coefficients were similar to those reported for the normative sample. The test also demonstrated good convergent and discriminant validity in their sample. For example, the MMPI-2-RF somatic/cognitive scales were associated with pain, pain-related disability, hours resting per day, and medication use. MMPI-2-RF scales measuring constructs relevant to anxiety (e.g., RC7, AXY) were associated with external criteria such as benzodiazepine use, whereas Externalizing domain scales were associated with substance use and smoking in the chronic low back pain sample.
Other Internalizing domain scales of the MMPI-2-RF were robustly associated with other indicators or mood and antidepressant use. Scores on the THD domain scales were associated with posttraumatic distress. Arbisi and colleagues (2011) indicated that patients with re-experiencing symptoms, hyperarousal, and distorted negative worldviews tend to score higher on the THD domain scales than controls. Taressavage and colleagues (2015) concluded that MMPI-2-RF scale score reliability coefficient estimates were similar to those reported in other samples and the test evidenced good convergent and discriminant validity when comparing scale scores to external criteria.

Limitations of the MMPI-2-RF

As with any instrument, there are limitations associated with use of the MMPI-2-RF in behavioral medicine settings. Because the test is still relatively new, descriptive data and psychometric information for many medical settings, such as oncology, cardiology, organ transplant evaluations, and rehabilitation settings are not yet available. Nonetheless, MMPI-2-RF scores have been validated in a number of medical settings reviewed here and research to support and improve the use of the MMPI-2-RF in medical settings is ongoing. It is reasonable to expect that this research will also yield positive findings.

THE CASE OF MS. W

Ms. W is a 42-year-old divorced woman who was assessed at a hospital as a candidate for bariatric surgery. She was severely obese, with a BMI of 66 kg/m². She also met diagnostic criteria for numerous medical conditions related to obesity, including typ. 2 diabetes, obstructive sleep apnea, and hypertension. Prior efforts to lose weight through diet and exercise had been unsuccessful.

The hospital required a psychological assessment prior to surgery. This included a medical chart review, a semi-structured psychodiagnostic interview, and administration of the MMPI-2-RF. The purpose was to determine whether psychological factors might interfere with Ms. W’s ability to comply with presurgical recommendations and rigorous postoperative behavioral guidelines. The MMPI-2-RF was administered prior to Ms. W’s semi-structured psychodiagnostic interview and medical chart review. (It is usually best to administer the test ahead of time so that findings can inform the interviewer.)

Ms. W’s obtained T-scores for the MMPI-2-RF scales are included in Table 23.3, with the exception of item responses that cannot be reproduced in this chapter because of concerns about test security. These score serve as the basis for the MMPI-2-RF interpretation that follows.

Table 23.3 Comparison of Ms. W’s MMPI-2-RF Test Scores to the Test Scores of Other Bariatric Surgery Candidates (Women; N = 435)

<table>
<thead>
<tr>
<th>MMPI-2-RF Scale</th>
<th>Ms. W’s T-Scores</th>
<th>Comparison Group Mean T-Scores</th>
<th>Comparison Group Standard Deviation</th>
<th>Percent Scoring at or Below Ms. W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validity Scales:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variable Response Inconsistency (VRIN-r)</td>
<td>43</td>
<td>48</td>
<td>9</td>
<td>44</td>
</tr>
<tr>
<td>True Response Inconsistency (TRIN-r)</td>
<td>57 T</td>
<td>52F</td>
<td>8</td>
<td>85</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>MMPI-2-RF Scale</th>
<th>Ms. W’s T-Scores</th>
<th>Comparison Group Mean T-Scores</th>
<th>Comparison Group Standard Deviation</th>
<th>Percent Scoring at or Below Ms. W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrequent Responses (F-r)</td>
<td>61</td>
<td>51</td>
<td>10</td>
<td>89</td>
</tr>
<tr>
<td>Infrequent Psychopathology</td>
<td>59</td>
<td>48</td>
<td>8</td>
<td>95</td>
</tr>
<tr>
<td>Responses (Fp-r)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrequent Somatic Responses (Fs)</td>
<td>74</td>
<td>51</td>
<td>11</td>
<td>97</td>
</tr>
<tr>
<td>Symptom Validity (FBS-r)</td>
<td>67</td>
<td>58</td>
<td>11</td>
<td>85</td>
</tr>
<tr>
<td>Response Bias Scale (RBS)</td>
<td>59</td>
<td>50</td>
<td>11</td>
<td>87</td>
</tr>
<tr>
<td>Uncommon Virtues (L-r)</td>
<td>47</td>
<td>51</td>
<td>9</td>
<td>47</td>
</tr>
<tr>
<td>Adjustment Validity (K-r)</td>
<td>35</td>
<td>52</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td><strong>Higher-Order Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional/Internalizing Dysfunction (EID)</td>
<td>75</td>
<td>51</td>
<td>11</td>
<td>97</td>
</tr>
<tr>
<td>Thought Dysfunction (THD)</td>
<td>48</td>
<td>47</td>
<td>8</td>
<td>71</td>
</tr>
<tr>
<td>Behavioral/Externalizing Dysfunction (BXD)</td>
<td>46</td>
<td>47</td>
<td>8</td>
<td>53</td>
</tr>
<tr>
<td><strong>Restructured Clinical (RC) Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demoralization (RCd)</td>
<td>66</td>
<td>52</td>
<td>10</td>
<td>89</td>
</tr>
<tr>
<td>Somatic Complaints (RC1)</td>
<td>65</td>
<td>57</td>
<td>11</td>
<td>85</td>
</tr>
<tr>
<td>Low Positive Emotions (RC2)</td>
<td>69</td>
<td>53</td>
<td>10</td>
<td>93</td>
</tr>
<tr>
<td>Cynicism (RC3)</td>
<td>57</td>
<td>46</td>
<td>9</td>
<td>92</td>
</tr>
<tr>
<td>Antisocial Behaviors (RC4)</td>
<td>54</td>
<td>49</td>
<td>9</td>
<td>78</td>
</tr>
<tr>
<td>Ideas of Persecution (RC6)</td>
<td>66</td>
<td>50</td>
<td>9</td>
<td>96</td>
</tr>
<tr>
<td>Dysfunctional Negative Emotions (RC7)</td>
<td>68</td>
<td>49</td>
<td>9</td>
<td>97</td>
</tr>
<tr>
<td>Aberrant Experiences (RC8)</td>
<td>47</td>
<td>47</td>
<td>8</td>
<td>69</td>
</tr>
<tr>
<td>Hypomanic Activation (RC9)</td>
<td>45</td>
<td>45</td>
<td>8</td>
<td>57</td>
</tr>
<tr>
<td><strong>Somatic Complaints Specific Problems Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaise (MLS)</td>
<td>81</td>
<td>65</td>
<td>10</td>
<td>97</td>
</tr>
<tr>
<td>Gastrointestinal Complaints (GIC)</td>
<td>46</td>
<td>53</td>
<td>12</td>
<td>70</td>
</tr>
<tr>
<td>Head Pain Complaints (HPC)</td>
<td>59</td>
<td>58</td>
<td>11</td>
<td>70</td>
</tr>
<tr>
<td>Neurological Complaints (NUC)</td>
<td>70</td>
<td>53</td>
<td>11</td>
<td>94</td>
</tr>
<tr>
<td>Cognitive Complaints (COG)</td>
<td>58</td>
<td>48</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td><strong>Internalizing Specific Problems Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suicidal/Death Ideation (SUI)</td>
<td>45</td>
<td>47</td>
<td>7</td>
<td>92</td>
</tr>
<tr>
<td>Helplessness/Hopelessness (HLP)</td>
<td>52</td>
<td>45</td>
<td>7</td>
<td>90</td>
</tr>
<tr>
<td>Self-Doubt (SFD)</td>
<td>65</td>
<td>53</td>
<td>11</td>
<td>90</td>
</tr>
<tr>
<td>Inefficacy (NFC)</td>
<td>58</td>
<td>49</td>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td>Stress/Worry (STW)</td>
<td>81</td>
<td>49</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Anxiety (AXY)</td>
<td>59</td>
<td>50</td>
<td>9</td>
<td>90</td>
</tr>
<tr>
<td>Anger Proneness (ANP)</td>
<td>66</td>
<td>48</td>
<td>10</td>
<td>95</td>
</tr>
<tr>
<td>Behavior-Restricting Fears (BRF)</td>
<td>71</td>
<td>48</td>
<td>8</td>
<td>99.5</td>
</tr>
<tr>
<td>Multiple Specific Fears (MSF)</td>
<td>78</td>
<td>51</td>
<td>8</td>
<td>100</td>
</tr>
<tr>
<td><strong>Externalizing Specific Problems Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juvenile Conduct Problems (JCP)</td>
<td>50</td>
<td>51</td>
<td>11</td>
<td>67</td>
</tr>
<tr>
<td>Substance Abuse (SUB)</td>
<td>50</td>
<td>46</td>
<td>7</td>
<td>83</td>
</tr>
<tr>
<td>Aggression (AGG)</td>
<td>51</td>
<td>45</td>
<td>8</td>
<td>86</td>
</tr>
<tr>
<td>Activation (ACT)</td>
<td>44</td>
<td>47</td>
<td>10</td>
<td>51</td>
</tr>
</tbody>
</table>

Table 23.3 (Continued)
### MMPI-2-RF Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Ms. W’s T-Scores</th>
<th>Comparison Group Mean T-Scores</th>
<th>Comparison Group Standard Deviation</th>
<th>Percent Scoring at or Below Ms. W</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interpersonal Specific Problems Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family Problems (FML)</td>
<td>63</td>
<td>49</td>
<td>10</td>
<td>91</td>
</tr>
<tr>
<td>Interpersonal Passivity (IPP)</td>
<td>52</td>
<td>50</td>
<td>10</td>
<td>75</td>
</tr>
<tr>
<td>Social Avoidance (SAV)</td>
<td>75</td>
<td>52</td>
<td>12</td>
<td>96</td>
</tr>
<tr>
<td>Shyness (SHY)</td>
<td>75</td>
<td>49</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>Disaffiliativeness (DSF)</td>
<td>78</td>
<td>48</td>
<td>8</td>
<td>99.5</td>
</tr>
<tr>
<td><strong>Interest Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aesthetic-Literary Interests (AES)</td>
<td>33</td>
<td>49</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Mechanical-Physical Interests (MEC)</td>
<td>38</td>
<td>44</td>
<td>6</td>
<td>38</td>
</tr>
<tr>
<td><strong>Personality Psychopathology Five (PSY-5) Scales</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aggressiveness—Revised (AGGR-r)</td>
<td>45</td>
<td>45</td>
<td>8</td>
<td>44</td>
</tr>
<tr>
<td>Psychoticism—Revised (PSYC-r)</td>
<td>52</td>
<td>47</td>
<td>8</td>
<td>78</td>
</tr>
<tr>
<td>Disconstraint—Revised (DISC-r)</td>
<td>38</td>
<td>46</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>Negative Emotionality/Neuroticism—Revised (NEGE-r)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introversion/Low Positive Emotionality—Revised (INTR-r)</td>
<td>67</td>
<td>52</td>
<td>11</td>
<td>89</td>
</tr>
</tbody>
</table>

**Note.** Permission to report comparison group data was obtained by the test author Yossef S. Ben-Porath; Ms. W responded to all test items. For TRIN-r, “T” stands for True and “F” stands for False.

Comparison group data excerpted from the MMPI-2-RF Technical Manual by Yossef S. Ben-Porath and Auke Tellegen. Copyright © 2008, 2011 by the Regents of the University of Minnesota. Reproduced by permission of the University of Minnesota Press. All rights reserved. “Minnesota Multiphasic Personality Inventory-2-RF” and “MMPI-2-RF” are trademarks owned by the Regents of the University of Minnesota.

## MMPI-2-RF Interpretation

Ms. W produced a valid and interpretable MMPI-2-RF protocol. Her pattern of responses indicated that she responded consistently and attended to the content of the items. There were no indications of over- or under-reporting psychopathology, cognitive, or somatic symptoms. Her resulting scores should provide an accurate representation of her current psychological functioning.

Ms. W’s scores indicate significant emotional distress. More specifically, she reports a lack of positive emotional experiences, significant anhedonia, and lack of interest. She is likely pessimistic, socially introverted and disengaged, lacks energy, and likely displays vegetative symptoms of depression. Ms. W also reported various negative emotional experiences. She reported multiple problems involving experiences of stress and worry, which likely include preoccupation with disappointments, difficulties with time pressure, and worries over misfortune and finances. She is likely to be stress-reactive and worry-prone and to engage in obsessive rumination. In addition, she reported multiple fears that may significantly restrict normal activity in and outside the home. She also reported multiple specific fears, such as fear of blood, fire, thunder, water, natural disasters, spiders, mice, and other animals. Finally, Ms. W reported anger-proneness, indicating that she is likely to have problems with anger, irritability, and low tolerance for frustration.

Ms. W reported feeling sad and unhappy and being dissatisfied with her current life circumstances. She is likely to complain of feeling depressed. She also reported self-doubt and is likely to be prone to rumination, to feel insecure and inferior, and to be self-disparaging and intropunitive.

Ms. W reported multiple somatic complaints, including vague neurological complaints. She is likely prone to developing physical symptoms in response to stress. Ms. W also reported a general
Ms. W reported significant persecutory ideation such as believing that others seek to harm her. She is likely to be suspicious of and alienated from others and may lack insight or experience interpersonal difficulties as a result of suspiciousness.

There were no indications of maladaptive externalizing behavior in Ms. W’s protocol.

Interpersonally, Ms. W reported that she does not enjoy social events and avoids social situations. She is likely to be introverted, have difficulty forming close relationships, and be emotionally restricted. She also reported being shy, easily embarrassed, and uncomfortable around others. She is likely to be socially inhibited and to be anxious and nervous in social situations. In addition, she reported disliking people and being around them, indicating that she is likely asocial.

Ms. W reported an uncommon lack of interest in two broad areas. She indicated little or no interest in activities or occupations of either an aesthetic or literary nature (e.g., writing, music, the theater) or a mechanical or physical nature (e.g., fixing and building things, the outdoors, sports).

Ms. W’s MMPI-2-RF scores identify several diagnoses for further consideration, including anxiety-related disorders, particularly those involving excessive stress and worry such as obsessive-compulsive disorder, agoraphobia, and specific phobias (including social phobia and avoidant personality disorder). In addition, anger-related and depression-related disorders should also be considered. Ms W’s test scores also indicated the need to consider the possibility that she has a somatoform disorder. However, the extent to which genuine physical health problems contribute to her scores on the RC1 and NUC scales should be considered. Lastly, disorders involving persecutory ideation should be ruled out.

Ms. W’s protocol also suggests a number of treatment targets and considerations. Her psychological distress should be an initial target for treatment. Specific targets in this domain may include her low self-esteem, self-doubt, anhedonia, and dysfunctional negative emotions, including excessive worry and rumination, behavior-restricting fears, and specific fears. She may also benefit from stress and anger management skills. In addition, her persecutory ideation, difficulties associated with social avoidance, and anxiety in social situations should also be considered as possible treatment targets. Given her level of distress and lack of positive emotions, Ms. W may benefit from a psychiatric referral to evaluate the appropriateness of psychotropic medication.

Finally, Ms. W’s readiness and motivation for treatment should be considered. She may be likely to reject psychological interpretations of somatic complaints and her malaise may impede her willingness or ability to engage in treatment. Her persecutory ideation may interfere with establishing rapport and treatment compliance. Moreover, her aversive response to close relationships may make it difficult to form a therapeutic alliance and achieve progress in treatment. On the other hand, her significant emotional distress and current health status may motivate her to commit to treatment.

Semi-structured Psychodiagnostic Interview

During the interview, Ms. W noted she was motivated for surgery by medical problems, pain, and limited mobility. She demonstrated a fair understanding of the surgery, risks, and benefits, and was judged needing to receive additional education in regard to nutrition after surgery and behavioral changes necessary due to some misunderstandings. Ms. W reported that she was not currently working and that she was receiving disability benefits due to obesity and depression. She stated that she is not very motivated to try to find work after surgery, but will see how she feels after she has lost some weight.
When reviewing her medical diagnoses, she reported only obtaining 4 hours of sleep on average due to her obstructive sleep apnea. She reported that she was nonadherent with using her continuous positive airway pressure (CPAP) device, stating she did not like how it felt. In regard to eating behaviors, she reported skipping breakfast, having a large lunch and dinner, and continuous snacking behaviors. Indeed, she reported behaviors associated with graze eating. She also reported drinking 1–2 carbonated and sugary beverages per day.

In regards to her mental health, Ms. W confirmed a medical chart diagnosis of unspecified depressive disorder and reported taking an antidepressant that was managed by her primary care physician (PCP). She reported that, although she thought she had benefited from medication management, she still experienced symptoms associated with anhedonia, psychomotor retardation, dysphoria, and low self-esteem. She also reported not feeling comfortable in her body, stating that she often feels others are judging her, making fun of her, and criticizing her for her weight. She indicated that these self-referential ideations keep her from going out to public places, such as the grocery store. She stated that her mother often does the shopping for her. She also reported having excessive worry, difficulty controlling her worries, anger and irritability when feeling worried, and worry that can often make it difficult to fall asleep at night. She denied current and past use of alcohol, tobacco products, and illicit substances. Her MMPI-2-RF results were then reviewed with her. Marked consistencies between her MMPI-2-RF profile and reported symptoms in the interview were pointed out. She was relatively open to feedback, but worried that her tendency to somaticize may result in some difficulties coping with the postoperative stress associated after surgery.

A number of concerns regarding her psychosocial functioning needed to be considered. Antidepressants tend not to be absorbed well after bariatric surgery (Roerig et al., 2012), and Ms. W did not appear to be responding well to her current dosage. She also was experiencing significant levels of anxiety-related symptoms that have gone untreated and impaired her ability to fall asleep, work, and support herself. Furthermore, she was not wearing her CPAP, solely because she did not like how it felt. Because of these reasons, the psychologist concluded that Ms. W needed to be seen by a psychiatrist for possible changes in her medication and she also needed to receive psychotherapy with a psychologist prior to being cleared for surgery. Finally, she was required to return to her sleep doctor to have her CPAP refitted and engage in four sessions of cognitive behavioral treatment for her graze eating behaviors. Ms. W agreed to follow through with these recommendations.

Case Discussion

As outlined earlier in this chapter, presurgical evaluations of bariatric surgery candidates are designed to identify barriers to successful surgical outcomes. Various forms of psychopathology ranging from emotional, thought, and interpersonal functioning are associated with suboptimal outcomes and behavioral disconstraint is associated with poor adherence to postoperative appointments, poor postsurgical eating behaviors, and suboptimal weight loss (Marek, Ben-Porath, & Heinberg, 2016). For this reason, objective psychological testing is recommended along with a clinical interview in bariatric surgery evaluations (Mechanick et al., 2013). In Ms. W’s case, the information in her medical chart and what she reported during the clinical interview were congruent with her MMPI-2-RF results, which also raised some additional concerns, discussed next.

Examination of Ms. W’s MMPI-2-RF Substantive Scales scores identified many areas of concern that needed to be addressed prior to being further considered for surgery. First and foremost, her scores on RC2 (69T), RCd (66T), and RC7 (68T) indicated that Ms. W was experiencing a number of symptoms associated with depression and anxiety. Indeed, she reported ongoing depressive symptoms during the clinical interview despite being medically managed for this condition by her PCP.
Moreover, her excessive worry and irritability had caused her additional sleep problems that were reflected in her MMPI-2-RF protocol (STW [81T], ANP [66T], and NEGE-r [77T]). Moreover, her scores on RC6 (66T) and a number of the Interpersonal Problems Scales, such as SAV (75T), indicated that Ms. W believed others were discriminating against her due to her weight and feared being embarrassed in social situations. Lastly, research reviewed earlier in this chapter shows that higher scores on the Somatic/Cognitive Scales of the MMPI-2-RF are associated with increased risk for excessive pain, nausea, and vomiting one and three months following the procedure (Marek, Ben-Porath, Merrell et al., 2014), indicating the advisability of closely monitoring her somatic symptoms postsurgery.

During a multidisciplinary team meeting following Ms. W’s psychological evaluation, members believed that Ms. W should be admitted to surgery as soon as possible due to her numerous medical diagnoses and higher than average (for a bariatric surgery candidate) body mass index (BMI). Her PCP agreed to moderate the doses of the antidepressants Ms. W was currently prescribed following the procedure. The surgeon performed a Roux-en-Y Gastric Bypass about a month after her psychological evaluation. At that time, she had not yet found a therapist for treatment of graze eating, depression, or anxiety-related symptoms.

At postsurgery appointments, Ms. W had lost weight at a rate that was expected given her pre-surgical BMI; however, other outcome assessments indicated that she was not doing well. At her 1-month follow-up, she reported a worsening of her depressive symptoms and grieved not having food as a coping mechanism. She also reported excessive pain that had caused numbing from her waist down to her feet. She reported excessive nausea, vomiting, and was not meeting her minimum protein and water intake. At her 3-month appointment, Ms. W continued to report numbness, excessive pain, nausea, and vomiting. She remained nonadherent to her protein and hydration intake, and was also attempting to graze-eat on sweets, which in turn caused her to become sick when she engaged in those maladaptive eating behaviors. She was again encouraged to seek therapy, whether it be with the psychologists on staff or with an outside provider.

In summary, Ms. W was prioritized for surgery before appropriate psychological interventions were implemented. As a result, Ms. W achieved suboptimal, short-term outcomes. Using the MMPI-2-RF in this case example provided useful information above and beyond what was obtained from Ms. W’s medical chart and relying on her semi-structured clinical interview data. Her MMPI-2-RF scores indicated that Ms. W’s depressive and anxious symptoms needed to be treated prior to surgery and monitored after. Although her PCP agreed to moderate her antidepressant medication, her depression worsened after surgery according to her self-report. The MMPI-2-RF test scores also indicated that Ms. W likely somatizes, and was at risk for developing excessive somatic symptoms postoperatively (Marek, Ben-Porath, Merrell et al., 2014). Indeed, Ms. W experienced a number of somatic complaints at her postoperative appointments, some of which were likely not related to the surgery.

**SUMMARY**

The MMPI-2-RF is gaining considerable recognition in the science and practice of behavioral medicine. In addition to its strong psychometric properties, the test can be administered in well under an hour (closer to 30 minutes if administered via computer). The instrument assesses a wide array of psychological constructs relevant to the diagnosis of psychopathology and identification of behaviors and interpersonal functioning relevant to behavioral medicine settings. A standard interpretative strategy and guidelines facilitate cross-examiner consistency in test score interpretation. A broad and growing empirical literature is available to guide MMPI-2-RF interpretation in a range of medical settings.
The MMPI-2-RF’s relatively short administration time and theoretically grounded scales, along with its strong conceptual and empirical foundation, offer a number of advantages for use in medical settings, including primary care, bariatric surgery, spine surgery, spinal cord stimulator, nonepileptic seizure, and chronic pain populations. Although many self-report measures are commonly used in primary care settings, most are narrow-band and focus on one diagnosis or construct. Within 25–35 minutes, the MMPI-2-RF can assess for a wide array of psychological, behavioral, and interpersonal functioning; aid or challenge a diagnostic consideration; provide clinical information on risk factors that patients may be sensitive to disclosing to their doctor or their psychologist when rapport has yet to be established; and generally assist in determining the appropriateness of referring the patient for specialty care. The test can also assess the extent to which a patient is over- or under-reporting symptoms, which can assist in determining the degree to which they may be accurately reporting their functioning. Overall, the MMPI-2-RF can aid in the diagnostic and treatment considerations in a variety of primary care settings when used in conjunction with other sources of information.

AUTHOR’S NOTE

Yossef Ben-Porath is a paid consultant to the MMPI-2-RF publisher, the University of Minnesota and Distributor, Pearson. As co-author of the MMPI-2-RF, he receives royalties on sales of the test.

REFERENCES


PART III

Examples of Integrated Healthcare Programs
The problem of pain persists, having defied attempts at resolution, and in the process suffocated the efforts of many to live meaningful lives and have loving relationships. Primary care is the optimal setting for addressing persistent pain (PP), and Focused Acceptance and Commitment Therapy (Robinson, Gould, & Strosahl, 2010; Strosahl, Robinson, & Gustavsson, 2012) offers a novel and powerful way of conceptualizing and intervening with PP. This chapter is an effort to share what has been learned about primary care treatment of PP in the authors’ combined 30 years of working as primary care behavioral health scientist practitioners. The chapter provides a brief summary of evidence and then shifts to an array of options for improving the outcomes of primary care services to patients with PP. These will include a variety of practice tools, ranging from scripts to guide new conversations between clinicians and patients to curricula for group medical visit services.

PP is the “new” chronic pain, and switching to this term may help us start a new, more positive conversation with patients and among each other, as professionals. Over the past several decades, the term chronic pain has become imbued with a variety of negative connotations, none of which generates the hope one would like to bring to the care of patients suffering from it. When speaking of chronic pain, patients tend to want providers to help them make it “not chronic.” Time will tell if use of the term PP makes a difference in the ability to take a more mindful perspective on pain and to promote that in patients. Given the state of pain care, it is worth the effort if it creates even a small increase in optimism or a just noticeable boost to the idea that it is possible to suffer from pain and, at the same time, live a meaningful life that includes awareness, compassion for self and others, and intentional choice.

The explosion of PP, intermixed with addiction to pain medications, is occurring in a socio-political-psychological-economic-medical storm. While explaining all of the factors and their synergetic relationships is beyond the scope of this chapter, it is hoped that the reader will take time to deepen his or her knowledge and broaden their perspective. The book Dreamland (Quinones, 2015) provides intelligence from multiple perspectives and fosters wise action. With greater understanding and with the tools offered in this chapter, the reader has a better chance of calming this storm, rehabilitating the victims, and fortifying the healthcare system and its providers.

PAIN PROBLEMS AND OPPORTUNITIES IN PRIMARY CARE

Pain that persists for months, years, and decades takes a toll on patients, their families, the healthcare systems and the providers of care. Before discussing the opportunities for improving primary care services for patients with PP, let’s take a look at current challenges in the practice of primary care medicine. The Triple Aim and the Quadruple Aim concepts provide a framework for our discussion.
Triple Aim Outcomes for Patients With Persistent Pain

Despite spending far more than any other country on health care, the United States lags far behind other countries in results. Berwick, Nolan, and Whittington (2008) offer the Triple Aim as a moniker for three keys to improving the US healthcare system (i.e., lowered cost, improved population health, and improved patient experience). As an example of the problems the system has, Berwick and colleagues present congestive heart failure, the most common reason for admission of Medicare patients to a hospital. Nearly 40% of patients presenting with congestive heart failure are readmitted within 90 days, even though well-designed demonstration projects have shown for a number of years that proper management of patients can reduce the readmission rate by more than 80%. Thus, owing not to a lack of knowledge or technology, but rather to various deficits and inefficiencies in the current system, these patients are not as healthy or satisfied as they could be, and are more costly to care for. In the case of patients with PP, the system historically supported costly interventions of limited value, such as repeated back surgeries for patients with low back pain, while denying coverage for less costly and more promising services such as cognitive behavioral therapy, acupuncture, massage therapy, and physical therapy.

Quadruple Aim

Bodenheimer and Sinsky (2014) expanded on the Triple Aim as a compass for optimizing health system performance by adding a goal of improving the work life of healthcare clinicians and staff. Physicians and other members of the primary care team report high rates of burnout and dissatisfaction, and the addition of an aim to improve provider experience is essential to achievement of the initial aims. Burnout is associated with lowered patient satisfaction and may also increase the cost of services. Undoubtedly, caring for patients with PP has been frustrating for most providers, particularly with the pressure on physicians to identify pain and then control or eliminate it with use of pain medications in the 1990s and the subsequent problems of increasing opiate dependence and opiate abuse among patients over the past 20 years.

The actual rates of physician burnout are surprisingly high, with a 2014 estimate placing it at 46% (Bodenheimer & Sinsky, 2014). Burnout rates are particularly high among family physicians and general internists. Survey results suggest that 68% of family physicians and 73% of general internists would not choose the same specialty if they could restart their career path (Kane & Peckham, 2013). These findings highlight the negative impact of today's situation, with the mismatch between what providers (as well as patients) wanted or expected and actual results, and the complete lack of sustainability of current approaches. Like most professionals, the driver for satisfaction among physicians is the ability to deliver quality (Friedberg, Chen, Van Busum et al., 2013).

Patients with PP may be among the more vocal groups of patients voicing dissatisfaction with care and complaining about poor quality. One of the most devastating experiences for a healthcare provider must be the loss of a patient to an avoidable death, such as an overdose on pain medications. This is certainly the most devastating experience for the parents and other relatives of the many lost to opiate overdose. This unfortunately is not uncommon for primary care physicians, and many of our readers probably know someone whose life has been touched by the loss of a friend or loved one to opiate misuse or overdose. Parents and providers are speaking out and communities are making efforts to change. There is a great deal we can do to improve care.

Improving Care

There’s one good thing about working to improve the quality of care for patients with chronic pain, and that is there’s lots of room for improvement. Here, the Quadruple Aim is used to organize possible
quality initiatives concerning PP. There is no need to embark on all of them; instead, one should choose one or two areas and enlist the support of coworkers in implementing a pilot study in their setting. As a part of evaluating one’s efforts, one must be sure to identify specific metrics that will be used to evaluate the baseline (if the metrics are currently in place) and the impact of the new efforts over time. Metrics that align with current primary care efforts are recommended (for example, see Stiefel & Nolan, 2012).

Also, an understanding of team members and how they work together is essential to improving care. On March 23, 2010, President Obama signed the Patient Protection and Affordable Care Act, commonly called the Affordable Care Act (ACA). It offers a comprehensive health insurance reform that is encouraging of strong and potentially helpful changes to the way care is delivered to patients with PP, including recreating the mix of clinicians providing team-based care in primary care. There will be increasing numbers of behavioral health providers in primary care, working in fully integrated roles. The recommended model is the Primary Care Behavioral Health (PCBH) model (Robinson & Reiter, 2007; Robinson & Reiter, 2015). This approach places a Behavioral Health Consultant (BHC) alongside medical providers in order to provide episodic and longitudinal brief services. As an expert in behavior change, the BHC strives to add evidence-based interventions to care plans to support health behavior change. The BHC often up-skills other members of the team in behavior change technology, while simultaneously learning more about medical conditions from medical providers of the team. Figure 24.1 provides the PCBH Big 10 Checklist for readers to use in assessing the extent to which efforts to provide behavioral health services in primary care align with the PCBH approach. One should use this on an ongoing basis and identify targets to enhance your consistency with this full integration approach.

The ACA also increases access to affordable health insurance options and encourages the formation of accountable care organizations (ACOs) to improve care. Now, millions of people who were previously uninsured (some of whom are victims of PP) are accessing care. ACOs, led by providers, have formed across the nation. Given that they are collectively accountable for quality and total per capita costs across the full continuum of care for a population of patients, they are creating payment systems that encourage quality and reduce costs. ACOs promote measurement of clinic and provider performance and use payment to incentivize changes designed to improve outcomes for patients with PP, as well as patients in other populations. The bottom line: the time is right to build a team capable of delivering quality care to patients with PP. And this team, of course, will be multidisciplinary.

Enhance Patient Experience
A foundational need for improving care to patients with PP is to explore a novel way of conceptualizing chronic pain that supports a new conversation with patients and their families. The new conversation needs to promote acceptance of a legitimate problem by the patient and all members of the care team. The conversation needs to provoke curiosity and interest and minimize fear and avoidance. Patients with PP need to feel welcomed and surrounded by a team of staff prepared to support them in a consistent manner. Providers need to broaden efforts of treatment to include specific lifestyle behaviors. Overall, the approach needs to lessen the amount of care time focused on pain elimination and increase time spent on assisting them with motivation and skill development for living meaningful lives, even with pain.

Improving Population Health
Historically, it has been difficult to identify patients with PP as a population. Typically, a patient population is identified by a diagnosis and patients with PP carry many different diagnoses, ranging from fibromyalgia to migraines to pelvic pain to low back pain. Central sensitization is a new
**PCBH Big Ten Checklist**

This checklist offers a way to evaluate your current efforts to use a behavioral health provider (BHP) as a member of your team to improve outcomes for patients and job satisfaction of all members of the team. Indicate yes or no, and identify targets for change (i.e., aspects of the program that you plan to change in the next 3 months).

<table>
<thead>
<tr>
<th><strong>In our program . . .</strong></th>
<th><strong>Yes/No</strong></th>
<th><strong>Target</strong>*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Are BHP services available during all clinic hours?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Does our BHP routinely complete 5+ same-day patient visits everyday?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Does the BHP help with patients of all ages (pediatric and adult) and any behavioral issue (i.e., chronic disease and preventive care needs, mental health and substance abuse, resource needs)?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Does the BHP complete 90% of their patient visits in 15–30 minutes?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Is the primary goal of a BHP visit to help a patient improve in functioning (i.e., ability to succeed as a partner, worker, student, parent, etc.)?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Does the BHP routinely complete 10+ visits/day?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Does the BHP chart SOAP notes in the medical record, where other team members can access them easily?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Does the BHP give same-day feedback to referring PCPs and RNs?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Does the BHP provide classes, groups, &amp;/or workshops for patients?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Does the BHC provide education to PCPs and staff about behavioral issues?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**BONUS QUESTION:**
Does the clinic have a pathway that describes what the BHP does to improve primary care quality initiatives (e.g., tobacco cessation counseling, SBIRT, depression, adolescent well child, etc.) and/or address high impact populations (e.g., patients with chronic pain)?

* Courtesy of Mountainview Consulting Group.

**Figure 24.1** PCBH Big Ten Checklist
conceptualization for PP. It is considered to be a common denominator for diverse medical conditions such as temporomandibular joint dysfunction, nonspecific low back pain, vulvar vestibulodynia, migraines, osteoarthritis, noncardiac chest pain, migraine, irritable bowel syndrome, and many so-called functional disorders. While not well understood by most medical providers, central sensitization offers a new way to identify patients with pain that persists beyond the time period expected and to offer more effective treatments and save money by avoiding unnecessary testing and related interventions.

Fortunately, many behavioral health providers have skills that can help alleviate some of the suffering individuals experience when central sensitization is perpetuating their medical disorders, and there is considerable support in the literature for these approaches. One of the difficulties in treating individuals with central sensitization, however, is the complexity of the phenomenon and the fact that it challenges common understandings of pain. Later in this chapter, readers will learn more about the use of this concept in explaining how pain originating in the peripheral nervous system can come to exert a strong impact on the central nervous system.

One should not wait for a common diagnosis to identify the large and growing group of patients with PP, but should begin now to develop a population-based focus for improving functional outcomes. Given the interaction between the experience of PP and lifestyle behaviors, there is a good chance of improving the outcomes for the population of patients with PP by targeting specific lifestyle behaviors, such as nutrition, exercise, pleasurable activities, and sleep. Later in this chapter, there is guidance on initiating a lifestyle change program for patients with PP.

Reduce Costs
It is understandable that clinicians recommend repeated tests and procedures to patients with PP, many of whom request just that. However, it is risky business, and patients may experience iatrogenic results. For example, a patient experiencing PP after an ectopic pregnancy may request to repeat procedures to again rule out the possibility of an active disease process. When a clinician consents to this, the patient may experience heightened anxiety rather than relief and may ask for anti-anxiety medicines as well as pain medicines “just until we know there’s nothing awful going on.” The negativity bias or fears of clinicians and patients can combine at times and move care in the direction of escalating costs with diminishing returns. This can and will change as one develops a new way of talking about pain and offers patients an array of options for working with PP.

A great deal of savings, in both money and suffering, may result from engaging patients with PP in primary care. All too often, patients become discouraged after interactions with primary care clinicians and disengage from care until they are in need of more urgent and costly care. In the process, they become more fragile, medically and psychologically, instead of more resilient. Engagement is a challenge, and one needs to be creative and persistent in building it.

Improve the Work Life of Clinicians and Staff
There are several strategies for moving the dial in this critical area. First is the redefinition of the primary care team—who, how many, and, of course, who does what? In addition to medical providers and nursing staff, the team will need behavioral health providers and, in some clinics, community health workers. The staffing ratio of behavioral health providers to PCPs will need to be adjusted to the needs of the clinic. Robinson and Reiter (2015) suggest a staffing ratio of 1:3 or 1:4 for clinics serving patients with greater mental and physical health needs (e.g., most health centers), while the ratio may be quite a bit lower for a clinic serving higher functioning patients (e.g., 1:7 or 1:8). They caution, however, that an inadequate ratio may compromise the quality and success of behavioral
health services, and this is of critical importance for clinics serving a significant number of patients with PP.

After redefining the team, all team members need to master the new way of talking about pain and learn key evidence-based interventions. When all team members understand the various interventions, all can support the patient in a continuous fashion. This is not to say that all will share the same level of expertise in all areas. Medical providers will always be the experts concerning medication strategies, and behavioral health providers are likely to know the most current evidence for behavioral interventions. It is to say that it is the responsibility of each team member to learn the basics of all interventions and to amplify the impact by supporting identified interventions in contacts with the patients. Another feature of this new way of working is to have all team members working at the top of their skill level, learning on an ongoing basis, and seeking supervision and consultation as needed. These strategies shift the burden of care from the medical clinician and distribute it among team members. This makes the job of the medical clinician more feasible and more enjoyable.

NEUROLOGICAL PERSPECTIVES ON PERSISTENT PAIN

As anyone who works in primary care can attest, chronic pain or PP is one of the more common conditions seen. With over 100 million Americans expressing concerns related to PP (Dzau & Pizzo, 2014; Institute of Medicine, 2011) and the majority of these individuals (63 percent) seeking care in primary care offices (National Center for Health Statistics, 2006), primary care provides most of the care for PP. Upwards of 20% of all outpatient visits are dedicated to PP, with 12% of all prescriptions being for treatment of PP (Alford et al., 2008). Problems associated with current treatments for PP tax a system already depleted of resources and represent a significant burden on communities and economies, with an estimated $100 billion lost due to lack of treatment or undertreatment of PP (Alford et al., 2008).

Pain, in an evolutionary context, is one of the most primitive biological functional systems in the human bodies. Family medicine residents learn to think of PP as the “STOP” signal for humans, suggesting discontinuation of an activity due to danger and/or harm. However, if pain signals a stop message, PP signals a persistent message of “STOP!” This is why the residual aspects of PP are often more debilitating than the actual pain, as patients develop a variety of comorbid concerns ranging from significant psychological impairment to sleep concerns to severe limitations in daily activities. Available data suggest that 80% of individuals with PP have daily functional disruption and over 60% have relationship difficulties (Vo, Marx, & Penles, 2008). These unfortunate consequences give impetus to the need for better understanding of the etiology of PP and the evidence for treatment of PP and comorbid conditions.

Etiology of Persistent Pain

A rudimentary explanation of pain involves a somatic experience that prompts a threat sensation related to the somatic experience and a subsequent emotional response of unpleasantness (Price, 1999). The biological pathway of pain has been identified and includes a stimulus that excites nociceptors, which initiates a cascade of signals to the spinal cord (SAMSHA, 2011). Once the message is transmitted to the spinal cord, multiple pathways break off to provide the comprehensive experience of pain. Specifically, information is sent to the somatosensory cortex to evaluate where the pain is
coming from and the intensity; information is also sent to the limbic system, which elicits the emotional response to the pain; autonomic centers of the brain also receive information that allows the body to be placed on “alert” (SAMSHA, 2011). The pain pathway also sends signals to other parts of the brain/body to produce an inflammatory response to the inflicted area in order to protect the damaged tissue.

Generally, pain can be broken into two categories: acute and chronic or persistent. Acute pain, which is the result of an injury that causes tissue damage, typically lasts less than three months. The cause of acute pain is usually easily identifiable and acute pain usually diminishes when treated. As stated earlier, acute pain is quite beneficial for humans and provides an alarm system to alert the person to a harmful act or situation. Without this alarm system, humans would be quite vulnerable and unable to detect when harm was occurring within their bodies.

Persistent pain is defined as pain lasting longer than three months or longer than what is typically expected with an ailment (Alder, 2000). Even when the pain culprit is removed or discontinued and treatment is applied, pain remains. Interestingly, medical evaluations and tests may not demonstrate any organic etiology causing the pain; nevertheless, the experience of pain remains. It is important to note that acute pain and PP are not mutually exclusive. That is, a person suffering from PP may also experience acute pain, and patients with PP are actually more susceptible to acute pain (SAMSHA, 2011).

Persistent pain can be classified as nociceptive or neuropathic pain. Briefly, nociceptive pain encompasses pain related to somatic complaints such as tissue damage (Hooten et al., 2013). Nociceptive pain can include pain resulting from mechanical, thermal, or chemical stimulation (SAMSHA, 2011). Functionally, nociceptive pain alerts the individual of danger; however, when prolonged, nociceptive pain can cause hypersensitivity to painful stimuli (Hooten et al., 2013). Neuropathic pain involves pain that results from damage to a neuropathway. This damage usually occurs in the somatosensory system and typically occurs secondary to a disease (e.g., diabetic neuropathy, autoimmune disorders; Hooten et al., 2013).

Theories about the development of PP have identified a number of potential mechanisms, such as peripheral sensitization, central sensitization, ectopic excitability, structural reorganization, primary sensory degeneration, and disinhibition (Devor & Seltzer, 1999; Levine & Reichling, 1999; Moore et al., 2002; Neumann, Doubell, Leslie, & Woolf, 1996; Woolf & Thompson, 1991; Woolf, 2004). Given the focus of this paper, the discussion will be limited to the mechanisms of peripheral and central sensitization. Peripheral sensitization refers to tissue damage that prompts inflammation resulting in change to the chemical environment, which results in increased pain levels and hypersensitivity to additional stimuli (Woolf, American College of Physicians, & American Physiological Society, 2004). Identification of the inflammation factors has prompted exploration of improving pain levels with diets that decrease inflammation (University of Wisconsin Integrative Medicine, 2007). Central sensitization refers to changes in the synaptic communication of nociceptor cells, which, similar to peripheral sensitization, heightens and lowers the threshold for firing of nociceptor receptors, resulting in a hypersensitivity to potential painful stimuli (Woolf & Thompson, 1991). Interestingly, the central sensitization mechanism is often not specific to a particular pain site and may prompt generalized pain (McAllister, 2016). This theory attempts to explain a common experience of PP patients, where pain spreads from an original area of injury (e.g., a back injury) to other areas (e.g., headaches, groin pain).

Predisposing factors for central sensitization involve both genetic and psychophysiological factors (McAllister, 2016). A predisposition for lower pain thresholds is a genetic factor. Psychophysiological factors, particularly the stress response, may also reduce pain thresholds and increase sensitivity of
pain. This theory is backed by growing evidence suggesting that adverse childhood experiences (ACE) predict multiple health conditions in adulthood, including the risk for PP (Anda, Tietjen, Schulman, Felitti, & Croft, 2010; Dube et al., 2009).

Persistent pain applies to different disorders with varied etiologies. The very broad category of musculoskeletal disorders (i.e., pain due to muscles, ligaments and tendons, bones and nerves; Cherney, 2013) falls under the umbrella of PP. Low back pain (e.g., lumbar degenerative disc disease; Ullrich, 2006), joint disease (e.g., arthritis), a very common problem, and other troubling pain syndromes (e.g., lupus, fibromyalgia, inflammatory bowel disease, chronic fatigue syndrome, headaches) are also a part of the group of problems that can be understood as PP.

**Conditions Associated With Persistent Pain**

Individuals who have PP are more likely to have comorbid physical concerns, overutilize medical services, and have poorer quality life and psychological functioning (Kato, Sullivan, Evengård, & Pedersen, 2006). While the link between depression and PP is well established, the direction of this relationship is less obvious and is most likely bidirectional. Individuals who have PP are more at risk for developing depression symptoms, and individuals who experience depressive symptoms are more prone to developing PP. This bidirectional pathway is largely attributed to the gate control theory of pain, which posits that both physical and emotional pain lower the human’s biological threshold for detecting pain, resulting in a hypersensitivity to pain (Deardorff, n.d.).

Perhaps less well known but important to a formulation for treating PP is research suggesting a higher rate of anxiety-related disorders among patients with PP than rates of depression (McWilliams, Cox, & Enns, 2003; McWilliams, Goodwin, & Cox, 2004). McWilliams et al. (2003) reported a strong link between PP and panic and posttraumatic stress disorder. It appears that frequency and intensity of traumatic events also predict future probability of developing PP (Casey, Greenber, Nicassio, Harpin, & Hubbard, 2008). This finding again supports the role ACEs play in developing PP.

Sleep-related disorders also appear to be related to PP. Similar to the bidirectional relationship between depression and PP, sleep disturbance appears to cause individuals to be more prone to pain; conversely, individuals with PP appear to be prone to developing sleep-related disorders (Smith & Haythornthwaite, 2004).

Lastly, the relationship between substance use disorders and PP, while complicated, appears to be alarmingly common. First, many of the factors that lead to increased risk for developing PP are also connected to substance use disorders (e.g., ACEs). The use of substances, while short-lived, is often a strategy for eliminating both emotional and physical pain. SAMHSA suggested a cycle of vulnerability where individuals in pain use a substance to alleviate their pain; after the effects of the substance wears off and the pain returns (often at higher levels), the individual uses the substance again. Given the human tendency to develop tolerance to substances, the impact of any substance will diminish, resulting in individuals using more of the substance to achieve results (i.e., dependence). This cycle complicates the patient and provider experience and interactions.

**TREATMENT FOR PERSISTENT PAIN**

Treatment for PP has evolved significantly over the past decade and there is greater emphasis on providing psychological interventions, patient education, alternative treatments (e.g., acupuncture) and rehabilitation in primary care settings as well as specialty settings. Delivery of nonpharmacological therapies that encompasses a team of providers is widely accepted, as these approaches have
been demonstrated to be effective in treating the plethora of concerns seen with PP (Dobscha et al., 2009; Kamper et al., 2015). Interventions include physical therapy, massage, family therapy, and various types of cognitive behavioral therapies. All of these innovations support a shift from working to eliminate pain to working to support patient improvement in functioning. In this section, information about Cognitive Behavioral Therapy (CBT), Acceptance and Commitment Therapy (ACT), Mindfulness-Based Stress Reduction (MBSR), and other several other psychological interventions will be presented. We will also discuss medication treatment briefly, as it will likely continue to be a part of pain treatment for a significant portion of patients.

Cognitive Behavioral Therapy

CBT has strong research support for a variety of conditions that fall under the umbrella of PP. Specifically, CBT protocols have shown to be effective in treating fibromyalgia, chronic low back pain, rheumatologic pain, and chronic headaches (Lumley, 2013). CBT treatments involve a multifaceted approach to chronic pain that addresses the entirety of the PP experience. Across PP related disorders, CBT focuses on educating the patient about the disorder, systematic development of self-management skills, initiation of lifestyle behavioral changes, and restructuring maladaptive cognitive schemas resulting from exposure to pain. Interventions may include behavioral activation, relaxation, and cognitive disputation, among others. As indicated previously, research regarding the effectiveness of CBT for PP is robust (Henschke et al., 2010; Richmond et al., 2015; Williams, Eccleston, & Morley, 2012). However, as is the case with many behavioral interventions, specialty protocols have yet to be adapted to the fast pace of primary care.

Acceptance and Commitment Therapy

ACT (said as the word) is a “third wave” CBT. The American Psychological Association Division 12 states that ACT has strong research support in treating general chronic pain and PP (Lumley, 2013). ACT encourages patients to disengage from their efforts to control and eliminate pain and to focus efforts on improving engagement in meaningful life activities. In the ACT approach, patients use exercises to learn a mindful perspective on pain, identify values, and develop skills to persevere in getting their lives back on track. ACT has been adapted to the primary care setting (Focused Acceptance and Commitment Therapy [FACT]; Robinson et al., 2010; Strosahl, Robinson, & Gustavsson, 2012), and dissemination studies are currently underway in the United States, Great Britain, Finland, and Spain.

The support for utilizing ACT with PP is robust and includes randomized controlled trials (Buhrman et al., 2013; Dahl et al., 2004; Thorsell et al., 2011; Wetherell et al., 2011; Wicksell et al., 2008; Wicksell et al., 2013), partially controlled trials (Johnston, Foster, Shennan, Starkey, & Johnson, 2010; McCracken, Vowles, & Eccleston, 2005; Vowles, Wetherell, & Sorrell, 2009), and effectiveness studies (McCracken & Gutiérrez-Martínez, 2011; Vowles & McCracken, 2008). The core processes of ACT protocols (e.g., acceptance, committed action to values, mindfulness) repeatedly have been shown to facilitate the positive clinical outcomes (e.g., reduction in perceived pain levels, improvements in overall functioning) observed in these studies.

McCraken, Sato, and Taylor (2013) reported on the usefulness of ACT interventions delivered in a brief (i.e., four-visit) group format to patients with PP receiving services in primary care clinics. The protocol of this group included “a combination of methods to promote psychological flexibility, including acceptance, cognitive defusion and values-based and committed action”
Seventy-three patients were enrolled in the study and randomly assigned to treatment or treatment as usual (TAU) conditions. Results showed the ACT group to significantly outperform the TAU group in measurements related to disability, depression, pain acceptance, and emotional functional. While the small sample size prohibits any generalization, the study is compelling because it shows ACT to be helpful for a variety of concerns related to PP (e.g., pain, emotional well-being, functionality) and it was achieved in a brief primary care group intervention.

In addition to ACT’s effectiveness in the clinical setting, a recent pilot study involving brain scanning suggests that ACT processes also influence brain activity related to pain. Smallwood, Potter, and Robin (2016) recently reported results in a study involving patients with PP and comorbid substance use disorders. Patients were randomized to patient education or ACT, and patients participating in the ACT intervention showed activity reductions, measured through fMRI, in areas of the brain that are believed to be associated with pain. While needing replication, this study suggests that ACT treatment impacts neurological and biological pain pathways.

**Mindfulness-Based Stress Reduction**

Mindfulness-based stress reduction (MBSR) is also a third wave CBT approach with a growing research base for treatment of PP (Cherkin et al., 2016; Cramer, Haller, Lauche, & Dobos, 2012; Marchand, 2012; Reiner, Tibi, & Lipsitz, 2013). Similar to ACT, MBSR focuses on improving patient acceptance and awareness of present moment experiences related to pain. Recently, Cherkin et al. (2016) completed a research project evaluating CBT and MBSR compared to TAU for 342 adults who had chronic low back pain. Both the CBT and MBSR treatment groups showed significantly greater improvement in pain reduction and overall functioning than the TAU group. There was no significant difference observed between the CBT and MBSR groups in regards to pain reduction and functioning, indicating MBSR may be an alternative to CBT for the treatment of chronic low back pain.

**Other Behavioral Interventions**

Compassion-Focused Therapy (CFT) and positive psychology are also other psychotherapeutic interventions that may be helpful in building a program to serve patients with PP. While the research concerning their use with patients with PP specifically is limited, their focus on functioning rather than symptom reduction appears consistent with the previously suggested approaches to PP with an evidence base.

**Compassion-Focused Therapy**

CFT is another third wave CBT. It incorporates components of developmental psychology, Buddhist practical philosophy, and evolutionary theory (Gilbert, 2009). As indicated in the name, CFT encourages individuals to develop a compassionate stance and flexibility in dealing with life challenges. Research supporting CFT and integration of mindful compassion practices in daily life has grown over the past decade. Hofmann, Grossman, and Hinton (2011) found compassion practices to be a useful additive to evidence-based treatments (e.g., CBT, ACT) for improving overall quality of life. Clinically, while still in its infancy, CFT has been shown to be effective in treating a variety of problems, including personality disorders (Lucre & Corten, 2012), psychosis (Laithwaite et al., 2009), and eating disorders (Gale, Gilbert, Read, & Goss, 2012).
While research utilizing CFT with individuals that have PP is lacking, research has shown CFT to be effective in reducing self-criticism, depression, anxiety, sense of inferiority, and shame in individuals who deal with chronic mental health problems (Gilbert & Procter, 2006). In a recent systematic review, Leaviss and Uttley (2015) concluded CFT to be a promising therapy in helping individuals who have a tendency to criticize themselves. It is reasonable to propose that these findings would generalize to patients who suffer from PP and the common comorbidities of anxiety and depression.

**Positive Psychology**

Positive psychology offers an alternative to the traditional focus of psychology on reduction of distress and amelioration of a disorder, and posits a focus instead on elements that support a healthy life or a sense of well-being. In introducing it, Seligman and Csikszentmihalyi (2000) described three levels of interest: the subjective, the personal, and the group level. While all levels are potentially of interest to building programmatic services to improve care to patients with PP, the subjective level is particularly inspiring of idea generation. Positive psychology, at this level, is about valued experiences, well-being, and content (in regards to the past), hope and optimism (in regards to the future), and flow and happiness (in the present) (Kashdan & Ciarrochi, 2013). Similar to CFT, positive psychology focuses on changing the way a person relates to pain as a past, present, and potential future experience and encourages involvement in meaningful life activities. While large RCTs are needed to evaluate the effectiveness of using positive psychology with PP, recent research supports its use to improve quality of life with patients with PP (Flink, Smeets, Bergbom, & Peters, 2015; Muller et al., 2016).

**Medications**

Even with advancement in development of medications for pain and further understanding of pain neurological pathways, most patients seeking treatment will experience only a 30% reduction in pain levels (Turk, Wilson, & Cahana, 2011). Unfortunately, using a substance to treat pain often results in increased use of the substance and no long-term pain relief. The inability to reduce an individual’s pain all too often leads to overprescribing of pain medications and this, of course, fuels the opiate epidemic (Dowell, Haegerich, & Chou, 2016). The Centers for Disease Control and Prevention (CDC) found that 165,000 human deaths were related to opioid pain medication overdose between 1999 and 2014 (Dowell et al., 2016). In 2014 alone, 18,893 accidental overdoses were attributed to prescription pain relievers (CDC, 2016). Further, four out of five heroin users started out using prescription painkillers—not heroin (Hedegaard, Chen, & Warner, 2015). Either directly or indirectly, it appears the over-reliance on opioid medications to treat PP has contributed significantly to the opioid/heroin epidemic in the United States, and now more than ever it is urgent that we make psychological treatments available to primary care patients with pain.

**THE HEALTHY BRAIN AND BODY TOOL KIT**

This section offers a variety of tools to support implementation of an array of new services to patients with PP. ACT is introduced in more depth, and then assessment options are reviewed. Scripts to promote new conversations about pain are suggested, and a care plan amenable to support by multiple members of the patient-centered primary care home is introduced. Materials for a group curriculum to support patient care plans are also provided. This how-to section concludes with the provision of tools for engaging PCPs in change and enhancing their resilience in practice.
ACT in Theory and Practice

As mentioned earlier in the brief introduction to ACT and the evidence for its application to the problem of PP, ACT is a newer cognitive behavioral therapy (Hayes, Strosahl, & Wilson, 1999, 2004). In addition to being helpful with pain, ACT is effective with medical problems experienced by patients with PP (e.g., tobacco dependence and obesity) as well as psychological problems (e.g., anxiety and depression). ACT involves the promotion of mindfulness, acceptance, and commitment practices and behavior change processes to human suffering. Pain is pain and life is painful, in different ways at different times, and in different contexts. Suffering is what is added on and this “extra” is targeted by ACT in treatment of PP. ACT encourages patients to take an active approach to living a meaningful life by making gains in core skills underlying psychological flexibility.

ACT interventions address cognitive fusion, experiential avoidance, and psychological flexibility. Cognitive fusion refers to the problem of being unable to separate one’s sense of self from one’s emotions, thoughts, and feelings. For example, a patient suffering with back pain may state, “This pain is killing me,” and then shift in his seat as if attempting to avoid being stabbed or injured. A more “defused” approach would be to describe the pain: “There is pressure and heat in my low back at this point.” A defused response to painful sensations provides better support for initiation of behaviors (e.g., diaphragmatic breathing, brief self-compassion meditation) that change what is happening in the nervous system.

Experiential avoidance refers to using cognitive, emotional, and/or behavioral strategies to avoid direct experience of unpleasant feelings, thoughts, and sensations. An example is the taking of pain medications to anticipate and avoid pain sensations or thoughts about pain, such as “I’m going to pay for this; I better take a pill; I don’t think the doctor understands—he doesn’t give me the pills I need.” Experiential avoidance often provokes behavioral avoidance. An example of behavioral avoidance is lying down much of the day to avoid triggering greater pain by moving too much or in the wrong way.

Psychological flexibility is the ability to choose a direction and behave in the world in ways that are consistent with that direction despite experiencing unwanted thoughts, feelings, and behaviors that are contrary to that direction. A patient with PP that attends a child’s sports game and continues to view the game, even with the appearance of more pain sensations, thoughts about the dangerousness of pain, and the emotion of fear, is demonstrating a high level of psychological flexibility. Figure 24.2 provides a schematic of the Healthy Brain and Body Core Processes. It identifies six core processes (being present, connect with values, engaging in committed action, flexible perspective taking, diffusion and acceptance) that work together to promote psychological flexibility. Table 24.1 depicts the Pillars of Psychological Flexibility, which result from collapsing the six processes into three pillars or response styles (open, aware, engaged). While the hexaflex is quite useful in understanding and intervening with patients with PP, the Three Pillars approach, suggested by Focused Acceptance and Commitment Therapy (FACT; Robinson, Gould, & Strosahl, 2010; Strosahl, Robinson, & Gustavsson, 2012), offers a more focused approach to promoting psychological flexibility in interactions with patients, as it is adapted for use in the brief treatment context of primary care. FACT supports efficiency in conceptualizing and intervening on an ongoing basis with patients over time.

Table 24.2 provides clinician guides for influencing the core processes of open, aware, and engaged in an effort to promote greater psychological flexibility in interactions with patients with PP. Table 24.3 provides a poster guide of similar information and will be useful for clinicians while learning these new therapeutic interventions. Conversations with patients can often
Figure 24.2 Healthy Brain and Body Core Processes
<table>
<thead>
<tr>
<th>Open</th>
<th>Aware</th>
<th>Engaged</th>
</tr>
</thead>
<tbody>
<tr>
<td>Able to engage in important actions and “just notice” when difficult thoughts, feelings, and sensations come up</td>
<td>Aware of the present moment, uses senses to stay in the moment</td>
<td>Able to describe values and talk about them, with sense of caring</td>
</tr>
<tr>
<td>Attitude of openness and curiosity about thoughts, feelings, and sensations that used to be troubling</td>
<td>Able to refocus attention on the present</td>
<td>Able to identify small goals that are consistent with values</td>
</tr>
<tr>
<td>Able to see that thoughts, feelings, and sensations come and go, allowing them</td>
<td>Able to see different perspectives on self and others and shift between them</td>
<td>Able to implement behavior change experiments</td>
</tr>
<tr>
<td>Open</td>
<td>Aware</td>
<td>Engaged</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>If “stuck,” be curious, model acceptance, notice “mind”</td>
<td>If confused, go to the present/direct attention to senses</td>
<td>If you see an opportunity, promote connection to values and action</td>
</tr>
<tr>
<td>Support openness and curiosity about previously avoided inner experiences</td>
<td>Promote flexible, voluntary, and purposeful attention to the now</td>
<td>Move toward identifying ongoing qualities of action that are meaningful in the here and now</td>
</tr>
<tr>
<td>Attend to thinking as an ongoing process, rather than the world structured by it</td>
<td>Support detection of different possible perspectives on self and others</td>
<td>Construct concrete behavior change exercises, follow up on specifics</td>
</tr>
</tbody>
</table>

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
## Table 24.3 Clinician Guide to Strategies to Promote More Open, Aware, and Engaged Responding

<table>
<thead>
<tr>
<th>What to Do When</th>
<th>Focus on</th>
<th>Strategies to Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversation with patients feels “stuck”</td>
<td>OPEN</td>
<td>1. Model acceptance (&quot;Let’s just sit still and let this play out.&quot;)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Notice “mind” (e.g., “What is your mind saying now?”)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Be curious about previously avoided inner experiences (“Is your ‘critic’ showing up again?”)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Attend to thinking as an ongoing process, rather than the world structured by it (“Can you just watch your thoughts, kind of like railroad cars moving by slowly?”)</td>
</tr>
<tr>
<td>Clinician feels “confused”</td>
<td>AWARE</td>
<td>1. Go to the present/direct attention to senses (&quot;What are you feeling in your body right now?&quot;)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Promote flexible, voluntary and purposeful attention to the now (&quot;Can you direct your attention to the emotion you are experiencing and how it is affecting you right now?&quot;)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Support detection of different possible perspectives on self and others (&quot;Can you for a minute imagine you are looking back at this moment, say a month from now?&quot;)</td>
</tr>
<tr>
<td>If you see an opportunity, promote connection to values and action</td>
<td>ENGAGED</td>
<td>1. Ask directly about values (&quot;What matters most to you?&quot;)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Identify ongoing qualities of action that are meaningful in the here and now (&quot;Will you for a moment tune into how it feels to love your partner/child?&quot;)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Construct concrete behavior change exercises, follow up on specifics (&quot;So, let’s figure out a specific behavior plan that you can implement to see if it connects you with an important value.&quot;)</td>
</tr>
</tbody>
</table>
feel like they are going nowhere (i.e., “stuck”), and this is a good time for clinicians to promote more openness, in themselves and in the patient. Simply modeling acceptance and acknowledging the challenge by noticing what’s going on in the mind can get an interaction back on track and promote opportunities for positive change. All clinicians feel bewildered (“confused”) at times, and when this occurs it is good to go to the center pillar and promote more awareness—what is going on in the present moment? Here, also, clinicians can support a flexible shift in the patient’s perspective in regard to self as well as others. The clinician needs to take every opportunity to go to the engaged pillar and these occasions occur when the clinician sees the possibility for promoting behavior change (“opportunity”). Strategies in this pillar include promoting a stronger connection with values and building skills and behavioral experiments to improve vitality. A better life is lived over time in many behavioral moments, one chosen action after another.

Assessment
Assessment is a cornerstone of all CBTs and is no less important in the FACT approach to PP. Three strategies are suggested, all of which are in the public domain and are feasible in the brief context of primary care. This is not a complete list, but instead is meant to be illustrative.

CDC Healthy Days Core Measure
This adaptation of the CDC Health Days Core Measure is helpful in assessing overall physical and mental health and in evaluating change over time. It can be used with people aged 12 and older. Figure 24.3 provides a list of the Healthy Days questions. This measure is often useful in monthly individual or class contacts with patients with PP. It is brief, thus supporting both the need for and efficiency in assessment. Patients often are interested in their progress, and the simplicity of this assessment helps them track their progress.

PEG
The PEG is a three-item scale that assesses pain intensity and interference (Krebs et al., 2009). PEG is an acronym, where “P” refers to “pain” intensity, “E” to interference in “enjoyable” activities, and “G” to interference in “general” activities. Similar to the Healthy Days, the PEG is brief and feasible in the primary care context. It may be helpful in transitioning patients and clinicians to new conversations about pain, as its questions concerning pain intensity may serve as a bridge to the more urgent need to help patients engage in enjoyable activities and maintain a good activity level, even with pain. The PEG is also valuable in gauging pain volatility in patients, and this is important because variability in pain may explain poor prescription opioid addiction treatment outcomes in persons with chronic pain. Specifically, one study found that patients with greater volatility in subjective pain during treatment have increased risk of returning to opioid use by the conclusion of an intensive treatment with buprenorphine/naloxone and counseling for 12 weeks in an outpatient setting (Worley, Heinzerling, Shoptaw, & Ling, 2015). Figure 24.4 provides a listing of the PEG questions.

Satisfaction
Measurement of patient and clinician satisfaction is an important beginning point in improving care to patients with PP, and an important repeated measure as it is perhaps the most amenable to change and can thereby energize change efforts. Figure 24.5 provides a copy of the Satisfaction Survey for
1. Would you say that in general your health is
   a. Excellent
   b. Very good
   c. Good
   d. Fair
   e. Poor
2. Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?
3. Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?
4. During the past 30 days, for about how many days did poor physical or mental health keep you from doing your usual activities, such as self-care, work, or recreation?

Scoring Directions and Interpretation
Questions 2 and 3 are combined to calculate a summary index of overall unhealthy days, with a maximum of 30 unhealthy days. Healthy Days are calculated by subtracting the number of unhealthy days from 30. An Able Days score can be calculated by subtracting the response to question 4 from 30. Share the results with patients and explain that the goal is to increase Healthy and Able days over time.
1. What number best describes your *pain on average* in the past week?
   1  2  3  4  5  6  7  8  9  10

2. What number best describes how, during the past week, pain has interfered with your *enjoyment of life*?
   1  2  3  4  5  6  7  8  9  10

3. What number best describes how, during the past week, pain has interfered with your *general activity*?
   1  2  3  4  5  6  7  8  9  10

**Scoring and Interpretation**

Clinicians may explain that pain levels are expected to vary over time and that the goals of treatment with persistent pain are to lower the interference scores, such that more enjoyment and general activity are present in the patient’s life.

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
The purpose of this survey is to further understanding of patient experience in receiving care for pain. We plan to ask these questions again periodically to assess the impact of our efforts to improve services to you. Thank you for completing this survey.

**Patti Robinson, PhD**

Today's Date: _____________________

Below you will find a list of statements. Please rate the truth of each statement as it applies to you. Use the following rating scale to make your choices. For instance, if you believe a statement is “Always true,” you would write a 5 in the blank next to that statement.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
<th>Example Statement</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Never true</td>
<td>My providers understand that my pain is real.</td>
</tr>
<tr>
<td>1</td>
<td>Very rarely true</td>
<td>My providers care about me as a person.</td>
</tr>
<tr>
<td>2</td>
<td>Sometimes true</td>
<td>At every visit, my providers show concern about my quality of life.</td>
</tr>
<tr>
<td>3</td>
<td>Often true</td>
<td>My providers have new ideas about how to help me.</td>
</tr>
<tr>
<td>4</td>
<td>Almost always true</td>
<td>My providers have many resources available to help them help me.</td>
</tr>
<tr>
<td>5</td>
<td>Always true</td>
<td>My services are helping my pain problems improve, slowly and surely.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I know how to help my provider help me.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>My provider and I agree about the nature of my pain problems.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>I would recommend the services I receive at my clinic to other patients with pain.</td>
</tr>
</tbody>
</table>

*Figure 24.5  Satisfaction Survey for Patients With Persistent Pain*
Patients With Persistent Pain. It includes nine questions for use in assessing patient satisfaction. The Satisfaction Survey for Primary Care Clinicians (PCC) Providing Services to Patients With Persistent Pain is provided in Figure 24.6. These are program evaluation measures and clinical experience suggests that they are acceptable to patients who read at the sixth grade level or above, and both sensitive to and supportive of change in delivery of services and experience in receiving and providing services.

**Lifestyle Risk Factors**

The level of debilitation patients experience with PP may be exacerbated by numerous lifestyle factors, including use of tobacco products, depression, sleep, and the use of alcohol. Tobacco use can worsen chronic pain patients’ levels of pain, as well as overall functioning (Weingarten et al., 2008). Use of tobacco products among patients with chronic pain is also associated with greater use of opioids (Hooten, Shi, Gazelka, & Warner, 2011). Symptoms of depression are common among patients with chronic pain. While available data do not address causation, patients diagnosed with major depressive disorder have been found to be four times more likely to experience disability with PP than those without the problem of chronic pain (Arnow et al., 2006).

Among patients with PP, poor sleep, use of alcohol, and being overweight are also associated with poorer quality of life (Arvidsson, Arvidsson, Fridlund, & Bergman, 2008). While data assessing the directional relationship between obesity and chronic pain are not available, the relationship between weight and chronic pain is documented in a number of disorders, including low back pain and osteoarthritis (Janke, Collins, & Kozak, 2007). Janke and colleagues hypothesize that as an individual gains weight, there is increased structural demand, which may result in bone and joint pain. Obesity is also related to metabolic problems (e.g., diabetes), which may also result in neuropathy and other pain-related complaints. Weight gain may additionally cause or increase pain as a result of decreased mobility and reduced activity (Janke et al., 2007). Given these identifiable lifestyle risk factors for patients with PP, we recommend that routine assessment include identification of risks and formulation of a plan to address them. Figure 24.7 provides a NEEDS Lifestyle Checklist (Robinson, Bauman, & Beachy, 2016).

**A New Conversation About Pain**

Yesterday’s conversations about pain led to unforeseen problems for patients with PP and vulnerability to dependence on opiate medications and for prescribing providers. Before pain became the “fifth vital sign,” it was a symptom and often seen in the context of one’s life, a natural result of engaging in repeated activities to meet basic needs. Farmers had back pain. Secretaries had wrist pain. Pain was not a vital sign but a part of life and possibly relieved by changes in behavior. Vital signs are heart rate, blood pressure, respiratory rate, and temperature. These are reliably measured and may be indicators of disease. Pain, as a symptom, is not a disease.

Pain came to be known as the fifth vital sign in the late 1990s, and this perspective evolved into a Joint Commission standard in 2001. At the time, there was an allegation of undertreatment of pain, along with a new idea that all pain must be closely monitored and treated—treated with pain medications that presumably could render patients pain free. During the decade following pain’s emergence as a vital sign, pain management clinics sprung up all over the United States, rates of opioid prescribing soared, and rates of death from use of opioids also climbed, touching families and communities of all sizes. In this troubling context, prescribers and patients alike have struggled and found it difficult to begin anew.
The purpose of this survey is to further understanding of Primary Care Clinician and Nursing experience in providing care for persistent pain. We plan to ask these questions periodically to assess the impact of our efforts to improve services and clinician experience. Thank you for completing this survey.

Are you a PCC?  Yes  No  Are you a nursing staff member?  Yes  No

Today’s Date: ______________________________

Below you will find a list of statements. Please rate the truth of each statement as it applies to you. Use the following rating scale to make your choices. For instance, if you believe a statement is “Always true,” you would write a 5 in the blank next to that statement.

Never true  Very rarely true  Sometimes true  Often true  Almost always true  Always true

1. My training prepared me adequately for working with chronic pain patients.
2. I enjoy working with chronic pain patients.
3. I have all the skills I need to work effectively with chronic pain patients.
4. I look forward to seeing chronic pain patients.
5. I feel that I am successful with chronic pain patients.
6. I want to specialize in treating chronic pain.
7. I usually have a new idea about how to help my most difficult chronic pain patients.
8. Pain medications are very helpful to my chronic pain patients.
9. I am able to refer my chronic pain patients to accessible, effective programs.

Do you have any ideas about how to improve our current Pain and Quality of Life Pathway program?

Yes  No

If yes, what?

➔ Figure 24.6 Satisfaction Survey for Primary Care Clinicians (PCC) Providing Services to Patients With Persistent Pain

© 2017, Handbook of Psychological Assessment in Primary Care Settings, Edited by Mark E. Maruish, Taylor & Francis.
Your quality of life and experience of pain may improve with a few changes to your lifestyle. This is a checklist that invites you to identify areas you might want to make a change in now, or at some point in the future.

<table>
<thead>
<tr>
<th>Lifestyle area</th>
<th>Possible Changes</th>
<th>Interested in Discussing a Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition</td>
<td>Eat better, eat less, eat more slowly, eat smaller amounts more often</td>
<td></td>
</tr>
<tr>
<td>Exercise</td>
<td>Going for short walks on a daily basis, learning stresses that reduce tension</td>
<td></td>
</tr>
<tr>
<td>Enjoyment</td>
<td>Make a list of things that make you smile, think of people you love and make a plan to get in touch</td>
<td></td>
</tr>
<tr>
<td>Don’t Smoke or Drink</td>
<td>Cut down or commit to not increasing how much you smoke or drink. If ready, set a quit date. Find a buddy.</td>
<td></td>
</tr>
<tr>
<td>Sleep</td>
<td>Eight hours is a goal. Learn to relax in bed, even if you are not sleeping. Develop a wind-down routine. Make your bedroom inviting for sleep.</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Table 1. The NEEDS Approach to Changing Lifestyle Behaviors in Patients With Persistent Pain (Robinson, Bauman, & Beachy, 2016).

**Figure 24.7** NEEDS Lifestyle Checklist for Patients With PP
Discussions about PP often involve too much finger-pointing and not enough creative thinking, so perhaps our best way out of the unproductive and sometimes health-harming conversations of yesterday is to “drop the rope” and start anew. One needs to identify what has not and is not working and open up to new concepts and new treatment. It is in this spirit that we humbly offer a new formulation of PP and guides for new ways of interacting. Key words in this discussion include brain and body health, brain change, and “chipping away” at change. These concepts are integrated into The Healthy Brain and Body Plan in Figure 24.8. This plan is intended as a patient previsit questionnaire. It is designed to support a new conversation, one that will inform a strong care plan.

**Brain and Body Health**

The application of the concept of central sensitization to pain with a peripheral nervous system basis offers us a helpful way of understanding pain. At the risk of oversimplification, one might say that what happens in the peripheral nervous system affects the central nervous system. This means that insults and injuries to the peripheral nervous system may cross over to the central nervous system and result in problems, just like a stroke might. Figure 24.9 provides a discussion tool for clinicians to use in explaining the concepts of central and peripheral to patients, as a part of a visit spent reviewing the patient’s responses to the Healthy Brain and Body Plan (Figure 24.8). Various uncontrollable circumstances, such as a history of trauma, may make some people more vulnerable to this crossing over. Including some elements of this conceptualization in a discussion with a patient enhances their understanding of brain and body health and lessens the chance of their feeling stigmatized because of their health challenges.

In the first section of the Healthy Brain and Body Plan (Figure 24.8), the patient is invited to indicate what he thinks causes his pain, as well as ideas about what is wrong with his body. This helps the clinician to understand the patient’s worldview as it relates to pain and draws out cognitions (thoughts, beliefs, values, etc.) that may add to the patient’s overall sense of suffering. Often, patients will identify beliefs about the cause of pain that are inaccurate and/or thoughts that evaluate the pain in an excessively negative way. In introducing the concept of central and peripheral pain, the clinician may be able to add to the patient’s conceptualization and promote development of a new perspective, one that is more curious about the brain–body connection and one that is more accepting of the pain experience.

**Brain Change: Neuroplasticity**

Neuroplasticity or brain plasticity is an umbrella term that refers to the ability of the brain to change throughout the lifetime of an organism. The process of neuroplasticity allows neurons in the brain to adjust neural activities in responding to new situations or changes in their environment, and to compensate for injury or disease. Reorganization in the brain takes place in many ways, such as that of “axonal sprouting,” which involves undamaged axons growing new nerve endings to connect neurons whose links were lost as a result of injury. Further, it is possible to engineer beneficial brain changes by providing repeated stimulation from behavior changes designed to heal the brain.

Many patients will resonate to a conversation about “brain training.” Most people are familiar with meditation or mindfulness and may already know about the health benefits of such. However, many patients with PP may not know how to begin brain training. The questions in the Brain Change section of the Healthy Brain and Body Plan are intended to identify activities that the patient can use to start experimenting with brain change. For patients interested in shifting perspective, the clinician might discuss the difference between the perspective of a participant on a roller coaster ride in comparison to the perspective of a person watching the roller coaster ride from 100 feet away, and
Healthy Brain and Body Plan

Dear Patient,

Thank you for answering these questions. Your answers will help us make a plan together, a plan to improve your health, one step at a time. We need to look at all parts of health—mind, body, and behavior.

Brain and Body Health

Injuries and the process of aging can hurt or brain and body health. Your answers to these questions will help us understand your views in this area.

1. What do you think causes your pain?

2. What do you think is wrong with your body?

Brain Change

Research is revealing more and more about the way pain in the body and pain in the brain connect with each other. Now we know that we can change pain in the brain by shifting perspectives. We also know that we can change pain in the body by changing what’s happening in our brain. Finally, we can also change what is happening in the brain and the body by changing our behavior.

1. What do you do that helps you change your perspective or view?

2. What do you do that helps to create a calm state?

3. What do you do to create a sense of joy or feeling of love?

Chipping Away at Change

It is very difficult to change our brain and body because it is hard to keep behaving in new ways and easy to keep behaving in old ways, even if the old ways don’t pay off. We need to use our values to inspire us. We also need partners to help us start and continue behaviors that will give us better brain and body health over time.

1. What new behaviors would you like to start to improve your health?

   Mind
   - Learning to focus on the present moment
   - Learning to use descriptive words more and negative judgments less
   - Learning to “notice” rather than get wrapped up in sensations
   - Building skills for mentally caring about or loving others
   - Building skills for mentally caring about or loving myself

   Other ideas:
   - Learning to relax
   - Learning to settle when upset
   - Learning to be in my body more, more aware
   - Learning to tune in to when my body feels good
   - Learning to sleep better
   - Learning to tune into what foods are good for my body
Other ideas:

Behavior

- Learning to plan and do activities that give me pleasure
- Learning to schedule and keep track of my behavior
- Learning to ask others for help when I need it
- Making more of an effort to help others
- Doing creative things more often

Other ideas:

2. Who will support you in making changes?

3. What values inspire you to try new behaviors?

Figure 24.8 (Continued)
Figure 24.9 Central and Peripheral Nervous Systems
then apply this to ways of experiencing pain. Further, the clinician might suggest that the patient use the taking of a slow, deep breath as a cue to shift from the front of the roller coaster to a back seat or even from the line for the Ferris Wheel. The brain reorganizes itself in response to illness or injury by forming new neural connections. A good book for patients interested in further developing skills related to mindfulness and flexible perspective taking is *In This Moment: Five Steps to Transcending Stress Using Mindfulness and Neuroscience* (Strosahl & Robinson, 2008). For patients interested in learning more about strategies for promoting a sense of calm, the clinician might use the CALM technique (available at http://behavioralconsultationandprimarycare.com) or recommend one of the numerous Internet applications available for training the brain to more readily experience a sense of well-being and calm.

The third area on the Plan attempts to tap into the patient’s experience of safety and love, for self and others. Compassion, for others and self, is of fundamental importance for many patients with PP. For patients interested in this area, we recommend that the clinician suggest 5–10 minutes of compassion-focused meditation daily. This approach involves repeating phrases of positive intention, perhaps beginning by directing these phrases to a good friend and then directing them toward the self. Phrases associated with compassion-focused therapy include: “May you be safe . . . May you experience physical happiness . . . May you experience mental happiness . . . May you have ease of well-being.”

**Chipping Away at Change**

The recommended cognitive-behavioral approach involves the traditional elements of all CBTs, including education, data collection, behavioral activation, and working with cognitions. In that it is a third wave approach or a newer CBT, mindfulness and committed action inspired by values is emphasized. The last section of the Health Brain and Body Plan, “Chipping Away at Change” (Figure 24.8), invites the patient to identify one or more change directions in one of three areas (Brain, Body, Behavior), to identify a person supportive of their learning, and to identify core values that dignify the difficulty of initiating and persevering in behavior change. The results of a conversion about the patient’s responses to these questions can inform development of a Healthy Brain and Body Care Plan.

**CREATING CHANGE: SKILLS FOR A HEALTHY BRAIN AND BODY**

Providing consistent, compassionate support for patient change over time is the heart of realizing better outcomes for patients with PP. In this section of the tool kit, we offer two tools that support this effort. The first is the Health Brain and Body Care Plan and the Second is the Health Brain and Body Class.

**The Healthy Brain and Body Care Plan**

The Healthy Brain and Body Care Plan is an ACT-consistent tool for clinicians to use to establish treatment goals, support specific behavioral experiments, track outcomes, and teach skills of importance to the individual patient. This plan is informed by the Bull’s Eye Plan (Robinson, Strosahl, & Gould, 2010), a tool that is useful for primary care patients with a variety of problems, including pain, depression, and chronic diseases. Figure 24.10 provides a copy of the Health Brain and Body Bull’s-Eye
Care Plan. It can be used as a visual aid in talking with patients. Table 24.4 provides directions for using this care plan. For an in-depth discussion of this approach and case examples illustrating it, see Real Behavior Change in Primary Care: Improving Outcomes and Increasing Job Satisfaction (Robinson, Gould, & Strosahl, 2010).

The Healthy Brain and Body Class

This class format uses the Bull’s Eye Plan to support behavior change and systematically teaches skills that promote psychological flexibility. The class may be offered as a class series or used in a group medical format, with meetings occurring on a monthly basis without a planned end. The group medical visit format will help clinics realize increased efficiency in assessing outcomes and teaching skills and provide patients with opportunities for socializing and supporting each other. In either format, it is best to allow for new participants at any class.

Table 24.5 provides the agenda for a typical class, but note that the agenda can be adjusted to take advantage of skill-training opportunities that arise in an ongoing group medical visit class (e.g., teaching skills in regards to context of an upcoming holiday experience). Because there may be new participants, the class needs to begin with a brief explanation of the class purpose (i.e., helping patients develop healthy brains and bodies), the psychoeducational format, and introductions. It is often useful to suggest a format for introductions, such as “first name and something that made you smile recently.” Assessment, using one or more of the outcomes suggested earlier, follows and then the topic of the day can be introduced. Classes need to include instruction in at least one skill and participants need to have a chance to practice this skill and then develop a plan to apply it in some type of behavioral experiment. The class will then shift to patients’ review of their ongoing Bull’s Eye work and the results of the behavioral experiments developed in the preceding class. It is important to allow time for participants to ask questions and discuss results of previous experiments.

There are typically six topics and each class covers one topic. The topics are the core processes identified in Figure 24.2, with “Being Present” as the first in the series and “Values Connection” being second. The topic order proceeds in a clockwise fashion around the Healthy Brain and Body hexaflex. Table 24.6 provides a listing of class topics, along with suggested exercises for each topic.

Brief descriptions of each of the skills listed in Table 24.6 are provided in the following section. For more information on suggested skill development exercises and additional exercises, read Real Behavior Change in Primary Care (Robinson, Gould, & Strosahl, 2010) or Brief Interventions for Radical Change (Strosahl, Robinson, & Gustavsson, 2012).

Topic 1: Being Present

Being present is challenging for all humans, but for patients with PP it is particularly difficult as it opens the door to closer experience of pain. The class teacher needs to introduce the present moment gradually, beginning with an invitation to experience the present moment by directing attention to sensory experience. Typically, one can start with hearing, as this is an easier focus for most people. After encouraging the participants to settle into a comfortable position and find a focus for their eyes two to four feet in front of them (or close their eyes if preferred), ask the participants to notice three sounds—one close, one at a distance, and a third more distant. The idea is to experience the sound, shift to the next, and so on, and eventually to have a simple awareness of all three. From there, the teacher can move on to another sense, such as vision, smell, touch, or taste. This exercise usually
Bull’s-Eye Care Plan

Values Description
(Love, Work, Play)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Consistency</td>
<td>Higher Consistency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Action Plan:

➔ Figure 24.10 Healthy Brain and Behavior Bull’s-Eye Care Plan
Table 24.4 Directions for Using the Bull’s-Eye Care Plan

1. Ask the patient to choose Love, Work, or Play as a focus for a short discussion about values. Ask the patient to explain what is important to him or her in each area of life. Look for patient to experience emotion during the talk about values, as this will help deepen the patient’s sense of engagement.

2. Listen closely, reflect what you heard, and then write a statement on the Bull’s-Eye Plan using the words (global, abstract) the patient used in talking about the value.

3. Explain that the Bull’s Eye on the target represents the patient’s hitting her/his value target on a daily basis (and explain for most of us fall far short of that on a day-to-day basis, but knowing the target and sharing it with others helps us make choices, set goals, and implement plans).

4. Ask patient to choose a number to represent how close to the Bull’s-Eye value statement her/his behavior has come over the past two weeks. This is the consistency score; it’s a baseline score in the first visit and then useful in assessing the results of behavior experiments from visit to visit.

5. Work with the patient to develop two specific behavior experiments (action plans) for the weeks between visits. Choose the experiments based on the patient’s belief that the action or behavior would bring her or him closer to important values identified in the Bull’s Eye or represent a move closer to the target.

6. At follow-up, ask the patient to identify barriers to engaging in planned behaviors and, if the patient was able to engage in the planned behaviors, to evaluate the extent to which they were value-consistent.

7. In initial care planning visits (if time allows) or in follow-up visits, teach a skill to help the patient improve skills for accepting distressing thoughts about pain (OPEN), being more skillful in finding the present moment and taking an observer perspective on pain (AWARE), or clarifying and connecting with important life values and lifestyle behaviors consistent with those values (ENGAGED).
Table 24.5 Healthy Brain and Body Class Agenda

1. Welcome and introductions.
2. Complete assessments.
3. Introduction of class topic and skill-training exercise.
4. Skill practice with a partner and development of a behavioral experiment involving use of new skill.
5. Record results of behavioral experiment from past class on Bull’s Eye document.
6. Discussion of previous behavioral experiments, questions, comments.
<table>
<thead>
<tr>
<th>Class</th>
<th>Topic</th>
<th>Skill Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>Being Present</td>
<td>Three Senses, Body Scan, Lake and Mountain Meditations</td>
</tr>
<tr>
<td>II.</td>
<td>Values</td>
<td>Love/work/play Bull's Eye, Tombstone</td>
</tr>
<tr>
<td>III.</td>
<td>Action With Commitment</td>
<td>Response-able, Start again</td>
</tr>
<tr>
<td>IV.</td>
<td>Flexible Perspectives</td>
<td>Self-stories and stories about others, chapters in a book</td>
</tr>
<tr>
<td>V.</td>
<td>The Defuse Difference</td>
<td>Sticky notes, pain-pain-pain version of happy birthday song</td>
</tr>
<tr>
<td>VI.</td>
<td>Accept Difficulties</td>
<td>Driving the bus as a physical metaphor exercise, compassion practice</td>
</tr>
</tbody>
</table>
takes three to five minutes, and at its conclusion the teacher can facilitate a discussion about what it was like to be in the present moment.

Another exercise that’s useful in promoting experience of the present moment is the body scan. It is often good to start with the feet and proceed toward the head, pausing at each area of the body to discover sensation. The teacher will want to promote a spirit of discovery and can perhaps do some training on the difference between evaluation or judgments and description prior to the start of the exercise. Then, participants can focus on describing sensations that arise as they scan their bodies, accepting and “just noticing” evaluations that will probably arise spontaneously, all the while continuing to attend to the body and its unending work to promote our health.

Several other favorite exercises for promoting present moment include the Lake meditation and the Mountain meditation developed by Jon Kabat-Zinn (1994). Teachers can use one of these on the present moment class day in a group medical visit. Most patients enjoy these activities and request repetition of them at meetings over time.

**Topic 2: Values**

When patients are disconnected from their values, they are unable to use them to guide their actions. Instead, patients will follow socially inculcated rules that substitute for personally meaningful, chosen life goals. Patients with learning histories heavily based in rule-following can easily lose sight of what they want their lives to be about, particularly if they are immersed in complicated life situations—like PP, which is impossible to manage by following overly general rules (e.g., “Your doctor should make it go away”). For many patients with PP, even if their pain miraculously disappeared, they wouldn’t know where to go with life, so clarification and strengthening connection with values is of fundamental importance.

Values are global abstract concepts that serve to motivate meaningful action. Most patients are willing to talk with medical clinicians about their values, and talking about values often deepens the relationship between the clinician and patient. Activities that enhance a patient’s clarity regarding important values may rectify the suffering that arises from weakened connections to personal ideals. The teacher can use one of two interventions to help patients form a stronger bond with their important values and identify discrepancy between current behaviors and valued directions.

The Bull’s Eye Care Plan is often a core part of ongoing interactions between PCPs and patients. This intervention has four components: (1) values identification, (2) identification of value discrepancies, (3) ongoing design and implementation of behavioral experiments to increase consistency between daily behaviors and valued directions, and (4) identification and remediation of barriers to implementing behavioral experiments. In introducing the Bull’s Eye in the class, the teacher can use a handout similar to the one provided in Figure 24.10 and follow the directions provided in Table 24.4. Once introduced, the patient handout can be used at every class, as it offers the patient an opportunity to reflect on past experiments and note results and then to plan future experiments. In group medical visit classes, where care is ongoing, the Bull’s Eye class worksheet can be saved between classes by the teacher and passed out at the beginning of the next class.

In introducing the Bull’s Eye, the teacher can explain that values are global, abstract concepts about what matters most to us in life, stating, “Values are like the bull’s-eye on a target; we don’t usually hit the bull’s-eye in a game of darts, but it gives us a direction. I’d like to know what your bull’s-eye, or value, is in one of three areas today; you choose. Would you like to look at your values concerning (1) love and loving relationships, (2) work or study or your connection with community, or (3) play—what you do to just have fun and feel alive?” The teacher can then ask the patient to
note the area of values focus on the handout and write key words to describe her values. The teacher can facilitate a discussion among participants concerning activities they already do that seem to bring them closer to the chosen values and ideas about other possible behaviors to increase that connection.

The teacher can then ask the participants to make a mark on the bull’s-eye target to indicate how consistent her behavior has been with her stated value over the past few weeks. Total consistency would be the bull’s-eye. It’s important to explain to the patient, “Most of us are not hitting the bull’s-eye but, rather, coming in somewhere out here” (pointing to one of the most distant rings). Sometimes patients respond, “Not even on the page.” When this happens, let the patient know it’s okay and that the point of the bull’s-eye is to create a focus so we can be more intentional in our day-to-day choices.

The Tombstone is another exercise designed to build the patient connection with values. The teacher may introduce it by saying, “I want to better understand your values, what matters most to you in life, and a good way of getting at this is for you to imagine that you are at the end of your life and being laid to rest. What do you hope your loved ones will put on your tombstone? What would they say about you and how you lived your life?” The teacher may want to use a worksheet for this, so that the participants can jot down their thoughts. It is also helpful to have patients share this with a learning buddy in the close and to have a one-on-one discussion. The teacher will want to link this activity with the process of discovering discrepancies and planning behavioral experiments to bring greater consistency, as with the Bull’s Eye exercise.

Topic 3: Action With Commitment

Sustaining value-consistent action relies on patients being able to acknowledge their ability to respond differently to troubling thoughts, feelings, and sensations and to troubling problems of living encountered in the outside world. The teacher will need to assist patients with building a large range of skills that promote sustained changes in behavior. Patients with PP often need assistance in building skills related to solving problems, planning with intention, and building stronger social relationships. Two class exercises are recommended as starting points in this important area. Both involve enhancing patient understanding of what choice means. The teacher, other participants, and the patient’s medical team are ideal “witnesses” for patient commitment to behavior change and “cheerleaders” for small successes.

“You are not responsible, you are response-able” is a great exercise for assisting patients with sorting out blame and fault, as these play a role in commitment to a course of action (Robinson, Gould, & Strosahl, 2010). The teacher can introduce this exercise by writing the word “responsible” and then “response-able” on a white board. Then, the teacher can explain that that today we often associate “being responsible” with blame and fault; however, the original meaning of “responsible” was “being alive.” Then, the teacher can facilitate a discussion about the importance of seeing available options at any given moment—to be response-able to see and choose vital, life-supporting options at any given moment.

Topic 4: Flexible Perspectives

Many patients with PP may have self-stories that keep them stuck in pursuing less meaningful lives that they would choose if they know they could choose. It is important to help patients learn to see stories about themselves as just “stories,” with some working to produce important outcomes much better than others. The teacher can introduce the idea of flexible perspectives by talking about all the different views possible on something that happens in a room with 10 people in it, as all will have different perspectives. Whatever our perspectives or stories about others or ourselves, they make perfect
sense in regard to our life context and experience. The question is, “Does the story we are telling at the moment produce the outcomes we want?” If not, then we may want to switch to a different perspective.

Several exercises work well in a class setting. The first is “Tell three stories about yourself.” Participants will work with a partner in this exercise, taking turns in telling three stories. The stories are in three different contexts: job interview, meeting someone interesting at a party, and defending yourself against a false allegation. The prompt for all three stories is, “Tell me about yourself.” Participants often resonant to this exercising, as most find it easy to tell stories. In wrapping up the exercise, the teacher will pose the question, “Which story is the true story about you?” This will facilitate a discussion about the ease with which we as humans make up stories about ourselves and about others. Our stories depend on the context we are in. The exercise debrief needs to encourage participants to look at their stories from a perspective of curiosity and interest rather than complete belief in them and to look at how they work in our life in terms of producing vitality. “Chapters in a book” is another helpful exercise for this class topic (Robinson, Gould, & Strosahl, 2010). It simply asks the participant to “name the chapter” in their book of life that they are living or writing at the moment and then to reflect on other chapters they’ve written and to notice how they feel as they flip through the book. The teacher can continue the exercise by asking participants to name the next chapter and perhaps identify other characters in the book. This metaphor can set the stage for development of new stories about self and others that support more vital outcomes.

**Topic 5: The Defuse Difference**

While somewhat challenging to convey, this is a very important topic for the class. The teacher may introduce this topic by talking about “favorite sayings” about life and asking participants to offer up their favorite saying while the teacher records them on a whiteboard. Then, the participants can talk about how well these sayings work in different contexts. It will be easy to draw out a few critical ideas, including the invisible operation of these sayings in our lives and our direct experience with using these to guide our behavior. Sometimes they work as guides and sometimes they do not, and when they do not it is very difficult for us to notice. Learning to defuse is the key to tuning in to see what our direct experience is telling us.

From this discussion, the teacher can move into an exercise inviting participants to notice what “sticky” thoughts come up in response to prompts, such as “What is wrong with you?” or “Be careful or other people will notice X.” These types of questions come up fairly often during the course of day-to-day life and it pays to know what our responses to them are because these are the sorts of things we can become quickly fused with, and when fused it is difficult to notice what is happening. The teacher can pass out a pad of paper with sticky tabs on the back and encourage participants to write down two or three of their sticky thoughts. After participants do so, the teacher can ask them to stick the note to their shirt or forehead and then to walk about the room talking with others. This exercise changes their relationship to their sticky thoughts and empowers them to be less fused to those thoughts in day-to-day life and freer to show up to their direct experiences in life. This exercise is called the “sticky thoughts” exercise.

A second exercise for this topic is singing the happy birthday song using only the word pain. The teacher will need to explain how words function to impact our internal and external responding, and that we can change the function of a word by changing the context we experience it in. Pain is usually experienced in a context of distress, and moving it to a context of play or joy changes its function in other contexts. In group medical visits with monthly meetings, it is a good idea to celebrate birthdays at the beginning of the class by singing the happy birthday song in its usual way and then by using only the word pain to sing it.
Topic 6: Accept Difficulties

A critical skill in flexibly pursuing meaning directions in life is the ability to accept difficult and previously avoided experiences while engaging in chosen behaviors that evoke those experiences. The physical metaphor of “driving a bus” is a great exercise for this. Basically, the person driving the bus states a valued direction and a behavior experiment consistent with that direction and then drives the bus in several ways. The driver also identifies two or three people to play the role of passengers on the bus, and the passengers act out what represent difficult experiences that are likely to show up when the driver starts driving toward a value-consistent goal. In the first driving experiment, the driver tries to control the passengers, get them to sit down and shut up. This usually results in even more noise from the passengers. The driver then attempts to accept the passengers and focus on the direction. Accepting the passengers often involves quietly acknowledging them (“Hello, Anxiety, I know you are here”; “Mister Self-Doubt, I see you are on board”) while directing more of the attention and effort on driving the bus. A didactic on practicing compassion is also helpful with the topic accepting difficult experiences, as core beliefs about the self may be painful (e.g., “I’m not good enough”; “I’m broken”) and interfere with chosen actions. In the exercise, participants experience the impact of using self-compassion statements discussed earlier in this chapter. With ongoing practice of self-compassion, painful thoughts and images can be more readily accepted and the person more able to take meaningful action in their presence.

Foundations for Flourishing Class

Positive psychology and ACT overlap in significant ways, and some clinicians may want to offer a second class. To encourage this we include information about the seven foundations of flourishing, examples of each, and a list of possible interventions. All of these are included in Table 24.7. The authors highly recommend Mindfulness, Acceptance and Positive Psychology: The Seven Foundations of Well-Being (Kashdan & Ciarrochi, 2013) to readers with an interest in developing this class at their clinic.

NEEDS Drop-In Class

While there are a number of lifestyle factors that may cause or exacerbate PP and result in decreased quality of life among chronic pain patients, less is known about the best mechanisms for improving lifestyles in these patients. More studies are needed, particularly effectiveness studies. It is likely that a combination of ACT, CBT, CBT Insomnia, and motivational interviewing will be effective in a class for patients with PP who are interested in looking a lifestyle changes. The NEEDS checklist introduced earlier can be used to identify patients for participation in a NEEDS Lifestyle Class. The recommendation is that the class be offered as an open-access group and that it meet on a weekly basis. Each class can focus on one of the areas—nutrition, exercise, enjoyment, don’t drink or smoke, and sleep. Topics can be listed, along with dates, in exam rooms or waiting areas, so that patients have the opportunity to self-identify and self-refer. Patients should be invited to attend any single class or all classes, repeating as many times as is useful for the patient.

The clinician teaching this class may use a variety of psychoeducational handouts to support behavior change (see http://behavioralconsultationandprimarycare.com for examples). Regarding nutrition, a mindful approach that encourages fresh foods, four light meals per day, and a focus on foods that reduce inflammation in the body is recommended. Because exercise has been shown to be helpful to improving a variety of disorders related to PP (Busch, Schachter, Peloso, & Bombardier, 2002; Fransen et al., 2015), establishing baselines in exercise routine and striving to complete short walks and gentle stretches routinely throughout the day is suggested. Clinicians may be able to start a weekly 1-hour walking group
<table>
<thead>
<tr>
<th><strong>Foundation</strong></th>
<th><strong>Example(s)</strong></th>
<th><strong>Interventions</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Helpful beliefs about self, others, and the world</td>
<td>Do you believe you can overcome problems and achieve goals?</td>
<td>Taking the power away from unhelpful thoughts (experience thoughts as passing events)</td>
</tr>
<tr>
<td></td>
<td>Do you view problems as a challenge or threat?</td>
<td>Adding in helpful thoughts</td>
</tr>
<tr>
<td></td>
<td>Do you believe you are a good person?</td>
<td></td>
</tr>
<tr>
<td>II. Mindfulness</td>
<td>Are you aware of your feelings, actions, surroundings, and thoughts?</td>
<td>Learning to be mindful (open and curious about experiences)</td>
</tr>
<tr>
<td></td>
<td>Can you describe the feelings you are having at a given point in time?</td>
<td>Learning to recognize emotions and describe them</td>
</tr>
<tr>
<td>III. Perspective taking</td>
<td>Can you take the perspective of others?</td>
<td>Learn to shift perspectives: participant, observer, witness</td>
</tr>
<tr>
<td></td>
<td>Can you take perspective on yourself?</td>
<td></td>
</tr>
<tr>
<td>IV. Values</td>
<td>What do you care about (values, personal strivings)?</td>
<td>Identify what matters most and talk about these things often</td>
</tr>
<tr>
<td></td>
<td>Do other people’s desires for you dominate your own?</td>
<td></td>
</tr>
<tr>
<td>V. Take action, even when difficult</td>
<td>In order to live according to what you care about, are you willing to have painful feelings and thoughts (e.g., nervousness, doubting yourself)?</td>
<td>Develop more courage—act according to your values even when it is difficult</td>
</tr>
<tr>
<td>VI. Control behavior</td>
<td>Are you able to control what you say and do so that you can pursue your dreams and goals?</td>
<td>Connect values with behaviors</td>
</tr>
<tr>
<td></td>
<td>Do you persist and bounce back from failure?</td>
<td>Set goals, figure out how to work with barriers</td>
</tr>
<tr>
<td></td>
<td>Are you able to change feelings in a way that helps you adjust?</td>
<td>Know how to change emotions (e.g., listen to music, go for a walk)</td>
</tr>
<tr>
<td>VII. Learning skills</td>
<td>How well do you solve problems and figure things out?</td>
<td>Improve memory or concentration</td>
</tr>
<tr>
<td></td>
<td>How well do you change your attention and screen out things that don’t matter?</td>
<td>Learn new things</td>
</tr>
</tbody>
</table>

meeting for interested patients to provide ongoing support for exercise. In the class focused on enjoyable activities, the teacher can offer lists of potentially enjoyable activities and pleasurable activities monitoring forms (both available on the Internet). Patients may be provided with information about a variety of approaches at the alcohol and tobacco class, including psychoeducational handouts, community resources (e.g., Alcoholics Anonymous, Smart Recovery, Free and Clear classes), and evidence concerning medication supplements. The CBT Insomnia (CBT-I) curriculum for the NEEDS class on sleep is recommended. Hopefully, many clinics will already have a class series for CBT-I and the NEEDS class will serve as an introduction to it. If not, then clinicians will want to try to offer the CBT-I interventions on an individual basis, as sleep is a critical factor in improving outcomes for patients with PP. For additional ideas on curriculum development, see Robinson, Bauman, and Beachy (2016).

Healthy Teams in Primary Care

Working in teams offers many advantages to caring for patients with PP. A team approach increases the likelihood of patients being able to access the expertise they need at the time they need it. Additionally, a team approach supports better continuity of care. Even with these advantages, day-to-day delivery of team care can be challenging. When the challenges are not addressed, patient care may experience care as less helpful. In this section, a few strategies for the team to use to maximize team health are explored.

Primary care is a fast-paced practice, and the team needs lots of compassion—for team members and for patients. Taken literally, compassion means “to suffer together.” Emotion researchers define this as the feeling that arises when one is confronted with another’s suffering and feels motivated to relieve that suffering. Compassion differs from empathy and altruism in that it involves the desire to help. Brain scientists are beginning to map the biological basis of compassion, and the deep evolutionary purpose of compassion is emerging. This line of research suggests that heart rate slows when one feels compassion and the regions of the brain linked to empathy, caregiving, and feelings of pleasure are activated. This increases the likelihood of our wanting to take action to help. Teams should develop practices that support daily experience of compassion. Here are some ideas.

Huddle Compassion

In this exercise, the team includes compassion in the huddle that begins the practice day. In reviewing the list of appointments, one or more team members can identify a desire to help a particular patient and then the team holds that desire in a moment of silence with each member trying to activate a sense of caring and possibility for making a difference for that patient. Patients selected do not need to be particularly challenging, but they may be. There does not need to be a great deal of detail regarding the action plan for helping identified patients, other than the typical huddle plans, because the focus is on simple intention and openness to helping in whatever way is possible once the encounter with the patient begins.

Teammate Compassion

This exercise involves setting a goal of feeling compassion for at least one teammate at least one time during the clinic day. This goal can be set silently at the beginning for the practice day in the huddle. Toward the end of the practice day, each team member may tell the team member(s) they felt compassion for and when that experience occurred. If the team member took action on the feeling of compassion, they may communicate that as well. Another important agreement is for team members to express gratitude to a team member that experiences compassion and attempts to act on it. Over time, each team member needs to make an effort to feel compassion for each of the members of the team and to communicate this.
Self-Compassion

It is perhaps easier for most to feel compassion for others than to feel compassion for themselves. Therefore, self-compassion should be cultivated by direct practice. We highly recommend that each team member spend a few minutes warming up their brain for compassionate practice by taking a few slow, deep breaths and sending these intentions to themselves: “May you be physically calm. May you smile often today. May you slow down. May you know gratitude.” There may be other intentions that each team member wants to practice, but these four are great beginners. It’s a good idea to post them in the team room, so that they can serve as reminders throughout the busy day.

SUMMARY

The purpose of this chapter is to help readers develop new services for primary care patients with persistent pain. Opportunities for improving care were suggested and evidence for taking specific approaches was presented. The Healthy Brain and Body tool kit is the most substantial part of this chapter. This section provides examples of assessments and interventions, many based on Acceptance and Commitment Therapy. Traditional CBT and motivational interviewing were also recommended. To support systematic delivery of services, the tool kit provides tools to use in individual and group visits, including the Bull’s Eye Care Plan, the Healthy Mind and Body Class, and the NEEDS Lifestyle Class. This chapter concludes with a few suggestions about clinician and team resilience, as this is of fundamental importance to development of the compassion needed to start and sustain new ways of working with pain. While most readers will not be able to implement all of the programmatic strategies suggested, the hope is that reading about them inspires readers to start something new and that that something inspires hope and optimism about the problem of persistent pain.

REFERENCES


With over eight million enrollees and over 1,400 hospital systems and community-based outpatient clinics located in all 50 states, Puerto Rico, the Philippines, Guam, and American Samoa, the Veterans Health Administration (VHA) is the largest integrated healthcare system in the United States (Kearney, Post, Pomerantz, & Zeiss, 2014a). The VHA functions much like an Accountable Care Organization (ACO), delivering care to a defined population. VHA care for most veterans is funded by a congressional appropriation and a global budget, rather than fee-for-service. VHA facility funding is primarily based on the number of veterans served, multiplied by a predetermined per capita rate, rather than by number and intensity of services provided. VHA is not a part of the military health system, which provides care to active duty members of the armed forces and their dependents. Rather, VHA is an option for care for individuals following military service, but with a few exceptions, it does not provide care for their dependents.

Traditionally, the veteran population has an illness burden greater than the general population, complicated by a high prevalence of mental illness (Agha, Lofgren, VanRuiswyk, & Layde, 2000; Trivedi et al., 2015). Over the past two decades, the wars in Afghanistan and Iraq have led to a significant increase in the number of VHA enrollees, and the prevalence of mental illness in the population served by VHA has increased as well. This change is demonstrated by a 63% increase in the number of veterans receiving mental health services from 2005 to 2013, which is three times faster than the overall growth rate of veterans using VHA services during the same time (VHA, 2014). Of those who used VHA services in 2013, 26% utilized mental health treatment, which is nearly double that of the general US population (VHA, 2014).

VHA mental health services are a critical component of the US public health structure, traditionally serving as a safety net care system for eligible veterans. In recent years, VHA has increasingly attracted other veterans who prefer VHA health care to that offered in the private sector or who cannot otherwise access the extensive services offered in special areas of need for veterans. These include programs for posttraumatic stress disorder (PTSD) treatment, amputee services, spinal cord injury services, and polytrauma services, which include treatment of multiple injuries to numerous body parts and organs as well as psychological comorbidities from blast-related injuries. Given the high percentage of individuals with mental illness in its population and the high overall burden of illness, VA has invested significant resources in mental health care, with $6.2 billion dedicated to mental health services in 2013, constituting 11.2% of VHA’s total healthcare budget (VHA, 2014). This mental health budget supports implementation of a range of services, from routine general outpatient mental health services to inpatient and residential care, including specialized treatment programs for PTSD and substance use disorders (SUD).
In the United States, most individuals with mental disorders present first in the primary care setting and are commonly not identified unless systematic screening is in place (Kessler, 2009; Kessler et al., 2005; Regier et al., 1993). Those who are identified either are most commonly not treated or receive inadequate treatment (Kessler et al., 2005; Young, Klap, Sherbourne, & Wells, 2001). The majority of those referred to mental health specialists either do not show up for treatment or do not engage after the first appointment (Wray, Szymanski, Kearney, & McCarthy, 2012; Zanjani, Miller, Turiano, Ross, & Oslin, 2008). To address this problem, many organizations have piloted approaches to improve mental health services in primary care including appropriate screening methods, care management, and colocated collaborative care approaches. VHA has adopted all of these approaches to improve engagement in mental health care through integrating primary care and mental health services.

HISTORY OF VHA MODEL

Early identification of mental health concerns through standardized, routine screening was a foundational first step for integrating mental health and primary care. Screening for depression, alcohol use, and PTSD became the norm within the VHA in the 1990s, and individual clinics developed appropriate clinical pathways into treatment for individuals with positive screens (Department of Veterans Affairs and Department of Defense [VA/DoD], 2009, 2010, 2015). Soon afterward, randomized controlled trials consistently demonstrated that most patients with uncomplicated depression could be effectively managed entirely within primary care (Dietrich et al., 2004; Katon et al., 2010; Unützer et al., 2002). Although there were several variations of this approach, termed “collaborative care,” all of the models tested with randomized controlled trials used nurses or other professionals as “care managers,” partnered with psychiatrists, to provide telephone follow-up for patients being treated for depression by primary care providers (Oxman, Dietrich & Schulberg, 2003). The care management services used in the collaborative care studies used fewer resources than referral to specialty care usually entailed and consistently demonstrated improved outcomes over care as usual (Unützer et al., 2002). VA subsequently developed and successfully tested two of its own approaches to care management: the Behavioral Health Laboratory (BHL) and Translating Interventions for Depression into Effective Solutions (TIDES; Oslin et al., 2003, 2006; Rubenstein et al., 2006).

At the same time that these care management approaches were being tested, many VHA clinics expanded their efforts to improve patient engagement in mental health care by colocating providers in primary care clinics. At the time, VHA had long provided psychological interventions for conditions common in the primary care population, such as pain and insomnia. These treatments, however, typically required that patients accept referral to mental health or other specialty clinics, usually located outside of primary care. As in other systems of care, no-show and nonengagement rates for such treatment were typically 50% or more (Kessler et al., 2005). As a result, many clinic appointment slots were wasted and the clinic capacity was limited for new patients who did accept mental health referrals (Kearney, Smith, & Pomerantz, 2015).

Despite many local efforts to bring mental health into primary care, simply colocating typical mental health services in primary care, while admirable and desirable, did not meet the needs of the population when clinicians were not trained to work in the primary care setting (Pomerantz, Corson, & Detzer, 2009). Mental health clinicians frequently maintained their traditional mental health service models (50-minute, weekly appointments), schedules quickly filled as they acquired a caseload, and waiting times for new appointments grew rapidly. Clinicians were then not available for open-access handoffs from the rest of the primary care team, also known as “warm handoffs.”
Increasing the delay between referral and initial appointment has been demonstrated to reduce the likelihood that an individual will show up for care (Gallucci, Swartz, & Hackerman, 2005; Van Voorhees, Wang, & Ford, 2003). Newer models of care were needed. Based on the assumption generated by the collaborative care studies of TIDES, BHL, and others, some clinics began to approach common mental conditions with abbreviated, solution-focused brief treatments. One of these initiatives, developed at the White River Junction VA in Vermont, demonstrated that such brief interventions by mental health clinicians embedded in primary care could overcome the problems associated with simple co-location. That program, begun in 2004, soon received national awards and recognition both inside and outside VA (American Psychiatric Association, 2005). It reduced no-show rates for specialty mental health appointments by 75% while reducing waiting times for patients identified with mental illness in primary care clinics from 6 weeks to less than 20 minutes.

Based on the success of TIDES, BHL, and the White River Junction model, VHA funded 97 facilities in 2008 to begin or expand integrated care programs in primary care clinics (Kearney et al., 2014a). This funded initiative, labeled Primary Care-Mental Health Integration (PCMHI), incorporates two key components: care management (CM) and colocated collaborative care (CCC). Care management is based on the collaborative care model of TIDES and BHL, as well as the non-VA trials such as those described earlier (Dietrich et al., 2004; Katon et al., 2010; Unützer et al., 2002). CM provides support for individuals being treated for mental health conditions by primary care providers, as well as for individuals treated by the embedded mental health clinicians of the CCC component of PCMHI. Colocated collaborative care embeds clinicians, typically psychologists, clinical social workers, psychiatrists, and advanced practice nurses within the primary care clinic setting with a focus on primary treatment of mild to moderate symptoms of depression, anxiety, PTSD, and alcohol use. CCC providers also address other conditions such as chronic pain, insomnia, hypertension, diabetes, and stress management where behavioral health interventions have a strong evidence base for assisting in improvement of symptoms. CCC providers use brief assessments (typically relying heavily on self-report measures in addition to information gathered in the clinical interview) and treatments (20–30 minute appointments spread over one to four sessions). CCC providers are members of the primary care teams, providing ongoing, consultative support to assure that the population needs are addressed. Rather than assuming individual caseloads, the primary care patient panel is the “caseload” and care or consultation is delivered to those in need.

A key component of the VA integrated care approach is maintenance of open or, at minimum, same-day access to care, which has been shown to improve access in healthcare clinics (Murray & Berwick, 2003). One goal of the PCMHI program is to assure that a brief functional assessment is provided and care initiated at the time a problem is identified during a primary care visit. Such immediate access avoids the attrition invariably associated with delays of any length between problem identification and mental health assessment and treatment. National program evaluation has demonstrated a significant increase in numbers of veterans with mental illness identified and treated in primary care since inception of the PCMHI initiative (Johnson-Lawrence, Zivin, Szymanski, Pfeiffer, & McCarthy, 2011; Wray et al., 2012). This evaluation has also demonstrated improved engagement in care for those whose needs cannot be met in primary care and require referral for specialized care (Bohnert, Pfeiffer, Szymanski, & McCarthy, 2013; Wray et al., 2012).

Soon after funding of the initial 97 programs in VA, the two components became policy requirements for all facilities and larger community-based outpatient clinics in the Uniform Mental Health Services Handbook (VHA, 2008). The policy directive specified use of TIDES, BHL, or evidence-based care management approaches approved by VA PCMHI leadership and embedded clinicians targeting common, uncomplicated depression, anxiety, and at-risk alcohol use. In 2009, additional funding was provided to facilities that had not previously been funded for PCMHI.
In 2010, VA began another transformation, this one of its primary care clinics to the Patient-Centered Medical Home model (Kearney, Post, Zeiss, Goldstein, & Dundon, 2011). In some sites, the existing PCMHI programs formed the nidus around which this transformation developed. The VA’s Patient Aligned Care Team (PACT) incorporates the structure and function of the Patient-Centered Medical Home. Care is organized around the patient and family. A core teamlet, consisting of a Primary Care Provider, Registered Nurse, Health Technician or Licensed Vocational Nurse, and a Clerk provide care for a panel of 1,200 patients. Several of these teamlets combine to form a larger interprofessional team that also includes social work, clinical dietetics, and PCMHI staff. In addition to the PCMHI staff, VHA hired health psychologists and social workers to serve as health behavior coordinators (HBCs) to work within the PACTs. These coordinators are expected to spend the majority of their time as consultants and trainers, helping primary care staff learn techniques, such as motivational interviewing, that they can apply in their work with patients. Other staff members, Health Promotion/Disease Prevention (HPDP) program managers, complement this work and help the PACTs to develop wellness and prevention programs. Over time, the relationship between PCMHI, HBC, and HPDP has evolved and, in many facilities, all work together as a single program within the PACT, sharing responsibility with the rest of the healthcare team for improving the health of the population. These providers have helped VA clinics accelerate the development of various interventions for weight management, chronic disease, smoking cessation, health coaching, and other initiatives.

In 2011, VA created a new office, the Office of Patient-Centered Care and Cultural Transformation (OPCC&CT) to develop a new approach to care. This new approach, known as personalized, proactive patient directed care (PPPD), represents a shift from disease-based care to promotion of health and overall well-being. OPCC&CT has developed a number of initiatives and processes to assist this change, beginning with a Personal Health Inventory (PHI) that is used to gather information from a number of different domains to capture a picture of the veteran’s overall health. These domains include the social determinants of health, such as housing, social relationships, spirituality, family relationships, leisure activities, diet, and emotional/behavioral wellness. The PHI serves as a prelude to identifying “what matters” to the veteran. These programs are steadily gaining traction in VA through several pilot initiatives, regional and facility-based trainings, clinician tools, and regularly occurring webinars. In this model, care does not necessarily begin with treatment of disease, but may instead begin by addressing housing or other concerns, focusing on whatever the veteran sees as his/her primary need. Care may often include complementary and integrative approaches such as acupuncture, yoga, or meditation. OPCC&CT has developed a “whole health coaching protocol” and has trained hundreds of clinicians and has begun to train peer support specialists as well.

**CURRENT IMPLEMENTATION FOR THE PCMHI PROGRAM**

The National PCMHI Evaluation Office is charged with evaluating the implementation of PCMHI across VHA (Kearney et al., 2014a). On an annual basis, the office has taken a survey of all VA PCMHI programs, expanding its coverage of the survey from all medical centers to every site of care where PCMHI is required, including community-based outpatient clinics (CBOCs) serving more than 5,000 primary care patients. The survey reviews the current status of PCMHI implementation covering fidelity to care management and colocated collaborative care models, staffing levels, clinical treatments provided, types of diagnoses treated in the program, PCMHI policies, and barriers and facilitators of program implementation. Current implementation in VHA has increased steadily
each year. As of February 2016, the program has provided direct clinical services to over 1.3 million veterans with over 5 million encounters (see Figure 25.1). Fiscal Year (FY) 2015 survey data indicate that further implementation of PCMH continues to occur, yet further efforts are still needed to achieve the full vision of the program. While steady progress has been noted, implementation of care management lags behind colocated collaborative care. At sites required to have both CM and CCC, care management is present at only 71.6% compared to CCC, which is present at 90.74% (see Table 25.1).

The work of the PCMH Evaluation Office has led to numerous publications that have enlightened our understanding of integrated primary care implementation (e.g., Pfeiffer et al., 2011; Wray et al., 2012; Zivin et al., 2010). However, to date, these efforts have been limited by the use of administrative and survey data in the absence of uniformly available patient self-report measures of treatment outcomes. The VHA national PCMH program evaluation capabilities would be enhanced by uniform utilization of self-report psychological instruments in PCMH. Such uniformity would allow self-reported patient and population-specific outcomes evaluations across VHA sites. Until there is a VHA policy mandating such practice, the evaluation of patient- and program-level outcomes must rely on the local implementation of these measurement practices. Therefore, the next section of this chapter will focus on the current implementation of measurement-based care (MBC) efforts in PCMH within VHA. Measurement-based care has previously been defined as the process of routinely gathering patient data throughout treatment and using this data to guide clinical care (Scott & Lewis, 2015). Within the context of integrated care, the term MBC will be used to refer to the routine collection of data to inform treatment decisions and systematically monitor clinical outcomes, from initial screening through completion of care.

*Figure 25.1* Cumulative New Patients and Outpatient Encounters for PCMH 2007–2015
As VHA continues to enhance and expand implementation of the PCMHI model within PACT, it has become increasingly important to implement measurement-based care in order to achieve service delivery aspirations (Kearney, Wray, Dollar, & King, 2015). Specifically, the features of measurement-based care concretely support the goals of the medical home, including providing infrastructure and mechanisms to accurately identify, track, and monitor both individuals and panels of patients, fully supporting the provision of team-based patient-centered, evidence-based care across time. In order to implement measurement-based care in PCMHI (Kearney, Wray, Dollar, King, & Vair, 2014b), the unique features associated with the role and functions of psychological assessment in primary care must be taken into account. Effective psychological assessment in the primary care environment requires a substantial departure from traditional psychological assessment practices (Scott & Lewis, 2015).

The type of assessment conducted in primary care must match the rapid service delivery setting; therefore instruments must be brief, easy to administer, quick to score and interpret, and focused on common concerns prevalent in the primary care setting (e.g., depression, substance use, and anxiety). The repeatable measures selected should be psychometrically sound and empirically validated within this setting and appropriate for use within the general population (Kearney et al., 2014b).

In addition to differences in the structure of psychological instruments appropriate for primary care, there are also important distinctions in the purpose and functions of psychological assessment within this setting. The key functions of psychological assessment in PC include (1) population-based screening, (2) triage to the most appropriate treatment setting, (3) individual monitoring across time, (4) tracking of patient outcomes across the entire primary care panel, and (5) communication to the
primary care team on patient and panel progress. (See Kearney et al., 2014b for detailed information on appropriate instruments for veterans in primary care.)

One advantage of MBC in primary care is the identification of patients who would benefit from mental health care through early identification of concerns in a standardized screening process. VHA has implemented a robust and standardized screening process for common mental health conditions in primary care, including depression, PTSD, and alcohol misuse, as well as screening for military sexual trauma and traumatic brain injury (Kimerling et al., 2007; VA/DoD, 2010). In addition to the use of psychological instruments for screening purposes, VHA has mandated that care management include standardized assessment and monitoring protocols (VHA, 2008). These required elements of psychological assessment in VHA for screening and outcome monitoring are described later. Additionally, although not mandated by VHA policy, many CCC providers routinely use psychological assessment as part of routine care delivery in their PCMHI programs.

**Population-Based Screening in VHA Primary Care Settings**

Population-based screening is the routine administration of specific tests or questions to all members of a population in order to identify risk for highly prevalent conditions (e.g., depression) that can be treated, or for which early prevention approaches are available (Kessler, 2009). The US Preventive Services Task Force (USPSTF; Siu & USPSTF, 2016), a panel of national experts in prevention and evidence-based medicine, recommends routine screening for depression in clinics when sufficient resources for appropriate treatment of depression are in place. Within VHA, population-based screening initiatives are tailored to specific populations (e.g., age, gender, disease prevalence) served and are implemented in settings where there are resources to treat any identified conditions. Depression, PTSD, and alcohol-related disorders are highly prevalent within veteran populations and VHA has allocated significant resources for treatment of these conditions both within primary care and specialty mental health care (VHA, 2014).

The VA/DoD (2010) clinical practice guidelines for the management of PTSD highlight the importance of routine screening within primary care for PTSD. Substantial research indicates that PTSD screening increases early identification and treatment referral (VA/DoD, 2010). More specifically, it is recommended that all new patients be initially screened for PTSD and thereafter on an annual basis, and screening that is more frequent may be clinically indicated (e.g., exposure to a traumatic event). VHA has implemented routine screening for PTSD in primary care based on these clinical practice guidelines. The VA/DOD Clinical Practice Guidelines for the Management of Substance Use Disorders (VA/DoD, 2015) includes the specific screening recommendation that all patients in general medical settings (such as primary care) receive screening for alcohol misuse on an annual basis. This recommendation is based on significant evidence that brief alcohol counseling in these settings is effective in reducing unhealthy drinking.

Clinics within VHA (Kearney et al., 2014a) are required to have access to mental health providers typically functioning as part of the primary care team, and are designed to provide primary care-based screening, assessment, brief treatment, and/or referral management in response to positive screens for depression (VA/DoD, 2009), as well as alcohol misuse (VA/DoD, 2015) and PTSD (VA/DoD, 2010). Within VHA, standardized screening instruments include the Patient Health Questionnaire-2 for depression (PHQ-2; Lowe, Kroenke, & Gräfe, 2005), the Primary Care-PTSD Screen for PTSD (PC-PTSD; Prins et al., 2003), and the Alcohol Use Disorder Identification Test Consumption Questions (AUDIT-C; Bradley et al., 2007) for identification of alcohol misuse. Screening for these conditions
occurs at least on an annual basis and may occur more frequently for individuals at high risk for specific conditions. Based on the unique risk factors associated with the veteran population, VHA has also implemented systematic one-time screening initiatives for select conditions. For example, early in the process of initially enrolling for VHA primary care services, veterans receive a one-time screening for traumatic brain injury (VA, 2010) as well as for military sexual trauma (Kimerling et al., 2007). Screening is most often administered by a PACT nurse or health technician and reviewed by the PC provider, who discusses positive screens with the patient and develops a plan for further assessment and treatment. This may include immediate introduction to the PCMHI provider for the additional assessment and follow-up. Figure 25.2 provides a flow chart depicting a typical clinical flow process in VHA for implementation of required population-based screening and measurement-based care with PACT.

Use of MBC in PCMHI Care Management in VHA

As previously described, mental health care management is a required component of PCMHI in VHA (VHA, 2008). Key elements of the model include (1) monitoring medication side effects, outcomes, and adherence to treatment; (2) decision support for primary care teams; (3) patient activation and education; and (4) referral management assistance to specialty mental health when needed (VHA, 2008). Care management involves case supervision by licensed independent providers and ongoing treatment of patients includes active collaboration with the primary care team. The most consistent implementation of MBC in VHA PCMHI to date is through care management program components. Therefore, the VHA roll-out of BHL and TIDES is the focus of this section.
A strength of the PCMHI CM approach is the standardized and routine application of validated self-report measures in veteran care. Through baseline and repeated measurement, care management uses evidence-based algorithms to assist care managers and health technicians in identifying appropriate patients for programs that provide medication monitoring, behavioral activation, cognitive behavioral, and motivational interviewing interventions (Oslin et al., 2003; 2006; Rubenstein et al., 2010). Further, both BHL and TIDES care management programs provide an infrastructure for creation of a patient registry and tracking system that allows not only for individual, routine monitoring of patient progress, but also for the systematic extraction of the program’s data to evaluate overall outcomes related to various mental health conditions. Currently within the VHA, 269 sites have implemented care management within their PCMHI programs. Delineated in the next sections are the descriptions of the instruments and protocols utilized in both the BHL and TIDES programs, along with descriptions of some of the unique features of each approach.

### Behavioral Health Lab

#### Staffing and Software Support

The Behavioral Health Lab care management program typically includes a health technician to collect data during the initial screening assessment, care managers (typically social workers, nurses, or in some cases psychologists), and a supervising prescriber. In some programs without health or psychology technicians, the care manager completes the initial screening assessments. A hallmark of this program is the BHL software, which provides a uniform portal for all data entry related to initial patient evaluation and follow-up monitoring, and supports review of individual and population outcomes across multiple sites of care. The software provides prompts to those conducting initial and follow-up encounters to ensure all required elements are covered. Decision-support algorithms are contained within the software and these algorithms can be modified based on local conditions. The software was developed by Capital Solution Design in partnership with the University of Pennsylvania and the Veterans Integrated Service Network (VISN) 4 Mental Illness Research, Education, and Clinical Center (MIRECC) at the Corporal Michael J. Crescenz VA Medical Center.

#### Initial Assessment

Veterans initially referred to BHL care management typically receive a phone call to complete a baseline screening assessment covering multiple areas (see Table 25.2); however, this initial screening is completed in person if care managers are colocated in the clinic. The purpose of this initial assessment

---

A strength of the PCMHI CM approach is the standardized and routine application of validated self-report measures in veteran care. Through baseline and repeated measurement, care management uses evidence-based algorithms to assist care managers and health technicians in identifying appropriate patients for programs that provide medication monitoring, behavioral activation, cognitive behavioral, and motivational interviewing interventions (Oslin et al., 2003; 2006; Rubenstein et al., 2010). Further, both BHL and TIDES care management programs provide an infrastructure for creation of a patient registry and tracking system that allows not only for individual, routine monitoring of patient progress, but also for the systematic extraction of the program’s data to evaluate overall outcomes related to various mental health conditions. Currently within the VHA, 269 sites have implemented care management within their PCMHI programs. Delineated in the next sections are the descriptions of the instruments and protocols utilized in both the BHL and TIDES programs, along with descriptions of some of the unique features of each approach.

### Behavioral Health Lab

#### Staffing and Software Support

The Behavioral Health Lab care management program typically includes a health technician to collect data during the initial screening assessment, care managers (typically social workers, nurses, or in some cases psychologists), and a supervising prescriber. In some programs without health or psychology technicians, the care manager completes the initial screening assessments. A hallmark of this program is the BHL software, which provides a uniform portal for all data entry related to initial patient evaluation and follow-up monitoring, and supports review of individual and population outcomes across multiple sites of care. The software provides prompts to those conducting initial and follow-up encounters to ensure all required elements are covered. Decision-support algorithms are contained within the software and these algorithms can be modified based on local conditions. The software was developed by Capital Solution Design in partnership with the University of Pennsylvania and the Veterans Integrated Service Network (VISN) 4 Mental Illness Research, Education, and Clinical Center (MIRECC) at the Corporal Michael J. Crescenz VA Medical Center.

#### Initial Assessment

Veterans initially referred to BHL care management typically receive a phone call to complete a baseline screening assessment covering multiple areas (see Table 25.2); however, this initial screening is completed in person if care managers are colocated in the clinic. The purpose of this initial assessment

---

### Table 25.2 BHL and TIDES Instruments

<table>
<thead>
<tr>
<th>Program</th>
<th>Area</th>
<th>Instrument</th>
<th>Citation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BHL</td>
<td>Depression and suicidal ideation</td>
<td>Patient Health Questionnaire-9 (PHQ-9)</td>
<td>Kroenke et al. (2001)</td>
<td>DSM criteria-based depression screening</td>
</tr>
<tr>
<td>TIDES</td>
<td>Depression and suicidal ideation</td>
<td>Paykel questionnaire</td>
<td>Paykel et al. (1974)</td>
<td>Review of prior suicide attempts and ideation</td>
</tr>
<tr>
<td>BHL</td>
<td>PTSD</td>
<td>PTSD Checklist (PCL)</td>
<td>Bliese et al. (2008); Weathers et al. (1993)</td>
<td>Assesses for PTSD symptoms based on DSM criteria</td>
</tr>
</tbody>
</table>

(Continued)
is to provide information on multiple data points that can be used by licensed independent providers (LIPs) to assist in treatment planning with the veteran. In addition to collecting background information on demographics and prior treatment, multiple instruments cover the following areas: depression, posttraumatic stress disorder (PTSD), suicidality, dementia/memory concerns, mania, psychosis, panic disorder, generalized anxiety disorder, alcohol and drug use, and chronic pain.

After the screening assessment is completed, a summary document is submitted to a LIP for review, and treatment planning can then occur, identifying the best resources for treatment based on these results and findings from initial discussions with the patient. For medication-related issues, a supervising prescriber, such as a psychiatrist, may review the case for relevant treatment planning and decision-making to support the patient and primary care provider involved in care. For cases deemed appropriate for intervention in PCMHI (e.g., mild to moderate depression/anxiety and alcohol misuse), treatment may be initiated with further intervention provided through care management, through the colocated collaborative care provider, or through continued treatment by the PCP.
Ongoing Monitoring and Panel Management

In addition to initial assessment of all new patients enrolled into BHL care management, patients who continue on to participate in one of the modules for treatment (e.g., depression, alcohol misuse, anxiety, referral management, or watchful waiting) will continue to receive routine assessment on progress related to the identified targeted condition. The schedule of follow-up calls and assessments is based on the condition being monitored and can be adapted to meet local requirements. During each follow-up call, the CM will administer the appropriate repeated measure and check for adherence to the treatment plan, patient concerns, and side effects if medication is involved. Primary care teams will receive updates on patient status and are involved if the treatment plan needs to be adjusted. Routine CM panel review meetings with the supervising psychiatrist allow for discussion of patient progress and identification of patients whose lack of improvement requires attention. All treatment plan changes are through collaboration with the patient, the CM, and the appropriate providers (PCP, PCMHI LIP, etc.).

The BHL software provides prompts to the CM to denote when follow-up contacts and assessments are needed in concordance with evidence-based models of intervention. Each module within the software has specific guidelines, decision trees, and follow-up schedules to guide treatment. The software can also be utilized to provide summary reports on treatment improvements (or lack thereof) on both individual patients and the population of patients enrolled in particular modules of care. A visual summary of progress, including graphs of outcomes on particular instruments, can be given to patients to provide feedback on their treatment. Patient registries within the software allow for a mechanism for data review by supervising clinicians on the population of enrolled patients. For example, symptom reduction, referral patterns, workload data, completion rate, and engagement rates assist leaders in evaluating program outcomes and program quality improvement processes. The patient registry also helps providers to track appointments and customize contents to flag cases of concern needing additional follow-up action. Example printouts by the BHL software can be found at http://www.mirecc.va.gov/visn4/BHL/docs/BHLSoftwareBrochure2013.pdf.

Translating Initiatives in Depression Into Effective Solutions (TIDES)

Staffing and Software Support

TIDES is primarily staffed by Registered Nurses (RNs) serving as Depression Care Managers (DCM) for implementation of both the initial screening and follow-up care management modules. Clinical supervision is provided by a prescribing provider, typically a psychiatrist. A DCM can typically follow approximately 150 patients over 6 months full time with a total panel of approximately 300 per year (Liu et al., 2007). Overarching panel suggestions include a recommendation of hiring one DCM per 7,500 to 10,000 primary care patients as a whole.

Unlike BHL, TIDES does not have a separate software system for utilization. Within TIDES, most DCMs track progress on self-reported instruments through secured Excel spreadsheets, which can be utilized to track individual and program-level progress and outcomes. However, in recent years, some facilities have implemented the BHL software to complete tracking while continuing to utilize TIDES materials for care management.
Initial Assessment
The initial TIDES assessment is conducted by the DCM typically by phone and takes approximately 40 minutes to complete. The DCM completes the PHQ-9 followed by standardized questions related to suicidal ideation, homicidal ideation, depression history/treatment, current antidepressant use, sleep and appetite symptoms, comorbid medical conditions, alcohol and drug use, anxiety and panic symptoms, PTSD symptoms, bipolar symptoms, functioning/support, and current self-care activities. If alcohol use is present, the AUDIT-C is also conducted. In addition to information gathering, the initial assessment provides an opportunity for the DCM to build rapport with the veteran and provides opportunities for motivational interviewing to assist with veteran engagement and activation. The case summary is then discussed with the supervising provider and then decisions for next steps are made for implementation.

Ongoing Monitoring and Panel Management
If care management is indicated, DCMs continue implementation of patient outcome measures, at each encounter, to assess individual patient progress over time, as well as to monitor overall program outcomes across a panel of patients (Chaney et al., 2011; Liu et al., 2003). Follow-up monitoring for those enrolled in DCM typically consists of six contacts over a 6-month period of time. Initial monitoring includes reviews of medication compliance and side effect monitoring after the initiation of an antidepressant medication. Feedback is given to the prescribing provider on progress during this time period and treatment goals and plans are developed with the veteran in collaboration with the PACT team. Interventions during the 6 months include patient education, addressing self-care processes, and increasing behavioral activation utilizing standardized materials on the TIDES share point, which include patient mailing materials. Treatment plans are updated as a result of patient progress monitored in routine standardized reassessment. Should patients require treatment that is more intensive than can be delivered in primary care, DCMs provide referral management services while the transfer to specialty care occurs.

Use of MBC in Colocated Collaborative Care (CCC) in VHA

Staffing and Software Support
A diverse number and type of professionals serve as CCC providers within VHA, with the majority of CCC providers being either clinical social workers or psychologists, followed by psychiatrists and advance practice nurses (APN; Wray et al., 2012). However, multiple other licensed mental health professionals can function in these roles, including marriage and family therapists (LMFT) and licensed professional mental health counselors (LPMHC). Although VA has recommended a staffing model of 0.5 CCC providers and 0.17 care managers per 1,000 primary care patients, few sites are yet staffed to that level.

As a distinct function of PCMHI, CCC differs from both care management programs (TIDES & BHL) and no specialized software program has been developed specifically for CCC. However, CCC clinicians have access to and routinely utilize assessment and tracking tools available to all VHA mental health clinicians via the electronic medical record, including Mental Health Assistant (a package of psychological assessments linked to the VHA electronic medical record). As described below some programs use software developed for care management (BHL software and secure Excel tracking documents) to further support CCC and incorporate routine assessment.

Initial Assessment
As described in detail elsewhere (Kearney et al., 2014a), the initial assessment within CCC is typically a functional assessment that is focused on the referral question from the primary care provider and the primary concern of the patient. This functional assessment occurs after the screening initially
conducted by other PACT team members. Distinct from initial assessments in other mental health settings, CCC initial assessments are brief and emphasize functional domains (e.g., occupational and social functioning) rather than symptomatology and diagnosis. Although not mandated and therefore not universal, many CCC providers supplement the functional assessment through use of brief standardized measures, such as the Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001), the PTSD Checklist-Military Version (PCL-M; Weathers, Huska, & Keane, 1991) and the AUDIT-C (Bradley et al., 2007).

Ongoing Monitoring and Panel Management
As there are no monitoring requirements specified by policy for CCC practice, there is great variation across VHA in terms of ongoing monitoring and panel management within CCC. Conceptually, CCC is a platform for care delivery rather than a specific intervention (e.g., as opposed to care management for depression). As a result, CCC practice addresses diverse presenting concerns with a wide range of interventions, complicating routine monitoring across a panel of patients. Typically, patients are monitored through in-person follow-up with the CCC provider in the primary care clinic. The CCC provider assesses progress on the PACT-based treatment plan through a combination of verbal review of progress on goals, clinical judgment, and use of standardized measures. Progress is then reviewed at PACT team meetings and treatment decisions are discussed with the primary care provider. The extent to which panel management occurs varies based on implementation of tracking processes within CCC, such as use of BHL software or other secure tracking databases.

Examples of MBC in CCC
Several VHA CCC programs have integrated routine outcome monitoring as part of their standardized process despite the lack of policy requirements to dictate this process. Although many clinicians appreciate this flexibility, the lack of specific guidance and requirements has created heterogeneity across the healthcare system, usually based on individual provider clinical preferences or local policy decisions. Thus, the degree to which CCC providers have implemented MBC varies greatly across the system, with some locations providing little to no systematic assessments, inconsistent use in other locations, and some locations fully implementing the practices of measurement-based care. Taken together the full extent to which systematic measurement-based care has been implemented within VHA CCC differs substantially by clinic location.

The White River Junction model is one such program that is well known for its incorporation of routine (i.e., every patient at every visit completes standardized measures) use of measurement-based care within PCMHI (Pomerantz et al., 2010). After the warm handoff from the primary care team, the veteran is provided with a portable electronic touchpad, which administers four structured assessment instruments, the PHQ-9, the PCL-M, the Generalized Anxiety Disorder-7 (GAD-7; Spitzer, Kroenke, Williams, & Lowe, 2006), and Medical Outcomes Study Short Form-12 (SF-12; Jenkinson et al., 1997). Typically, these are completed in the waiting room prior to the appointment and results are reviewed and discussed during the appointment with the provider. At follow-up visits, assessment instruments that had indicated clinically significant results at the initial assessment are re-administered. The results of the assessment instruments are used to track individual, cohort, and system outcomes. In addition, the scores help to focus, and thus streamline, the diagnostic assessment itself and the discussion of ongoing treatment with the patient (Pomerantz et al., 2010). Other programs across the VHA have implemented a similar process where the veteran completes routine standardized assessments using pen and paper while in the waiting room prior to the appointment or at the start of the appointment while the provider completes a brief chart review. These assessment results are then entered into an electronic database by either the provider or a technician.
Some locations have incorporated the use of all or select portions of the care management assessment protocols within the CCC context. For example, the Hampton VA provides open access to either a CCC provider or a Health/Psychology Technician colocated in the PCMHI clinic. Either team member may receive the warm handoff, depending on clinic flow and availability. If the CCC provider receives the warm handoff, he or she initially completes an individualized functional assessment (Hunter et al., 2009). Once the functional assessment is completed, the CCC provider offers the patient the option to meet that day with a Health Technician to complete an initial, standardized baseline assessment (see earlier BHL description) in person in the same location. If the CCC provider is not available immediately to take a warm handoff, the technician will receive the warm handoff and complete an initial structured baseline assessment in person using the BHL software. After the baseline assessment is completed, the CCC provider joins the patient to review the BHL results, complete a brief functional assessment, and collaboratively develop the treatment plan. Throughout the course of PCMHI care (regardless of whether the patient is being followed by a CCC provider or a care manager), the Health Technicians contact all patients who are receiving PCMHI services on a routine basis to conduct routine standardized assessments. The data are entered into a database that includes the initial assessment data, allowing for continued monitoring of symptoms and treatment response across time treatment. Modifications to the treatment regimen (including medication adjustments, brief interventions, or referral to more intensive psychological services) are informed by the regular review of data across time.

**BARRIERS AND FACILITATORS TO IMPLEMENTATION OF MBC IN VHA INTEGRATED CARE**

Full implementation of MBC in VHA is challenging, but the system has already made some significant inroads, particularly through its roll-out of care management. However, continued work is needed for MBC to become routine practice across all PCMHI programs. VHA experience with implementation of numerous clinical programs allows us to predict barriers, as well as facilitators, for the implementation of PCMHI MBC. These are explored in this next section.

**Barriers to Implementation of MBC in PCMHI**

A large success in VHA is that screening with brief measures is nearly uniformly implemented across the system, and veterans and primary care teams view this as routine practice. For example, rates of screening in FY15, Quarter 4 were at 96% for depression, 98.6% for PTSD, and 96.3% for alcohol use. Given that this component of MBC is so widely implemented, why are other components not yet in use? The Consolidated Framework for Implementation Research (CFIR) has been developed in an effort to organize the growing body of implementation science studies across a wide variety of interventions (Damschroder et al., 2009) and can be used to understand the challenges and facilitators that can be expected as VHA implements MBC. CFIR includes five major domains: intervention characteristics, outer setting, inner setting, characteristics of individuals, and process of implementation (see Table 25.3 for definitions of these domains).

**Intervention Characteristics**

For the context of this discussion, the term “intervention” refers to all aspects of MBC. Intervention characteristics include perceptions of the intervention’s source, the evidence for its use, its relative
<table>
<thead>
<tr>
<th>CFIR Domain</th>
<th>CFIR Construct</th>
<th>Brief Construct Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Intervention Characteristics</td>
<td>Intervention source</td>
<td>The intervention was developed internally to solve a problem or externally by vendor or research group</td>
</tr>
<tr>
<td></td>
<td>Evidence strength and quality</td>
<td>Stakeholders’ perceptions of the validity and quality of the evidence for the intervention</td>
</tr>
<tr>
<td></td>
<td>Relative advantage</td>
<td>The advantage of one intervention versus another or versus the status quo</td>
</tr>
<tr>
<td></td>
<td>Adaptability</td>
<td>Degree to which the intervention can be tailored to meet local needs</td>
</tr>
<tr>
<td></td>
<td>Trialability</td>
<td>Ability to pilot the intervention</td>
</tr>
<tr>
<td></td>
<td>Complexity</td>
<td>Perceived difficulty of implementation and the degree of difference from current practice</td>
</tr>
<tr>
<td></td>
<td>Design quality and packaging</td>
<td>How the intervention is bundled and presented</td>
</tr>
<tr>
<td></td>
<td>Cost</td>
<td>Cost of the intervention and of its implementation</td>
</tr>
<tr>
<td>2. Outer Setting</td>
<td>Patient needs and resources</td>
<td>The degree to which the patient is at the center of the intervention to be implemented, how well the intervention is perceived as meeting their needs and overcoming their barriers</td>
</tr>
<tr>
<td></td>
<td>Cosmopolitanism</td>
<td>The degree to which organization is networked with others</td>
</tr>
<tr>
<td></td>
<td>Peer pressure</td>
<td>The degree to which, at the organizational level, there is pressure to catch up to what others are doing</td>
</tr>
<tr>
<td></td>
<td>External policies and incentives</td>
<td>External incentives including: policy and regulations, external mandates, pay for performance</td>
</tr>
<tr>
<td>3. Inner Setting</td>
<td>Structural characteristics</td>
<td>Social architecture, age, maturity, and size of the organization</td>
</tr>
<tr>
<td></td>
<td>Networks and communications</td>
<td>Social networks within the organization, the sense of “teamness” and the effectiveness of communication within the organization in supporting these networks</td>
</tr>
<tr>
<td></td>
<td>Culture</td>
<td>The norms, values, and basic assumptions within the organization</td>
</tr>
<tr>
<td></td>
<td>Implementation climate</td>
<td>The organization’s absorptive capacity for change, the shared receptivity of individuals in the organization, and the extent to which the change will be rewarded, supported, and expected</td>
</tr>
<tr>
<td></td>
<td>Readiness for implementation</td>
<td>Differs from climate by inclusion of specific, tangible signs pertaining to implementation of the intervention including: leadership engagement, availability of resources for implementation and ongoing operation, and access to information and knowledge about the implementation</td>
</tr>
<tr>
<td>4. Characteristics of Individuals</td>
<td>Knowledge and beliefs about the intervention</td>
<td>Attitudes toward and value placed on the intervention, as well as familiarity principles and evidence about the intervention</td>
</tr>
<tr>
<td></td>
<td>Self-efficacy</td>
<td>Individual self-efficacy to implement the new intervention</td>
</tr>
<tr>
<td></td>
<td>Individual stage of change</td>
<td>Readiness for change at the individual level</td>
</tr>
</tbody>
</table>

(Continued)
advantage over the current standard practice, its ability to be adapted and tried out at the implementation site, its design and packaging, and its cost. Despite readily available instruments for some common disorders, brief, repeatable, and validated measures are not available for some less common disorders likely to be encountered in PC. Conversely, the number of options of measures available for common disorders can serve as a barrier as well. For example, in VHA, in the electronic medical record’s Mental Health Assistant (MHA), there are over 100 psychological assessments available for utilization. Thus, when policy does not dictate specific instruments, PCMHI individual providers within a facility may select and use different measures, making population-based outcomes measurement impossible. Finally, the practice of MBC in CCC involves multiple steps and complexity that require significant change from current common practice. The details of how measures will be administered and recorded, and at what points they will trigger treatment change for what conditions must be determined and agreed upon within each team, as no national policy or protocol outlines guidelines for instrument use. Addressing this level of complexity and implementing it in a busy PC clinic is clearly challenging.

Another pertinent intervention characteristic, which is a challenge for MBC implementation, is cost. Integrated care can be a difficult sell for leadership, which is constantly challenged to adapt to evolving priorities, staffing challenges, and a fixed budget. Although third-party payments are a small part of any VHA facility’s funding, work Relative Value Units (wRVUs) and other data are generally used as measurements of provider and clinic productivity (Kearney et al., 2015). Many critical activities, such as providing management advice during a team meeting or quick hallway consultations provide no workload credit as there are no Centers for Medicare and Medicaid Services (CMS) Current Procedural Terminology (CPT) codes for such activities. Additionally, it can be difficult to code for brief assessments and the health and behavior codes for assessment have a lower wRVU than other assessment CPT codes. If brief assessments cannot be readily employed in current treatment, which can be coded for, MBC implementation may be further slowed.

One way in which the administration of outcomes measures can be simplified is via technologies that allow patients to complete the measures on a tablet so that data is entered directly into the clinical database. However, access to these technologies, especially those with direct

<table>
<thead>
<tr>
<th>CFIR Domain</th>
<th>CFIR Construct</th>
<th>Brief Construct Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>Identification with organization</td>
<td>The way individuals view their organization and their own relationship to it</td>
</tr>
<tr>
<td>Other personal</td>
<td>characteristics</td>
<td>Individual traits that would be likely to influence willingness to change and success of the implementation process</td>
</tr>
<tr>
<td>5. Process of</td>
<td>Planning</td>
<td>The degree to which the process of implementation is determined prior to the start of implementation and the quality of those plans</td>
</tr>
<tr>
<td>Implementation</td>
<td>Engaging</td>
<td>The efforts taken to attract and involve individuals that should be included during implementation</td>
</tr>
<tr>
<td></td>
<td>Executing</td>
<td>The quality of efforts to accomplish the implementation plans</td>
</tr>
<tr>
<td></td>
<td>Reflecting and evaluating</td>
<td>Efforts taking to evaluate the process and success of the implementation based on qualitative and quantitative feedback</td>
</tr>
</tbody>
</table>
connection to the electronic medical record, is tightly controlled with the VHA. One strategy that has been successfully employed in VHA is the use of VetLink kiosks where patients check in for the appointment and can also complete questionnaires and screening instruments. Currently, kiosks are deployed at every medical center and at many CBOCs and expansion continues. Kiosks may allow a platform for administration of follow-up instruments for administration in the future. Further, VHA is continuing to pilot MBC technology and is currently piloting the use of tablets that do not contain patient health information but can be linked to the medical record by the clinician. Tablets were used in the implementation of CCC at White River Junction where providers initially found the routine collection of patient self-assessment measures to be only marginally useful (Zubkoff, Young-Xu, Shiner, Pomerantz, & Watts, 2012). However, at the end of a 6-month period of exposure to the information collected via this technology, they found both the format and the content of the patient self-report measures to be very useful and worthy of recommendation to other clinicians. It appeared that simply exposing clinicians to the technology and content of the information changed their attitude and improved the uptake of MBC. Therefore, while it is possible to overcome human and policy barriers with sufficient effort, like some other aspects of PCMHI across the VHA system, access to technologies to support MBC remains variable across sites.

Outer Setting
Outer setting factors include whether the intervention is perceived as meeting the cosmopolitanism of the site, peer pressure at the site level, and external policies and incentives. A major driver of healthcare practice and part of all VHA medical centers’ outer setting is policy. Currently there is a lack of policy requirements for standardized assessment outside of the CM component of PCMHI. Policy that has been developed and enforced for CM cannot be directly applied to CCC practice because of the differences is these two program components. Further, without the ability to extract data from all CM programs across the nation, it is difficult to measure full implementation of current MBC in the field. If MBC policies were put in place in VHA, along with consistent enforcement, it would help drive practice change in the direction of MBC.

Inner Setting
One of the challenges of implementation of any change in the very large system of the VHA is that the inner setting of each medical center differs—the culture, implementation readiness, structure, and processes of care. Therefore, local barriers to implementation of MBC will be highly dependent on the intricacies of each local system. Variable leadership support for changes in practice and differing structures and processes at each site will mean that there is no single pathway to implement MBC across the whole system.

The culture related to PCMHI at each VHA medical center has the potential to limit implementation of MBC. PCMHI programs may generate far fewer wRVUs than usual mental health care. Many critical PCMHI activities, such as providing patient management advice during a team meeting or quick hallway consultation, provide no workload credit, as there are no CMS CPT codes for such activities. The local culture may therefore value and provide more support to traditional general mental health and specialty mental health clinics. Despite demonstrated effectiveness in improving patient readiness for specialized treatment and reduction in the wasted clinic slots resulting from patients who do not show up or do not engage, many administrators resist moving staff into primary care, fearing they will have insufficient staff to provide more traditional treatments. As a result, high-functioning PCMHI programs may disappear when facility or clinic leadership changes.
Characteristics of Individuals

A final domain of potential barriers to implementation falls within the characteristics of the individuals who are adopting the change and includes their knowledge and skills related to the intervention, their self-efficacy and readiness to make the practice change, and their identification with their organization and other personal characteristics. Rogers’ (2010) studies of diffusion of innovations suggest that any innovation is likely to be picked up first by the innovators and early adopters. Innovators are well networked outside of their locality and intensely interested in new ideas. Compared to innovators, early adopters are more tied to their local systems. They are well respected by others and likely to become informal champions of change and local opinion leaders. Innovators and early adopters, however, make up the minority of individuals employed by the system. Judging by the current state of MBC, VHA, while further advanced in implementation of integrated care in general, is still in a state of early adoption related to MBC in particular. As barriers to implementation are overcome, MBC practice will be taken up by the early majority.

Staffing level is another key issue that may affect MBC implementation. Although VA has recommended a staffing model of 0.5 embedded providers and 0.17 care managers per 1,000 primary care patients, few sites are staffed to that level. Understaffing the program is likely to have a negative impact on the individuals who are attempting to provide PCMHI services. Overworked staff members who are struggling to keep up with demands of clinical work are unlikely to have sufficient time to be able to design and implement large changes in their processes of care, such as initiating MBC approaches. In addition to the mental health staff embedded in these settings, sites may be resistant to adopting PCMHI and MBC for mental health conditions in PC. Shifting from referral care to integrated care can also be a difficult proposition for overworked primary care staff, which is not accustomed to being part of their patients’ mental health care. Some sites have also rejected and resisted development of the PACT as well, preferring the autonomy of the traditional provider role. Entrusting a care manager to follow up with a patient by telephone with routine assessment rather than have a return appointment with a provider is difficult for many to accept. Primary care providers commonly have an initial view of PCMHI as an enhanced referral service that offers timely access in a convenient location rather than as a collaborative, shared effort. Despite this, even such a limited approach is progress and often a necessary step in the evolution toward full integration, shared responsibility, and shared MBC efforts.

Finally, in the case of MH providers, one challenge to MBC implementation may be the strong values placed on the process and outcomes of the clinical interview. Transitioning to a new model of intervention may be difficult for traditionally trained mental health staff (Pomerantz et al., 2009). Adopting brief assessments and treatments directed at a single problem is challenging for a clinician trained in the wide-ranging exploratory assessments commonly performed in mental health clinics. Many clinicians are trained to ask only open-ended questions and shun the structured self-report tools that support brief interviews and outcome management. MH providers may resist using structured patient-reported measures due to their belief that the clinical interview will lead to a richer understanding of their patients. With the exception of evidence-based psychotherapy protocols, few therapy training programs incorporate MBC. As a result, providers may not be well prepared to use outcomes measures meaningfully in day-to-day practice. Training resources will be needed to expand knowledge about the evidence base and recommended practices for MBC. Having the right staff with the appropriate competencies working in the program is essential. In addition to understanding and accepting this evolving model of care, clinicians working in primary care need to become comfortable in a fast-paced environment where being interrupted or asked for hallway advice is the norm and brief assessment is critical for success (McDaniel et al., 2014).
Facilitators of MBC in VHA PCMHI

It is possible to overcome barriers to implementation of new practice, and VHA has several notable facilitators to support implementation of MBC. First, while some aspects of technology support are challenging, VHA already has a unified electronic medical record that includes a feature known as the Mental Health Assistant (MHA). MHA includes a large number of psychological assessment instruments and links to the patients’ medical record so that scores can be pulled onto progress notes. It also includes a feature that allows a clinician to easily see scores over time and can thus help to facilitate MBC.

There is currently a growing movement in VHA for MBC implementation across all MH services, with a national workgroup working to determine standard instruments for national implementation across all VHA mental health services, policy development for implementation of MBC, and technology support for the initiative. As VHA specifies MBC standards of practice and develops and disseminates national policy, it will support the implementation of MBC by providing direct guidance to local leaders for implementation. Further, the implementation of interprofessional teams in VHA PACT will help encourage MBC in PCMHI. PACT-model care supports a culture of MBC for all medical conditions. Use of MBC can be explained to patients as similar to taking blood pressure measurements and adjusting care based on the results. In addition, PCMHI staff using MBC can more easily communicate and collaborate with the rest of the PACT team about their patients. Once MBC becomes standard practice in these interprofessional teams, its benefits in supporting the principles of PACT, team communication, and high-quality care are likely to help this important practice change be sustained.

Beyond basic implementation of PCMHI and MBC, many programs have developed innovations that continue to move integrated care forward in VHA. An innovation that may advance the uptake of MBC through a social marketing (Stead, Gordon, Angus, & McDermott, 2007) approach is placing MBC tools directly into the hands of veterans. VHA has developed 11 award-winning apps including the PTSD Coach, Prolonged Exposure (PE) Coach for PTSD, Cognitive Behavioral Therapy for Insomnia (CBT-i) Coach, Acceptance and Commitment Therapy (ACT) Coach, Cognitive Processing Therapy (CPT) Coach for PTSD, Moving Forward (a Problem-Solving Therapy app), Parenting2Go, and Mindfulness Coach. Many of these include assessment instruments for users to track progress, and some can transmit this information to the patient’s provider (e.g., PTSD Coach can administer a PCL and transfer that information to the therapist if the patient selects this option). Many of these mobile apps are available for utilization in association with a clinician-guided course of treatment and can become part of the initiative for full implementation of MBC in VHA. Even without specific protocols for clinician-guided use, these apps are a useful supplement to the PCMHI provider’s toolkit as they provide self-management support to those primary care veterans who may need less intensive services. As users of these apps become accustomed to tracking their progress using structured instruments, they are more likely to expect a similar approach to their care when they choose to engage in mental health treatment in PCMHI or in traditional mental health services.

The last domain of the CFIR pertains to the facilitation of MBC in PCMHI in the actual process of implementation including planning, engaging stakeholders, and executing and evaluating the implementation. At the national level, VHA has placed considerable effort and resource into supporting full implementation of MH programs. As described earlier in the chapter, VHA provides both training for evidence-based psychotherapies and program implementation support through external facilitation. These resources are nonexistent in most other settings.

In summary, potential barriers to MBC implementation are numerous both in VHA and in the community of mental health providers outside of VHA. However, by investing in PCMHI implementation nationally, VHA has devoted considerable resources to efforts that will lay the
groundwork for MBC. As an organized system of healthcare facilities, VHA continues to have numerous potential facilitators to MBC implementation, especially in the area of its ability to drive external facilitators of practice change. Most importantly, VHA has growing understanding and capacity for implementation facilitation that can lead to sustainable practice improvements in MBC.

CONCLUSION

The VHA has become a national leader in implementation of integrated primary care. The next step will be advancing VHA to full implementation of MBC. This accomplishment will allow VHA to move beyond program evaluation based on administrative outcome reporting (such as no-show reduction, wait time monitoring, and encounters in a program) to program evaluation based on patient and population outcomes generated from patient self-report data. A critical foundation that must be developed to ease administration burden for patients and providers will be the creation of a technological infrastructure to support such a roll-out (Robinson & Reiter, 2007). This technological infrastructure would include interoperability with the VHA’s electronic medical record and ease of data extraction for program and patient evaluation purposes. Further, identification of specific instruments to be used for each desired targeted conditions will be critical (Kearney et al., 2015b). Standardization of MBC requirements and operating procedures through national policy would further support the implementation process. As was learned from implementation of PCMH in VA, however, mandating MBC will not be sufficient for true change to occur (Kearney et al., 2015b; Leykum et al., 2007; Nutting et al., 2009). Broad stakeholder involvement will be critical for MBC to be fully realized in VHA PCMH, but also in any system desiring to take on this goal. The burden of implementation of MBC as standard practice for integrated primary care is well worth the far-reaching benefit of providing program outcomes from the view of the patient, upon which our work is appropriately centered.

REFERENCES


Since 2008, the Virginia Commonwealth University (VCU) Psychology Department, which includes both an APA clinical and counseling psychology PhD programs, has been gradually evolving its clinical training model to one that is predicated on integrated primary care. With time, our training program has become perhaps the largest in the country in terms of volume of integrated clinical care provided by trainees and the number of trainees who receive integrated care training. We have developed a threefold mission (Table 26.1): to develop clinicians who are equipped to meet future workforce needs in primary care, to use our trainee and faculty resources to address the extensive unmet mental health needs among the underserved in Richmond, and to add to the empirical base of literature supporting the effectiveness of the integrated care model in safety net settings. This initiative began with a focus on adults through our students with a clinical health psychology concentration. However, a recent redesign has extended our training to include trainees in our child/adolescent concentration and service to underserved children and their parents. To advance this mission, we have created a Primary Care Psychology Training Network (PCPTN) that now includes six clinics serving both children and adults, including three that serve largely indigent populations at our own urban academic medical center and three community-based clinics (Table 26.2).

This transformation would not have been fully realized without the forward-thinking advocacy of the American Psychological Association (APA) leadership in setting up the first-ever funding program for psychologists from the Health Resources and Services Administration (HRSA), designated as the Graduate Psychology Education (GPE) program. Beginning in 2002, these funds were made available for APA-accredited programs that “foster an integrated approach to health care services and address access for underserved populations by training psychologists to work with underserved populations and in areas of emerging need” (HRSA, 2016). Currently, the annual budget for GPE is $8.9 million and there are 31 grantees, including doctoral trainings (like the PCPTN), internships, and postdoctoral training programs.

Table 26.1 Three-Part Mission for the VCU Primary Care Psychology Training Network

<p>| Mission Part I | To equip future psychologists to meet the workforce demands of a changing healthcare system that includes integrated primary care as well as responding to calls for increased interdisciplinary training to improve health care. |
| Mission Part II | To address unmet mental health and behavioral health needs by providing brief, pro bono evidence-based behavioral health services in Richmond safety net primary care clinics. |
| Mission Part III | To undertake research to test brief behavioral interventions in primary care and demonstrate the overall efficacy of the integrated care model. |</p>
<table>
<thead>
<tr>
<th>Year Joined the PCPTN/Initial Funding Source</th>
<th>Type of Clinic</th>
<th>Patient Visits Per Year</th>
<th>Unique Patient Care/Training Aspect</th>
<th>On-site complementary specialty mental health services available</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCU Academic Medical Center Internal Medicine Residents Clinic (IMRC) 2009, Virginia Health Care Foundation</td>
<td>Residents Training Clinic/Department of Internal Medicine/Indigent care prime focus</td>
<td>17,000 annual visits</td>
<td>Co-training with Internal Medicine Residents, Pharmacy Residents and Medical Students</td>
<td>none</td>
</tr>
<tr>
<td>Daily Planet for Homelessness Primary Care Clinic 2010, Health Resources and Services Administration (HRSA) Graduate Psychology Education Program (GPE)</td>
<td>Federally Qualified Health Center</td>
<td>11,000 annual visits</td>
<td>Focus on homeless population/3 blocks from VCU Psychology Dept.</td>
<td>Outpatient mental health clinic on-site</td>
</tr>
<tr>
<td>VCU Academic Medical Center Complex Care Team Clinic 2011, VCU Health System funding</td>
<td>VCU Health System Primary Care Clinic for High-Utilizing Uninsured Patients with Multiple (&gt;7) Chronic Conditions</td>
<td>4,000 annual visits</td>
<td>Small, cohesive interdisciplinary primary care team providing intensive services; average of 9 sessions with behavioral health trainees per year</td>
<td>none</td>
</tr>
<tr>
<td>Fan Free Clinic 2013, HRSA GPE</td>
<td>Free Clinic with 30 paid staff and 500 volunteers</td>
<td>6,000 annual visits</td>
<td>Unique free clinic culture</td>
<td>Outpatient mental health clinic on-site</td>
</tr>
<tr>
<td>Hayes E. Willis Health Center 2015, Richmond Memorial Health Foundation, Virginia Health Care Foundation</td>
<td>VCU Health System Department of Family Medicine Clinic located in underserved community</td>
<td>9,000 annual child and adult visits</td>
<td>Community-based clinic serving predominately African American and Latino adults and children</td>
<td>MSW located at clinic and 1 day per week psychiatric nurse practitioner funded by grant to PCPTN</td>
</tr>
<tr>
<td>VCU Academic Medical Center Pediatric Residents Clinic 2016, HRSA GPE</td>
<td>Residents Training Clinic/Indigent care a prime focus</td>
<td>24,000 annual child visits</td>
<td>Child cases only; co-training with Pediatric 2016 newly designed team care unit residents; 2016 newly designed team care unit</td>
<td>Pediatric psychiatrist available 1 day per week</td>
</tr>
</tbody>
</table>
The genius of HRSA GPE funding is that it is a “two-for-one” proposition, whereby GPE supports the interprofessional training of psychology graduate students while also providing mental and behavioral health services to significantly underserved populations. Additionally, the GPE program requires that co-training occurs with two or more other health disciplines, such as medicine, nursing, pharmacy, or social work. This dovetails with the universal agreement among health policy experts that increased interprofessional training is a critical component to improving healthcare outcomes (Rozenzsky, 2012). Perfectly fit and timed to match our developing vision for VCU PCPTN, GPE has been largely responsible for its success.

The forward-thinking aspect of GPE was its recognition that healthcare reform was dramatically shifting resources and incentives toward greater investment in primary care, where costly chronic health diseases could be addressed earlier and more cost-effectively. Part of that vision included an emphasis on addressing the behavioral factors (e.g., medication adherence, exercise, diet, sleep) that cause or worsen many chronic conditions as well as the mental health issues (e.g., depression, anxiety) that exacerbate chronic health conditions.

Cost savings is a critical, win-win aspect of the new movement toward integrated care. The Milliman Report (Melek, Norris, & Paulus, 2014) used a large national insurance database to show that medical costs for treating patients with chronic medical and comorbid mental health or substance use disorders (MH/SUD) is two to three times as high compared to patients with chronic conditions who do not have comorbid MH/SUD disorders. The additional healthcare costs incurred were estimated to be $293 billion in 2012 across all patients in the United States, primarily due to increased use of medical services rather than behavioral services. The authors argue persuasively that this creates a large opportunity for medical cost savings through integration of behavioral and medical services and the treatment of previously untreated or undertreated MH/SUDs, resulting in possible savings of $26–$48 billion a year after factoring in costs of integrated treatments. Furthermore, these potential cost savings are more pronounced in underserved populations, because of higher levels of comorbidity and more limited access to behavioral treatments. These types of savings are exactly what healthcare reformers had in mind when they set in motion the gradual change from fee-for-service, “volume-based” healthcare reimbursement to “value-based” reimbursement, predicated on clinical outcomes and quality indicators, including behavioral health.

The integration of mental health and behavioral health into primary care is already a high priority for Medicare, the Veterans Affairs Health System, and the Department of Defense health system for active military and the Medicaid system, which has developed a waiver system for traditional fee-for-service payment models for states to pilot test programs that include integrated services and other innovative methods for improving care outcomes. Private health insurance plans are beginning to experiment with these same payment models and are closely watching the publicly funded health plans to see what cost-offsets are realized with value-based reimbursement models. Additionally, the movement toward a patient-centered medical home (PCMH) model of primary care in both the public and private sectors has fueled the hiring of psychologists to help meet certification requirements. As of 2011, primary care practices seeking recognition as PCMHs must track the use of evidence-based interventions for the treatment of at least one condition related to unhealthy behaviors, mental health, or substance abuse (National Committee for Quality Assurance, 2011).

In spite of the upcoming workforce demand for psychologists trained to work in integrated primary care, there is a serious shortage of clinicians properly trained to work in this setting (Hall et al., 2015; O’Donohue & Maragakis, 2014). Furthermore, it has been shown that trainees who are not exposed to integrated behavioral care training at the earliest, most formative stages of doctoral training are less likely to pursue careers in primary care psychology (De Master, 2011). Although formal specific competencies for practice in integrated primary care and guidelines on when to provide training in these competencies are still being developed, there is no question that psychologists trained
in primary care integration are in short supply and in high demand on the current job market (Hall et al., 2015). In the Veterans Affairs Health System alone, there have been hundreds of jobs created by their investment in the patient-centered medical home and integrated primary care models. Similarly, many large primary care practices that have recently adopted the PCMH model have invested in creating positions for psychologists, in part because of their anticipation of the new payment systems and because they have been influenced by other large practices that have included psychologists on their teams. We have had a similar feedback loop whereby our first wave of trainees in integrated care has proven to be very successful on the internship, postdoctoral fellowship, and job market, fueling student interest as well as faculty commitment to grow the PCPTN even larger.

Beginning in 2010, the PCPTN program has received three consecutive 3-year cycles of GPE funding, totaling over $1.8 million, from the HRSA GPE initiative to support this gradual transformation of our training model. Those funds have been used primarily to provide six faculty members with salary coverage to work with the PCPTN as well as 32 advanced trainees with paid fellowship positions to provide clinical services and leadership at our six primary care settings. We have also received valuable start-up and pilot funding from the Virginia Health Care Foundation and Richmond Memorial Health Foundation, two organizations invested in improving care for underserved populations in Richmond, for start-up efforts in specific clinics (see Table 26.2 for funding sources). During that process we have essentially redesigned our training model to focus on developing basic competencies that go beyond traditional clinical psychology doctoral training to include brief assessment and interventions for team healthcare settings, a population-based care focus, and interprofessional consultation and collaboration skills.

The overall outcomes of the PCPTN thus far include over 85 students trained in this model, over 10,000 sessions of pro bono care provided in six different underserved primary care settings, and a host of trainee-designed collaborative research projects that have provided evidence supporting the efficacy of integrated care for patients, trainees, and the healthcare system. The 2016 GPE grant for the PCPTN will greatly expand its clinical capacity due to the latest GPE requirement that 60% of funding goes toward student fellowships. We therefore anticipate being able to provide approximately 5,000 sessions per year, essentially tripling our prior number of pro bono sessions provided to the community on an annual basis. This chapter will provide some background on the demand for this training model, highlight its essential competencies and unique advantages, describe some interprofessional training activities, and provide evidence for favorable trainee outcomes.

**PCPTN FOCUS ON UNDERSERVED COMMUNITIES**

VCU is ideally located for training in providing integrated care to underserved and disadvantaged communities. The university is centrally located in the heart of an urban environment with numerous challenges related to economic, educational, and health disparities. Richmond, Virginia, has a poverty rate of 26% according to the US Census Bureau. Additionally, 56% of the urban population is African American with another 6% being Latino. Richmond has a long and complex history of social, class, and racial division. While some of these divides have lessened, much of the city remains a Medically Underserved Area (MUA) and Health Professional Shortage Area (HPSA) as defined by HRSA. The results of these challenges can be seen in Richmond’s lagging health indicators across areas such as longevity, infant mortality, and HIV infection rate (Virginia Health Equity Report, 2012). Similarly, although there is a high degree of known mental health morbidity in Richmond, only a small fraction of these patients are currently able to receive care from mental health professionals due to a shortage of agencies providing subsidized mental health services to low-income and minority populations.
Bringing university resources and skills to bear in meeting the needs of the underserved, indigent population in Richmond is core to the VCU mission and a key component of the University’s strategic plan. VCU is a top 50 community-engaged university according to the Carnegie Foundation. The leadership at VCU is highly committed to the community engagement mission, as exemplified by President Michael Rao’s recent statement that “we are proudly inseparable and indistinguishable from the communities we serve.” In 2014, VCU was selected as the top university for community engagement and honored with the C. Peter Magrath Award. The PCPTN represents an ideal opportunity to provide accessible behavioral health services within these communities and has received excellent support from the VCU administration. The VCU clinical and counseling PhD programs are the only accredited psychology training programs within 50 miles of the Richmond metropolitan area, so there is a moral and ethical obligation to leverage our training resources to address substantial unmet health needs in our community.

The VCU Department of Psychology overall has a substantial focus on both health psychology and on underserved and minority populations. The health emphasis cuts across all programs in addition to our Clinical and Counseling programs, with even our Developmental and Social Psychology programs conducting health-related research. The department was committed enough to the health and minority focus to create a separate research-only Health Psychology doctoral program that admitted its first class in 2010. Approximately half of our Department’s full-time, core faculty conducts research on multicultural psychology, with most of them focusing on health disparities in underserved populations. Our department is truly a unique place where students have access to many additional mentors in their area of research, and the focus on health disparities allows many faculty and students to collaborate across labs, across programs, and with other departments at the university. This focus also allows us to recruit an extremely diverse graduate student body, as some of our programs have had an approximately 50% racial/ethnic minority student body.

Our faculty on the PCPTN, in particular, are well matched to the task of both training students to work with underserved populations and for recruiting trainees from under-represented groups to train in these settings. To enhance the cultural and linguistic competency and to meet the needs of the diverse, underserved communities in Richmond, we have several faculty members on the team with specialized expertise in diverse populations, including Dr. Rosalie Corona (Latino health, bilingual), Dr. Carla Shaffer (Latino health, bilingual), Dr. Heather Jones (African American youth), and Paul Perrin (LGBT, international, and homeless). These four licensed faculty members have expertise in developing and implementing interventions for these specific underserved populations, respectively, making them ideally suited to enhance the cultural competency aspect of our training program. Additionally, three of these faculty, Drs. Corona, Shaffer, and Jones, are themselves from under-represented groups (Latina, Latina and African American, respectively) and therefore are ideally positioned to facilitate the recruitment of additional under-represented minority students into our PCPTN training program—another priority for GPE funding and our Psychology Department. Thus far, 11 of the 85 VCU PCPTN trainees, and 8 of our 25 trainees who received funding, have been African American or Latino. All the minority students who have graduated thus far are practicing in underserved settings, and we have recruited three new bilingual Latino students to begin training in the counseling and clinical programs in 2016.

The six safety net clinics (see Table 26.2 for their unique characteristics) included in the PCPTN provide more than 61,000 visits annually. Table 26.2 describes each of those clinics and their unique training and service opportunities. In the following sections we describe five specific underserved groups that these clinics serve: the homeless, Latino families, children and adolescents, the LGBT community, and uninsured individuals with multiple chronic medical conditions.
Homeless and Those at Risk for Homelessness

Within three blocks of the VCU Department of Psychology is the Daily Planet federally qualified health center—a full-service holistic center for the homeless population in Richmond. In 2015, the Greater Richmond area had approximately 700 homeless adults and children at any given time (Homeward, 2015), with a much larger group at risk for homelessness. According to the 2015 report by Homeward, 78% of these individuals were unemployed, 75% had a history of incarceration, 51% had a history of drug/alcohol problems in their lifetime, 42% reported a current mental health diagnosis, 41% reported a long-term disability, 27% had been victims of domestic violence at some point in their life, and 23% were veterans. Providing mental health services within a primary care setting for homeless individuals represents an innovative way to increase services to this population nationally. Among the homeless population, these intertwined mental health and behavioral issues are likely to be even more pronounced, and therefore the benefit from integrated services even greater (Kwan, Ho, Preston, & Le, 2008). The leadership of the Daily Planet is committed to providing holistic, integrated care and, as a result, reached out to the PCPTN in 2013 to see if there may be an interest in partnering together. One of the unique aspects of the Daily Planet is the fact that in addition to providing 11,000 primary care visits it also has a vibrant mental health clinic that provides more than 15,000 mental health visits per year. That unique setup allows for our behavioral health trainees in the primary care clinic to make rapid referrals to specialty mental health when the need is there and also to coordinate care with specialty mental health providers when they are seeing both types of providers. When this is the case, the behavioral health consultant is usually focusing on health behavior issues such as smoking cessation, weight loss, insomnia, or coping with chronic pain.

The Latino Community

Two recent papers examining integrated care in a safety net clinic serving a primarily Latino population have made a compelling case regarding the effectiveness of integrated care for reaching an otherwise underserved Latino population (Bridges et al., 2014, 2015). The PCPTN is strategically positioned to provide such integrated services to Richmond’s growing Latino population and provides training for culturally sensitive integrated care for the underserved Latino community. Richmond’s Latino population has followed the trend of other cities where there had not traditionally been a large Latino presence, growing substantially over the last two decades. In 2010, Latinos accounted for 6% of the city’s population at about 13,000 residents according to the US Census Bureau. Since 2000, the city’s Latino population has increased by 95% and continues to grow by larger numbers each year. The growth in the number of Latino children in Richmond is even more striking, with the number of Latino kids in the Richmond School District doubling every year between 2010 and 2013. There are now approximately 1,200 kids in Richmond schools whose first language is Spanish. The school system has struggled to find bilingual teachers and staff, and providing school-based counseling services for this population is even more difficult given the lack of bilingual service providers. The Southside of Richmond is where the Latino population has clustered, which places the Hayes E. Willis Health Center (see Table 26.2) in optimal proximity to serve these residents. In fact, a large percentage of pediatric services for these Latino children is currently provided by the Hayes E. Willis Clinic, which presented an excellent opportunity for integrating mental health services.

A community needs assessment by Dr. Rosalie Corona and colleagues (a supervisor for the PCPTN) indicated that 75% of Latino adults in the Richmond area were worried about mental health problems,
and 88% said that mental health services would be useful (Corona et al., 2009). Focus group members and key informants emphasized the process of immigration as a key factor in mental health concerns. Specifically, they noted that the immigration journey was stressful and traumatic for families. For example, parents discussed losses of loved ones and physical and sexual abuses. Acculturation-related stressors were also mentioned by parents and community members. Children were described as acculturating to American values and behaviors at a faster rate than their parents, which was perceived to cause stress/conflict in the family as family roles potentially shifted. A second unpublished survey by Dr. Corona to determine the extent of the safety net for mental health care found there are very limited bilingual services available in the Richmond area (R. Corona, personal communication, July 19, 2016).

Latinos have been found to utilize fewer mental health services than non-Latino whites, even after controlling for a lower overall incidence of mental health problems (Cabassa, Zayas, & Hansen, 2006). Other studies also demonstrate that Latinos are less likely to receive evidence-based interventions compared to non-Latino white groups in the same settings (Alegría et al., 2008; Young, Klap, Sherbourne, & Wells, 2001). It has been argued that this gap can be partially and effectively addressed nationally by an increase in integrated primary care services targeting this population (Bridges et al., 2014, 2015). However, it has also been noted that in order to be effective, such services have to be provided in a linguistically and culturally sensitive manner as well. The funding obtained from the Richmond Memorial Health Foundation to build an integrated service at the Hayes Clinic, addressing the Richmond Latino population, was done so with the goal of attracting bilingual trainees to the VCU PCPTN. It had been our experience that bilingual trainees often wish to train in California or in larger cities that have a more substantial Latino population. With the addition of our Hayes training site, as well as a newly licensed bilingual supervisor for that site, we were successful in attracting three new bilingual doctoral trainees out of the 16 new trainees admitted to the clinical and counseling psychology doctoral programs—more than we have ever been able to recruit in a single year in the history of the Department.

Children and Adolescents

Integrated pediatric primary care offers a great opportunity to provide behavioral interventions during an early stage in development, providing opportunities for prevention of mental health and medical morbidity later in life. As Kolko (2015) noted in his recent editorial in *JAMA Pediatrics*, “the integration of mental or behavioral health services in pediatric primary care is by all accounts a national priority at this time” (p. 894). In pediatric primary care for low-income and underserved populations, the gap between the mental health needs and mental health care and referrals is even wider than in adult settings. Garrison and colleagues (1992) found that while low-income families have higher levels of psychosocial needs, medical providers often are less likely to address psychosocial needs in this population relative to more affluent populations.

Table 26.5 shows results from an anonymous resident survey conducted in 2013, which demonstrated that residents felt very positively about the impact of the PCPTN on patient care, their overall training and willingness to work in safety net primary care in the future if integrated care is part of that experience. Using child psychology faculty expertise and students in our Child/Adolescent training track in the Clinical Psychology PhD program to launch a successful integrated care service in a safety net pediatrics setting has been a long-held goal of the PCPTN. The launch at Hayes has been very successful in terms of frequency of child and adolescent referrals and satisfaction among family medicine physicians and other team providers regarding the added value of having integrated care services.
The recent funding from GPE included support for relaunch of our integrated behavioral services at the VCUHS Pediatrics Clinic in the fall of 2016. This clinic now has 23,000 visits per year (30% growth since 2010). The key barriers in obtaining a sustainable number of referrals in our 2010 effort were the absence of a shared clinical space, the lack of a faculty “champion” to serve as a liaison to the trainees and psychology faculty supervisor, and a lack of continuity in the Psychology faculty supervision. The new launch at VCUHS has the benefit of being led by our expert in underserved pediatric behavioral health, Dr. Heather Jones, and is now located in a newly constructed facility that is fully designed for interprofessional interactions among all trainees and preceptors. We believe the lack of proximity of our staff to the residents and their preceptors was identified as a major factor in the absence of sufficient referrals. Additionally, to improve our PC psychology team’s capacities and facilitate a successful launch of this service, we will have two advanced child/adolescent track trainees assigned exclusively to this clinic each year. Due to funding limitations we only had more junior, concentration level trainees at the site when we attempted the same service in 2010.

LGBT Population

Founded in 1968, the Fan Free Clinic (now called The Health Brigade) was the first free clinic in Virginia and one of the first in the country. During the early years it focused on women’s access to oral contraceptives and provided primary care to the poor and uninsured. In the 1980s, it provided support services in the wake of the AIDS epidemic, establishing the first community-based HIV/AIDS outreach program in the state. For the past 10 years, it has specialized in providing primary care services to the transgender population and has expanded its mental health services. It serves local residents who live at or below 200% of the federal poverty level, which is $48,600 for a family of four, according to the US Department of Health and Human Services. The Fan Free Clinic serves about 10,000 people annually through its medical clinic (6,000 visits per year), mental health services, and HIV-testing services. The Fan Free Clinic provides an opportunity to work with underserved patients in a unique clinical environment where there is much to be learned by trainees. The culture of this free clinic is unique in that it is staffed by 30 paid employees and 500 volunteer workers and clinicians annually. These “mission-driven” volunteers usually function in a part-time capacity as a supplement to their regular employment or donate their skills as part of a retirement plan that includes giving back to the community.

PCPTN PARTNERSHIP WITH THE COMPLEX CARE CLINIC

The VCU Health System has earned high praise for its Virginia Coordinated Care (VCC) program, a care coordination program initiated in 2002. VCC is funded by the health system and enrolls over 20,000 uninsured individuals in the Richmond Metropolitan area in order to provide them free access to primary health care. The goal of this national model was to provide each indigent patient with a community-based primary care physician, to enhance the health of these patients and, at the same time, reduce, or at least control, the per capita cost of care delivered. In 2011, the VCC program established the Complex Care Clinic (CCC) to improve the quality of care and decrease the costs associated with the sickest patients from the VCC population. To qualify for the program, these patients needed to have five or more chronic illness diagnoses and a history of high utilization of specialty clinic, emergency room, and inpatient medical care. These patients often experienced hospitalizations for what are termed “primary care sensitive” diagnoses (e.g., diabetes, asthma, and hypertension),
conditions that when properly managed should not result in acute care needs. These hospitalizations are most often a function of poor access to care as financial, social, and other barriers to adherence.

“We’ve created this as a population health management initiative,” said Sheryl Garland, MHA, vice president for health policy and community relations for VCU Medical Center, “This Complex Care Clinic is really our first venture into true management of a population.” The CCC staff members are a comprehensive team of healthcare professionals that includes physicians, nurse practitioners, pharmacists, social workers and our behavioral health trainees who provide wraparound, holistic services for each patient. The intensely collaborative visits often include a wide range of these practitioners and address the barriers to care in order to connect the patient to the right resources and services. All patients are required to see the behavioral health consultants at least once for an evaluation and most of the patients continue to receive some services over time from the psychology trainees. The clinic’s model is to engage patients in their care to a level they had not achieved previously and to build a relationship of trust with the team through intensive engagement. This intensive team involvement with a small team that works closely together represents an even more intensive collaborative team experience relative to our other sites.

During the first year of operation, the CCC effectively achieved improved patient outcomes and reductions in utilization (Virginia Commonwealth University Office of Health Innovation Data and Analysis, 2013). Specifically, preliminary findings showed a decrease in hospital costs of $3.9 million for the 443 patients paneled to the CCC clinic between November 1, 2011, and October 31, 2012. These cost savings were attributable to a 38% reduction of inpatient stays and a 28% reduction in emergency department use. Measures of patient health outcomes indicated that as costs and utilization decreased, patient health over the same year has improved. For example, the percentage of patients within the CCC clinic with hemoglobin A1C under control increased from 45% to 55%, and the percentage of patients with their blood pressure under control increased from 31% to 50%.

While the CCC leadership has emphasized that our PCPTN trainees play a key role in the positive outcomes and have backed up this belief with funds to support a psychology trainee fellowship, research and evaluation is needed to determine the degree to which improvements are attributable to the inclusion of psychology within this interdisciplinary framework. Recent data collected by our service as part of a trainee dissertation (Worthington, 2015) found that among the 250 patients who were routinely engaged in receiving care at the CCC, 88 (35%) were also routinely engaged with the psychology trainees in receiving services. The mean number of psychology visits for the psychology-engaged patients was 10.81 (SD = 7.75) with a median of nine visits. Although this number is much higher than our median of two visits per patient in our other larger integrated clinic (Sadock et al., 2014), the intention of this service overall is to provide much more intensive clinical services for a smaller number of patients.

BUILDING A BETTER CLINICIAN: UNIQUE ADVANTAGES OF INTEGRATED PRIMARY CARE TRAINING

For effective practice in integrated primary care settings, there is a wide range of additional competencies that go beyond the training generalist clinical skills that are emphasized in most specialty mental health training settings. These skills, we argue, build a better overall clinician, regardless of whether they ultimately end up practicing in integrated care or in other settings. Detailed taxonomies created by Blount and Miller (2009), Robinson and Reiter (2006), and the APA (2015) consensus document on Competencies for Psychology Practice in Primary Care articulate the unique competencies that are instilled in clinicians as a result of integrated care training. It is beyond the
scope of this chapter to provide a detailed account of these competencies, but it is worth specifying several of them that are difficult to obtain outside of this unique type of training. Blount and Miller (2009) articulated several that can only develop in a milieu, such as primary care, where concise and fast-paced manner is the norm. These skills include: writing prompt and concise notes that focus on specific functional problems and recommendations that can be supported by other team members; routinely articulating to other team members how a behavioral health professional can help in a wide variety of primary care cases; effective use of “curbside consultation” to communicate with physicians and other team members; learning to “get to the point” when speaking with physicians, who have extensive time demands and communicate in that manner on a routine basis; and speaking sensitively and with clarity about a patient’s behavioral/psychological issues with a physician, in front of the patient, to facilitate a shared understanding and to model such conversations for physicians. Other competencies that are unique to this setting are the many skills needed to discuss health problems that straddle the psychosomatic domain, in creative ways that do not alienate the patient and yet get patient buy-in to focus on behavioral goals. Blount and Miller (2009) articulated this skill as being able to discuss “bodily symptoms that have no medical findings with patients in a way that promotes curiosity and coping in relation to the illness” (p. 118). Communicating in this domain is difficult because one is working against the Western bias of mind and body dualism. Persuasive ways of breaking through this bias and promoting a hypothesis-testing, experimental approach to trying out behavioral methods for managing physical symptoms across a wide range of health problems is a crucial art form, often using metaphors, for working in primary care settings.

Robinson and Reiter (2007) added the important concept of “practice management skills,” including using 30-minute sessions efficiently, staying on time when conducting consecutive appointments, completing treatment in four sessions or less, use of an intermittent visit strategy when suitable, using flexible patient contact strategies (e.g., including telephone), and appropriately triaging mental health and chemical dependency services. Two challenging consultation skills they articulated are aggressively following up with physicians when indicated (very difficult for junior trainees in particular) and focusing on recommendations that reduce physicians’ visits and workload.

The intervention skills are also challenging to master and because of the wide range of issues encountered, there is a great emphasis on learning on-the-fly, starting with learning one module from an evidence-based therapy approach and then learning the next module in time for the next session with the patient. The most commonly used evidence-based treatments have been described in a prior paper and are summarized in Table 26.3. Notable is the fact that there is a balance between ones that address traditional mental health issues and ones that address behavioral health issues (e.g., adherence, smoking cessation, weight loss, insomnia, nonpharmacologic pain management)—something we call the “two buckets” of our integrated behavioral services.

Primary care requires a unique set of clinical skills, and we try to emphasize that in our training approach. Generally, compared to a more traditional community mental health clinic, primary care requires a more flexible, adaptable clinical style. Primary care clinicians need to be able to think on their feet and be able to quickly switch tasks. We promote a mindset that one must be “interruptible” and able to thrive in chaos to succeed in primary care. This flexibility also extends to our “open chart” model in which patients are invited to schedule to see us anytime. All our sites work on a “No wrong door” policy, meaning that no matter the referral question our trainees will meet with the patient. This approach allows our team to provide more complete services for these underserved populations in addition to providing our students with a greater breadth of experience. As a result of this policy, our trainees are often on the front line for identifying serious mental health conditions and managing
### Table 26.3 Typical Interventions in the PCPTN by Visit Focus

<table>
<thead>
<tr>
<th>Problem areas</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>All problem areas</td>
<td>Introduction of services</td>
</tr>
<tr>
<td></td>
<td>Psychoeducation</td>
</tr>
<tr>
<td></td>
<td>Supportive counseling</td>
</tr>
<tr>
<td></td>
<td>Self-monitoring</td>
</tr>
<tr>
<td></td>
<td>Goal setting</td>
</tr>
<tr>
<td></td>
<td>Problem-solving</td>
</tr>
<tr>
<td>Depression</td>
<td>Play Your Cards Right (Lang, Norman, &amp; Casmar, 2006)</td>
</tr>
<tr>
<td></td>
<td>Behavioral Activation (Hopko, Lejuez, Ruggiero, &amp; Eifert, 2003)</td>
</tr>
<tr>
<td></td>
<td>Cognitive Behavioral Therapy (CBT; Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td></td>
<td>Interpersonal Intervention (Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td></td>
<td>Assertiveness Training (Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>Relaxation Training (Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td></td>
<td>Graded Exposure (Barlow, Farra, &amp; Cohen, 2002)</td>
</tr>
<tr>
<td>Pain</td>
<td>Introduction to the Gate Control Theory Activity (Keefe &amp; Somers, 2010)</td>
</tr>
<tr>
<td></td>
<td>Pacing and Behavioral Activation (Keefe &amp; Somers, 2010; Kerns, Sellinger, &amp; Goodin, 2011)</td>
</tr>
<tr>
<td></td>
<td>Relaxation Training (Keefe &amp; Somers, 2010; Kerns et al., 2011)</td>
</tr>
<tr>
<td>Smoking/substance</td>
<td>Stimulus Control (Penberthy, Wartella, &amp; Vaughan, 2011)</td>
</tr>
<tr>
<td></td>
<td>Urge Surfing (Penberthy et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>Motivational Interviewing (Britt, Hudson, &amp; Blampied, 2004)</td>
</tr>
<tr>
<td>Insomnia</td>
<td>Sleep Restriction (Rybarczyk, Lund, Garroway, &amp; Mack, 2013)</td>
</tr>
<tr>
<td></td>
<td>Stimulus Control (Rybarczyk et al., 2013)</td>
</tr>
<tr>
<td></td>
<td>Sleep Hygiene (Rybarczyk et al., 2013)</td>
</tr>
<tr>
<td>Weight loss</td>
<td>Weight Loss CBT (Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td></td>
<td>Motivational Interviewing (Britt et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>Stimulus Control (Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td>Medication adherence</td>
<td>CBT (Cully &amp; Teten, 2008)</td>
</tr>
<tr>
<td></td>
<td>Motivational Interviewing (Britt et al., 2004)</td>
</tr>
</tbody>
</table>

### Table 26.4 Psychology Trainee ($N = 54$) Ratings for a Subset of Questions From 20 Questions Assessing the Impact of Training

<table>
<thead>
<tr>
<th>Mean Score</th>
<th>“My training experience in Primary Care Psychology . . .”</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6</td>
<td>1. Substantially enhanced my interdisciplinary communication skills.</td>
</tr>
<tr>
<td>3.2</td>
<td>2. Led me to view primary care as a top choice for a career path if a job opportunity was available.</td>
</tr>
<tr>
<td>3.4</td>
<td>3. Led/will lead me to look for an internship that includes a similar rotation.</td>
</tr>
<tr>
<td>3.5</td>
<td>4. Greatly enhanced my confidence in working with challenging situations.</td>
</tr>
<tr>
<td>3.4</td>
<td>5. Provided excellent training in administering brief interventions.</td>
</tr>
<tr>
<td>3.9</td>
<td>6. Enhanced my belief that integrating psychologists into primary care will be important for reducing health care costs.</td>
</tr>
<tr>
<td>3.8</td>
<td>7. Overall, I am very satisfied with the Primary Care Psychology training experience I had.</td>
</tr>
</tbody>
</table>

*Note. Mean answers for graduate student trainees ($N = 54$) anonymously surveyed between 2012 and 2015 (0 = Strongly Disagree, 1 = Moderately Disagree, 2 = Neutral, 3 = Moderately Agree, 4 = Strongly Agree).*
crisis/at-risk patients, something that does not happen in many training clinics that screen patients for serious conditions (e.g., psychosis, suicidality). Further, formal termination of therapy services is rare in this setting. Often patients improve and no longer need to see us, but they are always encouraged to (and often do) come back for further sessions later on to maintain their progress or reinforce and relearn important skills.

Students also benefit from the flexibility of seeing patients at different time points in the course of their treatment, which provides them practice with quickly updating their knowledge base on a particular evidence-based intervention. We have a comprehensive resource book and Blackboard website that includes instructions, basic overview of treatment, screening measures, and patient education/self-monitoring handouts for each major evidence-based treatment we provide that facilitates this “just in time” learning. “Just in time” learning complements formal didactics and informal in-clinic teaching that provide a basic overview over each of these topics as well. This flexible and reiterative learning process allows students to start with a breadth of exposure to a wide variety of evidence-based treatments while also developing basic clinical skills and using basic interventions that are widely applicable across clinical problem presentations (e.g., relaxation skills, cognitive restructuring). Over time, they become increasingly proficient with each individual evidence-based treatment.

Primary care work also requires a strong ability to self-manage, as emotions can get intense in this busy setting where not all individuals (both staff and patients) involved may be skillful in navigating the chaos. Highly upsetting patient encounters (e.g., risk of harm, reporting of significant trauma, patients living in high-risk violent neighborhoods, frightening health diagnoses) are also not uncommon in the primary care setting, so considerable self-care skills and supportive relationships between students, supervisors, and clinic staff from other disciplines are also critical to success in this setting. Psychology trainees and supervisors can be helpful in modeling a balance of professionalism and compassionate collegiality for teammates. Our students have been asked to present self-care and stress management skills training for the staff several times over the years, and it is not uncommon

| Table 26.5 2013 Anonymous Survey of Impact of PCPTN on Internal Medicine Residents (N = 34) |
|-----------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|
| EFFECTIVENESS | Strongly Disagree | Somewhat Disagree | Somewhat Agree | Strongly Agree |
| Having psychology clinicians present has enhanced the care received by patients at Primary Care. | 3% (1) | 0% (0) | 23% (8) | 74% (25) |
| The presence of psychology clinicians has significantly enhanced the training at the Primary Care residency program. | 3% (1) | 6% (2) | 60% (20) | 30% (10) |
| TRAINING EXPERIENCE | | | | |
| CAREER PLANNING IMPACT | | | | |
| I am more likely to continue working in underserved care if I can work with a program similar to the Primary Care Psychology Program. | 19% (6) | 13% (4) | 50% (16) | 18% (6) |
| EFFICIENCY | | | | |
| I am able to see a greater proportion of my patients within the 20-minute appointment times as a result of the Primary Care Psychology Program. | 23% (7) | 32% (10) | 39% (13) | 7% (2) |
in team-based settings for psychology staff to become the “glue” that keeps teams healthy and working together in positive ways. In a similar way, we serve as models to our patients in communication skills, emotional self-regulation in frustrating circumstances, and helping one another cope with difficult and upsetting health information. In some cases, our role is focused on helping a family cope with an individual’s diagnosis, particularly when the individual’s capacity is hindered by cognitive or psychiatric status. In many ways, it can be helpful to think of the clinic as a “unit” in that we are responsible for all the patients seen, not just those we are seeing for direct services. We contribute to the care of every patient in the clinic by providing interdisciplinary “cross-pollination” both in terms of intervention delivery and formal and informal real-time consultations. This is a unique practice model compared to traditional environments for psychologists; for those who thrive in this environment, the rewards of interdisciplinary work are boundless.

**Unique Aspect of Supervision and the Vertical Team**

One unique advantage of primary care training is how well it lends itself to a vertical team model of training, similar to the one used in medicine and other healthcare disciplines. There are multiple ways our PCPTN program benefits from including students at all levels of training, even brand new clinicians. We use a precepting supervision model that capitalizes on the multiple levels of experience that may be present in the clinic (individuals present vary by clinic setting). The levels of preceptorship available include a licensed psychologist supervising all trainees, a postdoctoral fellow (at some clinics), senior students with several years of training in the primary care setting who have been selected as and funded with GPE and other grants for 20 hours a week as student leaders in the program, mid-level students, and new students. Precepting in this manner allows for live modeling and teaching of clinical skills as well as real-time delivery of knowledge needed to enhance the clinical encounter. Rather than using recording devices, supervisors and trainee preceptors can directly observe sessions with students, and provide corrective feedback and modeling during sessions. Further, students can also participate in co-therapy sessions with supervisors and trainee preceptors. Supervisors can assign cases to “match” the training needs of individual students to ensure a breadth and depth of exposure for each student and to shape clinical skills in areas that may need development.

Precepting is further beneficial in allowing for students to interrupt sessions in the case that they need more information or assistance from the supervisor in managing the clinical encounter. Particularly in the early stages of training, we encourage all students to consult with their supervisor before ending any session during which risk of harm to self or others was reported or detected. In our clinics, reporting of suicidal ideation occurs daily, thus our trainees get early high-volume exposure to assessing suicidal risk, developing safety plans, and/or referring to emergency psychiatric services. These types of encounters are highly stressful for early clinicians, and having the benefit of an on-site supervisor greatly improves student confidence and comfort with managing these difficult patient scenarios. Starting students early in their clinical training means that by their later years in the program, they are mature clinicians with diverse skill sets and the ability to teach the newer students alongside the supervisors and postdoctoral fellows. Further, having multiple levels of preceptorship allows advanced trainees (postdoctoral fellow and advanced students soon to leave for internship) with opportunities for leadership, teaching, and growing toward their own professional goals.

Our preceptorship model simulates that used in medical training, nursing, pharmacy, and other medical professions, thus further acculturating our students to training and working in medical settings. This approach emphasizes real-time clinical learning using real cases, which in our setting allows for precepting across disciplines. Psychology students collaborate with medical and pharmacy
residents frequently, providing a real-time exchange of knowledge that not only enhances the learner’s experience, but also enhances patient safety (e.g., collaboration between a medical resident and psychology student in determining a safe schedule of decreasing alcohol intake for an epileptic patient who is abusing alcohol). More formally, trainees from other disciplines shadow our students to learn more about delivery of mental and behavioral health services in the primary care setting. This prepares them to be informed referral sources and collaborators in supporting psychology interventions and principles in working to achieve best outcomes for patients. Finally, our students also shadow other disciplines during their sessions to increase their knowledge of health conditions, required self-management skills, and potentially unexpected effects of medications (e.g., antidepressants for chronic pain) that can be beneficial to their patients. This knowledge prepares them to appropriately collaborate with other disciplines in providing behavioral interventions and to be strong contributors to patient care in future interdisciplinary settings. Overall, the precepting model has proven beneficial in providing both formal and informal opportunities for students to learn in real time with real cases and to gain knowledge from a diverse set of professionals.

Skill Development Specific to Safety Net Primary Care

Training in a primary care safety net clinic requires an additional set of considerations and skills beyond other primary care settings, even beyond more traditional clinics that predominantly serve vulnerable and underserved racial/ethnic minority groups as well as low socioeconomic status populations. Safety net training in particular requires psychology trainees to consider and even integrate into clinical work the various social systems that affect patients’ lives, including unstable or nonexistent housing, chronic or acute stressors, inability to pay for medications, food insecurity, exposure to the elements and trauma, among many others. For example, one of our trainees had to figure out how to implement cognitive behavioral therapy for insomnia (CBT-I) with a homeless patient who slept at night by the river. His primary fear was that a snake would slither across him and wake him up in the middle of the night, as had legitimately happened to him several months before. The trainee successfully helped teach him relaxation strategies to calm his nerves when he was falling asleep, provided him with psychoeducation about stimulus control, and helped strategize with him about potentially safer places he could sleep at night. Many of our other trainees have had similar experiences trying to help patients deal with insomnia while trying to sleep in homeless shelters, where dozens of people all sleep in bunks in a single room and create a substantial amount of noise, understandably leading to trouble sleeping.

In part because of these extremely challenging life circumstances that safety net patients are often in, patients sometimes present to primary care with high severity, comorbid, and often undertreated or untreated psychopathology. Safety net trainees have to be ready to handle crises, to consult immediately with supervisors, and to make appropriate referrals to longer term or more comprehensive mental health services. Suicidality is an extremely common reason for a warm handoff from another provider, so trainees must be extremely proficient in suicide assessments and ready to consult supervisors. Several times a year, one of our supervisors provides formal didactics on suicide assessment, with many other informal suicide trainings occurring on-site or on-call.

Conversely, safety net clinics often involve working with patients who are psychotherapy naïve. Despite the high degree of physical health issues in the population, many patients have had limited engagement with health clinics and as a result, very low exposure to psychologists or other mental health professionals. Patients are often confused about why they are being referred to primary care psychology, particularly when they have been socialized to seeing their problems (e.g., chronic pain, sleep problems, diabetes management, gastrointestinal issues) as being physical in nature. Trainees
often have to start with the basics of psychoeducation about the connections between mind and body and describe to patients the various ways in which their health behaviors can affect their physical health. Trainees many times often have to be more concrete in their approach to psychoeducation and to treating presenting problems. Traditional insight-oriented psychotherapy in particular does not work well in safety net clinics, where patients’ concerns are often much more pressing or specific than trying to understand the deeper motivations for their behaviors.

Safety net clinics also provide trainees with excellent opportunities to work in Federally Qualified Health Centers, free clinics, and other sites where there is a premium on social justice. Many safety net staff members choose to work in their clinics because they value giving back to the community and bringing services to populations that are all too often marginalized in traditional healthcare settings. It is an impactful experience for students to interact with these providers, as they regularly inspire values of social justice within our trainees and show them how to actually live out those values in a professional capacity. Many of our safety net trainees feel they are legitimately making a difference in their patients’ lives because even providing a small amount of help to a patient who is marginalized and in a vulnerable situation can have a remarkable impact. Furthermore, when these services can be provided in a culturally and linguistically competent manner, supported by supervisors with expertise in these areas, it can be a tremendous career-shaping, formative experience for trainees.

One highly unique aspect of training in safety net clinics is that trainees learn to function as de facto social workers or patient advocates, as often patients are in need of tangible assistance or program resources. Trainees have to learn how to advocate for their patients within the healthcare system, particularly when their patients are used to being marginalized or unheard. Patients often do not know how to ask their providers for certain types of care they need, so helping patients clarify or write down questions they want to ask their providers in upcoming appointments is a common intervention in primary care psychology. Also, the presence of many unmet needs tends to come out in primary care psychology, and often our trainees are aware of resources (e.g., housing assistance, food pantries, programs, showers) that patients are unaware of. So simply providing a list of additional places where patients can go to get help, or helping patients make appointments at those places, is an important, tangible contribution to patients’ lives.

CASE EXAMPLE

The following case illustrates the challenges and rewards of providing behavioral health interventions in a safety net primary care setting.

A 34-year-old black female, AJ, was referred to the PCPTN team for smoking cessation assistance by her primary care physician (PCP). Her PCP noted that she has been smoking for 20 years, up to two packs a day, and has untreated asthma. Helping AJ quit smoking was particularly important because she was awaiting spinal fusion and laminectomy and her surgeon informed her she must be tobacco-free for 2 months prior to surgery. Further, her PCP was concerned that AJ has not taken any of her medications for 2 months and now has dangerously high blood pressure and is reporting an increase in depressive symptoms. Her PCP had administered a Patient Health Questionnaire-9 (PHQ-9) during their appointment, and AJ had scored moderately severe depression (19/27) and denied suicidal ideation (SI).

Initial contact with AJ was during her appointment with her PCP through a warm handoff, a 10-minute meeting between one of our team members and the patient. During this initial meeting, AJ reported that currently her stress levels were extremely high and that was the most pressing issue she was interested in addressing. She felt that if her stress could be managed it would be easier to quit
smoking and engage in other healthy behaviors. She also reported that part of the reason she had not taken her medications was financial difficulty. At this point, she was scheduled with our team for the following week and for an initial appointment with social work for the same day to explore financial support options. A referral to long-term outside counseling was made, but AJ declined due to financial constraints.

Between records and the first complete session, AJ’s full history was obtained. She reported smoking for 20 years, having cut back from two packs a day to one pack currently. She smoked to relieve stress primarily. She reported being molested as a child by her uncle, who has since passed away and no charges were ever brought against him. At the time of this appointment, she lived in the house he molested her in with her grandparents and mother. Due to a car accident that resulted in her current back problems, she has been out of work for 5 years and is the primary caregiver for her grandfather, who was recently diagnosed with Parkinson’s disease. Her grandparents are aware that she is a lesbian and do not approve, resulting in them treating her like a servant around the house. Her only break from her current situation is going to appointments or going outside to smoke. She experiences chronic back pain and spends much of her free time lying in bed in the room she shares with her mother. Further, AJ described passive suicidal ideation, feeling like her desire to live was disappearing. She figured that because she could not afford her medication, why bother taking it and just allow herself to die. The primary intervention of the first session was supportive counseling, validating her stress and distress. Motivational interviewing (MI) was incorporated to affirm the actions she had taken to better herself and identify reasons to continue attending appointments. Her primary motivation was to not let her past and her grandparents dictate how her life turned out. She wanted to take back control. Finally, due to her SI a safety plan was established to manage her risk and address any possible changes in her SI. This plan included identifying friends and family she could talk to, providing her with hotline numbers, and establishing a plan for someone on the primary care team to reach out to her if she missed an appointment. She readily agreed to the safety plan.

AJ regularly attended appointments every 2 weeks at our clinic for 2 months with two follow-up booster sessions. The initial sessions relied heavily on MI to encourage her to restart taking her medications and increase her desire to quit smoking. She was able to get on a prescription payment program and began taking her medications after three appointments. Relaxation techniques were also taught early on, and she readily took to diaphragmatic deep breathing and guided imagery. These techniques were initially applied to her pain management, but she began to use these to manage stress instead of cigarettes and set a quit date for during month two of treatment. Once AJ had a handle of relaxation and was engaged fully in treatment, cognitive restructuring was introduced to address negative self-talk and beliefs stemming from her abuse and depression.

During treatment, AJ was referred to a number of other care teams at the clinic. She met with on-site pharmacy for a smoking cessation consultation and began using nicotine gum as a part of her quit plan, hoping it would accelerate her progress. Pharmacy and her PCP were also consulted about pain and depression management. They agreed to start her on duloxetine due to its depression and pain benefits. They also switched her blood pressure medication to one better suited for her. AJ met with on-site social work on three occasions to get assistance with her medications and to assist her in finding a new place to live.

Throughout treatment, AJ was assessed using a number of subjective and objective methods. Her depression and anxiety/stress were measured each visit with the PHQ-9 and Generalized Anxiety Disorder-7 (GAD-7). Scores on both measures declined from moderately severe (PHQ-9) and severe (GAD-7) to the mild range for both measures. AJ tracked the total number of cigarettes smoked each day, which decreased steadily until she quit completely. Finally, her blood pressure, which was monitored biweekly by her PCP, steadily declined into a healthy range (115/75). Student clinicians also noted
that AJ was more engaged in sessions, showed improved affective range, and reported more positive experiences by the end of treatment.

AJ returned for two follow-up booster sessions at 2 and 4 months after quitting smoking. These sessions served to reinforce her stress management techniques, check in on stressors, monitor psychotropic medications, and reinforce positive thought patterns. At the final session, AJ reported that she had been scheduled for her back surgery, had not smoked in 4 months, was living on her own and maintaining her relationship with her family, and had begun dating. She still experienced periods where depressive symptoms would appear and occasional desires to smoke but reported that she was able to manage them. Her abuse history was not addressed during most of the treatment due to AJ’s reports that she was past it. At her final visit, abuse resources were provided in case she felt she need treatment in the future. She was proud of her progress at the final session and plans to check in with our team as needed in the future.

INTERPROFESSIONAL TRAINING ASPECT OF THE PCPTN

An additional critical goal of the PCPTN is to meet the demand for an expanded workforce that is educated, trained, and prepared to practice in an interprofessionally focused, team-based delivery system. As Belar (2016) notes, interdisciplinary skills have “become an acknowledged critical element of collaborative team-based primary care—a centerpiece of healthcare reform” (p. 153). In order to instill interdisciplinary competencies in future clinicians, training programs need to be willing and able to move out of the traditional “silo” approach to training psychologists. Interprofessional practice includes the features of multidisciplinary and interdisciplinary models, but it also emphasizes shared duties, co-training, and developing team-based competencies (Janicke, Fritz, & Rozensky, 2015). There is a strong consensus among health policy experts that team-based, coordinated primary care provides better clinical and financial performance while reducing clinician workload (Rozensky, 2012). However, in spite of these apparent benefits and national calls for interprofessional training (Interprofessional Education Collaborative Expert Panel, 2011), collaborative training between psychologists and other healthcare disciplines remains relatively uncommon.

Psychology training programs have been called on to “reach across the aisle” to initiate and build co-training programs that create a new generation of trainees highly competent in healthcare delivery (American Psychological Association, 2015). To address the national need for increased interprofessional training, we place our trainees in three large training clinics that involve multiple informal and formal opportunities for “learning from and learning with” trainees from other disciplines, with a primary focus on internal medicine, pediatrics, and pharmacy residents. Much of the cross-discipline learning comes from informal interactions around clinical care of shared patients. Physician, nurse, and pharmacy preceptors located in our training clinics also provide secondary supervision, teaching, and consultation to our psychology trainees, and our on-site psychology faculty serve in the same capacity for the pharmacy and medical residents. In that regard, occupying a shared conference space for note writing and precepting is a crucial element in our success in cross-training. A recent study (Gunn et al., 2015) confirmed our experience that clinic layouts that promote interprofessional “bumpability,” and opportunities for on-the-fly communication led to four times as many integrated care conversations between primary care physicians and behavioral health providers.

As a further enhancement, the Internal Medicine residency program at VCU has employed a formalized requirement for shadowing our trainees in order to learn about issues that we address through behavioral treatments. Similarly, the same residency program has employed our service to provide a 10-minute training session, the “Clinic Minute,” during weekly morning rounds on a behavioral
health topic. Conversely, faculty from Medicine and Pharmacy also attend our supervision sessions periodically to provide didactics. For example, Pharmacy faculty and residents provide two lectures per year on psychopharmacology to our trainees. This type of sharing of teaching resources is one of the many benefits of interprofessional training. These training collaborations have evolved over the years as the faculty and the leadership in the residency program have seen the benefits of including psychology in their training experiences.

Internal Medicine residents, who are generally given a large voice in shaping their residency experience, have voiced strong support for working with the PCPTN. Table 26.3 shows results from an anonymous resident survey conducted in 2014, which demonstrated that residents felt very positive about the impact of the PCPTN on patient care, their overall training, and their willingness to work in safety net primary care in the future if integrated care is part of that experience. One resident who trained in our integrated care program even accepted a job at a safety net clinic on the condition that they hire a psychologist to begin providing integrated care. Residency training by physicians in an integrated behavioral care (IBC) clinic has been shown to have an impact on beliefs and future practice (Garfunkel et al., 2011). Among the 174 pediatrician alumni, those who trained and practiced alongside psychology trainees and their faculty preceptors were more likely to report feeling prepared for collaborating with mental health professionals compared to residents who trained at the same site prior to implementation of IBC or who trained concurrently at a site that lacked IBC. Additionally, alumni who trained in the IBC setting reported more experience in consultation, joint treatment, and collaborative patient care with mental health colleagues in their current practice.

Training in collaborative care is particularly critical for addressing behavioral health and overall health disparities in underserved populations. Specifically, the lack of collaborative care between mental health professionals and primary care physicians has been even more problematic with low-income and underserved patients, who have higher rates of noncompliance with prescribed medical care, unhealthy behaviors (e.g., smoking, obesity, high rates of alcohol use, poor sleep hygiene), and long-term dependency on pain or sleep medications that are designed for short-term treatment. Healthy behaviors and medical adherence for these individuals are further undermined by social stressors, which are known risk factors for poorer health. Additionally, cultural differences between patients and providers may pose significant barriers to communicating about mental health concerns. Thus, training that enhances communication and collaboration between medical and behavioral health professionals is even more critical in this setting.

In addition to co-training described earlier, our training program uses a formalized team-based learning (TBL) curriculum that includes three 2-hour sessions per year with the internal medicine, pediatrics, and pharmacy residents at VCU. The purpose of the TBL sessions is to have trainees work in small teams to collaborate together on formulating treatment plans for a fictional patient with common primary care diagnoses that include significant biopsychosocial components. To date, these TBL sessions have addressed such diagnoses as noncancer chronic pain, depression, dementia, diabetes care, insomnia, anti-coagulation treatment, HIV treatment and adherence, substance use disorders, generalized anxiety/panic disorder, and smoking cessation. These TBL sessions all include the following steps:

1. Each learner is assigned brief readings on state-of-the-art practice (e.g., with smoking cessation) prior to the session.
2. On the day of the session, learners are placed in teams composed of members from each discipline: psychology, pharmacy, and medicine.
3. Within their teams, learners individually complete an 8- to 10-item test that is based on the assigned reading. They then complete the test as a team and collaboratively discuss the answer choices.

4. Learning teams then work together on case-based, application-focused assignments. These group assignments will be designed to foster give-and-take discussions among team members in the context of treatment planning decision-making that requires teams to make a decision on what they feel is the “best” treatment plan option (A, B, C, or D) from a series of options that are all designed to be plausible.

5. Throughout the learning session, teams will come to a consensus on the answer to a question and then will be asked to display their answers using cards labeled A, B, C, or D.

6. Trained instructors facilitate discussion between groups asking teams to support their responses.

7. At the end of the 2-hour session, each team member is then asked to complete an anonymous evaluation of the components of the TBL, as well as the degree to which learning objectives are met.

The PCPTN has completed 15 TBL sessions to date, and the program has received excellent support from the leadership of all the training programs, as well as high levels of attendance and ratings of the learning experience by trainees (Bishop, Phillips, Lee, Sicat, & Rybarczyk, 2015). To contribute to the national effort to disseminate best practices in interprofessional training, four of the modules created for this series have been published in a peer-reviewed journal (Bishop et al., 2015; Dixon et al., 2013; Hobgood et al., 2015; Sicat, Willett, Breden, Rybarczyk, & Flack, 2012).

RESEARCH CONTRIBUTIONS BY THE PCPTN

As noted in Table 26.1, a primary mission of the PCPTN is to contribute the knowledge and evidence supporting integrated behavioral services, particularly within safety net clinics targeting underserved populations. This constitutes the third leg of our three-legged stool of the PCPTN—clinical service, training, and research. The specific aims of this mission include measuring clinical outcomes, measuring cost-offsets related to behavioral services, and examining high-priority healthcare issues where psychologists can make a contribution (e.g., professional burnout in safety net care and care of chronic pain patients). Because our funding sources like GPE are focused on training and clinical service, our research thus far has been primarily driven by student thesis and dissertation research. This is a win-win for the PCPTN and the students who are able to fulfill their research requirements in a setting where they are also providing clinical services. Furthermore, the medical staff and administration at these clinics have been very supportive of these opportunities to contribute to the evidence base for integrated behavioral care. These outcome data are also instrumental in providing evidence to funders as to the return on their investment. As can be seen in Table 26.6, which includes all of our completed student thesis and dissertation studies to date in primary care, their research has covered many important areas thus far. Next we will highlight some of our key empirical contributions, specifically in the area of clinical outcomes and cost-offset from providing integrated behavioral services.

A large body of studies thus far has demonstrated the effectiveness of brief behavioral interventions in the primary care setting for the two most common mental health issues of anxiety and depression (Cape, Whittington, Buszewicz, Wallace, & Underwood, 2010). However, these studies do not address the effectiveness of services in “real-world” integrated care settings, where treatment sessions are generally shorter and more widely spaced, and are provided to patients with comorbid and
overlapping conditions, especially in a safety net clinic serving underserved populations. Among the handful of studies that have examined the impact of integrated services on patients in a real-world setting (Bridges et al., 2014, 2015; Bryan et al., 2012; Corso et al., 2012; Davis, Corrin-Pendry, & Savill, 2008; McFeature & Pierce, 2012; Ray-Sannerud et al. 2012; Sadock, Auerbach, Rybarczyk, & Aggarwal, 2014a; Sadock, Perrin, Grinnell, Rybarczyk, & Auerbach, in press), results have been favorable in terms of brief interventions having a significant impact on depression, anxiety, or global functioning following brief treatment. However, only three of these real-world studies have examined outcomes in a safety net setting (Bryan et al., 2012; Corso et al., 2012; McFeature & Pierce, 2012). Among that small group of studies, only one used a standardized self-report measure (McFeature & Pierce, 2012) and none has employed a comparison control group or long-term follow-up measurement of outcomes. The main focus of our research thus far has been filling this gap in the literature. We believe our studies described next have made a novel contribution to the literature thus far in this area.

Table 26.6 Master’s Theses and Dissertations Completed by PCPTN Trainees

<table>
<thead>
<tr>
<th>#</th>
<th>Title</th>
<th>Author(s)</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Evaluation of psychological services provided in a university-based primary care clinic</td>
<td>Sadock, E. (2012)</td>
<td><a href="http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=3780&amp;context=etd">http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=3780&amp;context=etd</a></td>
</tr>
<tr>
<td>3</td>
<td>Evaluation of psychological services for anxiety and depression provided in a university-based primary care clinic</td>
<td>Sadock, E. (2014)</td>
<td><a href="http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=4605&amp;context=etd">http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=4605&amp;context=etd</a></td>
</tr>
<tr>
<td>7</td>
<td>Chronic pain causal attributions in an interdisciplinary primary care clinic: Potential discrepancies between patient and providers</td>
<td>Jensen, B. (2016)</td>
<td><a href="http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=5134&amp;context=etd">http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=5134&amp;context=etd</a></td>
</tr>
<tr>
<td>8</td>
<td>Determinants of stress and effects on performance in internal medicine residents</td>
<td>Braun, S. (2016)</td>
<td><a href="http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=4873&amp;context=etd">http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=4873&amp;context=etd</a></td>
</tr>
</tbody>
</table>

Table 26.7 Most Frequently Endorsed SRRS-R1 Items by a Sample of Primary Care Psychology Patients at ACC (N = 102)

<table>
<thead>
<tr>
<th>#</th>
<th>Top 10 most frequently cited stressful life events during the past year</th>
<th>% of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Experiencing financial problems/difficulties (bankruptcy, credit card debt, college costs, tax problems)</td>
<td>46.51</td>
</tr>
<tr>
<td>2</td>
<td>Death of close family member</td>
<td>41.86</td>
</tr>
<tr>
<td>3</td>
<td>Attempting to modify addictive behavior of self (smoking, alcohol, drugs, etc.)</td>
<td>40.70</td>
</tr>
<tr>
<td>4</td>
<td>Major injury/illness to self (cancer, AIDS, etc.)</td>
<td>31.40</td>
</tr>
<tr>
<td>5</td>
<td>Change in residence</td>
<td>29.07</td>
</tr>
<tr>
<td>6</td>
<td>Death of close friend</td>
<td>27.91</td>
</tr>
<tr>
<td>7</td>
<td>Major injury/illness to close family member (cancer, AIDS, etc.)</td>
<td>26.74</td>
</tr>
<tr>
<td>8</td>
<td>Discovering/attempting to modify addictive behavior of close family member (smoking, alcohol, drugs, etc.)</td>
<td>23.26</td>
</tr>
<tr>
<td>9</td>
<td>Being fired/laid off/unemployed</td>
<td>23.26</td>
</tr>
<tr>
<td>10</td>
<td>Assuming responsibility for sick or elderly loved one</td>
<td>18.60</td>
</tr>
</tbody>
</table>

1The Social Readjustment Rating Scale-Revised (Hobson et al., 1998).
Our first study to contribute to the identified literature gap was a master’s thesis project conducted by Elizabeth Sadock (Sadock et al., 2014a). This was an overall program evaluation of the first 2 years of HRSA-supported services at the ACC Residents Clinic. The published paper (Sadock et al., 2014a) included a description of the program, including types of patients served, their presenting problems, and treatments administered, as well as evidence of the impact of behavioral health services on primary care patients’ emotional adjustment and progress on behavioral goals. Intake and follow-up measures of depression, anxiety, smoking, insomnia, chronic pain, and weight loss were reported for 452 adult patients (mean age = 52; 59% African American; 35% uninsured) who were provided brief interventions (mean visits = 2.2) over a 16-month period. Although limited by the lack of a control or comparison group, these findings indicated that clinically significant effects were found for depression and anxiety in approximately 40%–50% of the patients receiving treatment. Promising effects were also found for a smaller group of patients receiving smoking cessation and weight loss treatment. Although these effects did not lead to a full remission in many cases, we argued they were acceptable given the brevity of treatment and the complex biopsychosocial issues that these patients have. In fact, in a companion study where we sampled the stress events exposure for 102 of these patients, cumulative exposure to a major stressful event was far higher than in the normative national population, with close to half of the respondents having had major financial problems and more than 40% having experienced death of a close family member in the prior year (Sadock, Auerbach, Rybarczyk, Aggarwal, & Lanoye, 2014b).

Our second study, a dissertation by Elizabeth Sadock (Sadock et al., in press) evaluating clinical outcomes at our largest clinic (the Internal Medicine Residents’ Clinic; IMRC), was designed to address the absence of any previous real-world study that examined long-term follow-up of outcomes in a safety net population. In this study, we compared the depression and anxiety scores of predominately low-income and minority patients who underwent brief interventions ($N = 147$) to those of patients from a demographically similar comparison clinic without integrated psychological services ($N = 139$), matched on baseline levels of anxiety and depression and length of time between assessments. As can be seen in Figures 26.1 and 26.2, participants from our clinic with integrated psychology services experienced significantly greater decreases in depression and anxiety scores than participants...
in the control clinic. The differences were most notable with anxiety symptoms (Figure 26.2), which are known to be more difficult to manage long-term with pharmacologic approaches. Even though aggregate changes for the intervention group do not result in remission of symptoms in most cases, these relatively modest outcomes must be weighed against limited access to mental health services and a frequent reluctance to seek traditional mental health services among underserved populations. These services are not being provided in place of traditional mental health services, but as a first step in care for those in high need for a starting point for behavioral health services.

Interestingly, one unintended positive outcome of this study was the realization of a new partnership with the clinic used as the comparison clinic, the Hayes Clinic. The fact that our data from this study showed that their patients did not fare as well as the patients in our community clinic in terms of managing depression and anxiety caught the attention of the leadership at Hayes and convinced them to invest considerable time into partnering with us to write a successful grant for resources to provide an integrated service beginning in 2015. Prior to our study, the medical director of the clinic and other family medicine physicians at the clinic felt that they had managed depression and anxiety effectively with the pharmacologic approaches, community referrals, and the in-house social worker they had at the clinic.

Our third study of clinical outcomes was a master’s thesis by Rene Grinnell (Sadock et al., in press). It served as an attempt to answer the question of whether our modest clinical gains posttreatment (Sadock et al., 2014) were sustained, reduced, or enhanced over time following the discontinuation of brief treatment. As noted earlier, no prior studies have examined treatment durability for brief integrated services in a safety net clinic. The only two previous “real-world” integrated care studies showed favorable long-term outcomes at a military primary care setting for active service members (Ray-Sannerud et al. 2012) and a British primary care clinic (Davis et al., 2008). But both studies employed a mail survey with low response rates of 10.5% (Ray-Sannerud et al., 2012) and 40% (Davis et al., 2008), raising questions about response bias.

In our study, we collected current PHQ-9 and GAD-7 scores from 47 adult primary care patients who had completed brief primary care treatment for depression and/or anxiety (mean sessions = 4.2) between 6 and 18 months prior to the follow-up data collection. Data were also collected on treatment received by patients in the interim and ratings by the patient of the helpfulness of the services.
provided. To reduce bias in our study, we collected data via telephone (to increase participation rate) and did not leave messages for patients who could not be reached directly (to avoid bias of those choosing to return phone calls). Among those who were reached by telephone, 95% participated in the survey. As can be seen in Figures 26.1 and 26.2, participants had declining depression and anxiety over time, with lower scores at the last session and substantially lower scores after longer term follow-up ranging from 6 to 18 months. Approximately 40% of the patients received other forms of treatment after terminating with our clinicians, but receiving therapy from other sources in the interim did not affect scores over time. Nearly 47% of the patients reported that the primary care psychology services were “very helpful” and another 23% reported that they were “somewhat helpful.” Overall, these data suggest that the behavioral treatments provided in integrated care may have their most significant effects after therapy is discontinued, as patients are able to implement the problem-focused approach that is emphasized over time. The last session they attend, when “posttreatment” scores were collected in our previous studies, may actually represent a time of higher levels of distress. As noted previously, patients are encouraged to cancel sessions if they are doing better and not in need of treatment.

Our final study was a thesis by Autumn Lanoye (Lanoye et al., 2016), which tested the hypothesis that our integrated behavioral services would result in reductions in costly medical utilization (i.e., emergency room visits and hospitalizations). Using the same large clinic and patients studied previously (Sadock et al., 2014), we conducted a large-scale retrospective pretreatment and posttreatment analysis with quasi-experimental control group constructed using propensity score matching. That technique consisted of creating a group of untreated patients that were matched 1:1 to each of our treated patients on baseline demographic, medical utilization, and diagnostic variables. The assumption in this technique is that those who did not receive treatment were not referred for behavioral treatment because their physicians made far fewer referrals or the patients were not detected as having a treatable problem because they had not been screened. We ended up with a group of 1,440 adult patients at a safety net primary care clinic, 720 of whom received integrated behavioral services and 720 of whom received medical treatment only. As can be seen in Figure 26.3, our analysis confirmed our hypothesis by showing that rates of preventable inpatient utilization decreased significantly among IBH-treated patients from 1 year before to 1 year after intervention (or before and after the matching date for the control group). Primary care utilization remained stable in both treated and control groups as hypothesized, indicating that the changes were not due to overall reduced utilization in either group. We inferred that the collective effects of these interventions may be a reduction in preventable inpatient stays, which are due to complications from chronic medical conditions. Our

![Figure 26.3 Number of Preventable Inpatient Hospitalizations by Group and Time](image-url)
interventions are aimed at improving treatment/medication adherence, improving general health behaviors, and treating mental health symptoms that may interfere with treatment adherence or contribute to poor health behaviors.

We believe the findings are important in building the case for integrated care in general, and for integrated care in safety net settings in particular. At an estimated cost of $10,000 per visit, the reduction of approximately 50 hospitalizations over a 3-year span relative to a matched control group (using propensity score matching) saved approximately $500,000 in primarily Medicaid and VCU indigent care costs. That easily covered the cost of our training program and supervisors, who account for only one full-time equivalent at the ACC clinic. This type of cost-offset research for integrated care services has been very difficult to undertake due to the necessity of access to inpatient and outpatient billing records for a given group of patients. The few prior studies that have been conducted are limited in scope because they only examined patients with high rates of utilization (i.e., “frequent attenders”; Cummings, O’Donohue, & Ferguson, 2003) or patients with a specific mental health diagnosis receiving a specific treatment protocol (Felleman, Athenour, Ta, & Stewart, 2013; Unützer et al., 2008). Thus, the cost-offsets achieved by these studies cannot be generalized to the clinic operation in its entirety. Furthermore, no studies could be found that have evaluated the cost-effectiveness of these services delivered in a safety net primary care setting.

We will continue to examine our outcomes as the PCPTN grows its database of outcomes across our six clinics and begin providing expanded services to approximately 5,000 adults and children in the upcoming year. A high priority will be a dissertation project being conducted by the fourth author (Z. Radcliff) examining the impact of integrating our services on the children and their parents at Hayes Clinic. This study will be an examination similar to our program evaluation study conducted for the adults we serve at the VCUHS Internal Medicine Residents’ Clinic and will review behavioral health and physical health outcomes. Further, qualitative data from patients and clinic staff will be used to supplement the quantitative outcomes. As part of routine practice, we conduct screenings of all child and adolescent patients with the Vanderbilt Assessment for Children (Wolraich et al., 2003), administered to parents of children aged 6–12 years to screen for ADHD and internalizing problems; the Pediatric Symptoms Checklist 17 (PSC-17; Gardner et al., 2007), a youth self-report aged 12–18 that screens for internalizing, externalizing, and attention symptoms; the Edinburgh Postnatal Depression Screen (Cox, Holden, & Sagovsky, 1987), completed by new mothers from the child’s first newborn visit to the 4-month well-child check; and the Modified Checklist for Autism in Toddlers (M-CHAT; Robbins et al., 2014), administered to parents to screen for autism spectrum disorder as the American Academy of Pediatrics recommends at 18 and 24 months. Behavioral health services are then provided to those patients and their families as needed. These measures are also to monitor treatment progress of patients being seen by the PCPTN team.

In the future, we plan to continue studies like the ones described earlier. Expecting to provide more than 5,000 sessions per year through the PCPTN going forward, we have ample opportunity to aggregate data and study outcomes as part of our mission. Having a large volume of patients is particularly important to doing integrated primary care program evaluation, because the aforementioned population-based approach to clinical care has an overarching goal to provide brief interventions to a wide range of patients. As noted earlier, this approach will inevitably lead to small changes in a large number of patients instead of large changes in a small number of patients.

**SUMMARY**

We believe that our experience in developing the PCPTN over the past eight years serves as a model and roadmap for other training programs to consider. As can be seen in the APA directory of doctoral
training program with elements of primary care training (American Psychological Association, 2016), VCU is one of only 10 programs nationally that is offering such an emphasis and is perhaps the largest one nationally in terms of number of trainees and hours of clinical service provided. The public and the profession need many more of these training programs to accommodate the changing landscape of primary care and to fulfill the promise of using integrated care to reach individuals at earlier stage in the development of mental health disorders and behavioral health problems. In making this assertion, we are emphatic in stating that integrated care is not meant to be a replacement for traditional mental health services. Rather, it serves as a vital supplement to provide services in a more time-sensitive, accessible, and cost-effective format. By employing the efficiency of integrated care with brief, timely, problem-focused behavioral interventions, psychologists can help fill the gap in behavioral health needs of a much larger population of patients than is possible in specialty mental health clinics. In a stepped care approach to mental and behavioral health, integrated care serves as an earlier, lower intensity form of intervention. These lower intensity, more convenient treatments are provided first and then more intense therapeutic strategies are applied in a progressive manner until satisfactory results are achieved (Mack & Rybarczyk, 2011).

Two pragmatic aspects to PC as a training setting not mentioned earlier are efficiency and cost-effectiveness of training. Doctoral students have myriad time pressures exerted on them throughout their training. They have research to complete, courses to take, assistantships, or employment outside of their training to financially support themselves, and they have to accrue enough clinical experience and client contact hours to be competitive for internship applications and beyond. Trainees in integrated care get excellent efficiency from their time spent on the service by virtue of the fact that they are almost always seeing patients during their assigned hours on PC. The schedule is booked to accommodate same-day “handoffs” as well as the typical number of daily cancellations and no-shows (a reality of safety net primary care in general). When there are more “shows” than usual, psychology trainees perform like other primary care healthcare professionals and shorten their sessions to accommodate the time pressures. Because there is no expectation of the “50-minute hour” in integrated PC, the patients typically do not feel shortchanged.

In terms of cost-effectiveness, one of the main barriers reported in the 2011 APA Primary Care Training Task Force survey of primary care training of APA-accredited training programs was the lack of resources for making a shift toward training in primary care (APA, 2012). The argument must continue to be made for public financing of doctoral level programs like the PCPTN and other programs nationally that serve as a model (see the Antioch University specialty PsyD program in New England). We also believe that even without direct funding from safety net health clinics or federal grants, training endeavors smaller in size could be financed in a large proportion of accredited training programs by shifting resources from costly university-based clinics operated by departments of psychology to primary care settings. University-based clinics run by departments are highly expensive to operate, especially when considering the relatively small number of clients that are provided with care each year. Our experience, and those reported by other training programs, is that safety net settings will gladly trade office space, computers, and other support staff resources in exchange for training programs providing high-quality pro bono services to their clientele. Most safety net settings put a high premium on providing training opportunities and greatly value an infusion of talented clinicians in training. Based on the extensive training benefits noted earlier and the fact that integrated primary care in safety net settings is more likely to provide care to those truly in need at no cost, rather than those savvy and resourceful enough to obtain reduced fee services at a small university-setting training clinic.

Understandably there will be some inertia to maintain the current model of practicum training occurring predominantly in specialty mental health clinics in the initial, foundational years of their training. University-based training clinics have become a convenient domain for faculty to supervise
in, to have control over selecting clinical cases that are viewed as optimal for training in traditional psychotherapy skills, to conduct clinical research in proximity to campus, and for departments and universities to lay claim on offering public benefit with treatment provided at reduced costs. The argument of these authors is that clinical psychology training will be improved and at the same time the public will be better served by greatly expanding the number of training programs that use safety net primary care settings as a foundational aspect of their training.

The perspective of the trainees on their experience has been crucial in developing and growing the PCPTN. Advanced and more junior trainees alike have reported a very high level of satisfaction with their training experiences, as the trainee survey reported in Table 26.4 notes. One particularly noteworthy finding is that students training in the PCPTN have expressed a strong interest in pursuing a career in integrated care, even though they may not have started out with that interest. As a result, the trainees who have graduated having trained in our program have “voted with their feet” in terms of what types of internships and positions they have landed in. Among the 55 students who have completed the doctoral program, more than 85% of trainees completed internships that include PC psychology/underserved care; currently eight graduates have leadership, staff positions, or postdocs in PC psychology, and another 36 have positions in medically underserved settings. As a snapshot from one year, 83% of the 2014–2015 PCPTN program completers are in positions providing services in medically underserved settings.

We believe upwards of 50% of the current in-house training of doctoral psychologists could be shifted to safety net integrated care settings, resulting in better trained clinicians and greater public benefit. We believe this can be accomplished without additional costs by a small amount of retraining and then subsequent redeployment of faculty to primary care supervision. Because overheads for operating safety net clinics are borne by the clinics themselves, there would be cost savings from reducing resources required by universities to operate their own clinics. It is possible that such a reduction in costs would in turn benefit departments and trainees alike. Additional research opportunities would follow from providing a venue for faculty and students to conduct research projects, including high-impact theses and dissertations (discussed earlier), which address social determinants of health, integrated care, primary care reform issues, and interdisciplinary elements that are otherwise not possible in the typical university training clinic. As can be seen in this chapter, our experience has been a game changer for the VCU clinical psychology and counseling training programs and faculty involved.

To move forward with such an agenda would require some progressive thinking and action on the part of the psychology training community. Changes need to happen in program philosophy, faculty incentives, course requirements, and practicum placements. There must also be a willingness to move out of traditional silo approaches to training psychologists. APA has led the way with an advocacy for this type of transformation, but other organizations within the psychology training community need to step up and do the same. Further funding and educational resources need to be created to facilitate this shift in training resources. One vital step that Division 38 (Health Psychology) in APA took in that regard in 2016 was the creation of a set of high-quality “plug and play” primary care psychology course modules designed for doctoral trainees and faculty wishing to develop their knowledge and skills in this area (see http://www.health-psych.org/ to obtain access). As of the summer of 2016, they have created four exceptional “foundation” modules, which included high-quality graphics, videos, and discussion topics, and plan to add as many as 20 more in the upcoming year. These modules would enable any program to easily add a course in primary care psychology to their curriculum. We applaud Division 38 for taking such action and hope that it sets the stage for more training programs to update their curriculum to join in this important transformation of clinical training. We look forward to seeing more indications in the near future that our field is ready to make this imperative, large-scale pivot in the clinical psychology training paradigm.
ACKNOWLEDGMENTS

The project described in this chapter was supported by Grant Numbers D40HP19638 and D40HP25724 from the Health Resources and Services Administration (HRSA), an operating division of the US Department of Health and Human Services. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Health Resources and Services Administration or the US Department of Health and Human Services. The project was also supported by grants from the Virginia Health Care Foundation and Richmond Memorial Health Foundation.

REFERENCES


Neftali Serrano et al.
Implementation at a FQHC

This chapter focuses on the successful decade-long implementation of an integrated care program, centered on the Primary Care Behavioral Health (PCBH) model, at a Federally Qualified Health Center (FQHC). FQHCs are private, community-based primary care clinics that focus on providing care to the medically underserved. The federal government provides enhanced funding and oversight to these clinics, which are present in every state, to assist the clinics in providing this care. The basic outlines of the service delivery model are described below along with the key factors believed to have led to its successful and sustained implementation.

HISTORICAL OVERVIEW

Access Community Health Centers in Madison, Wisconsin, began its journey into integrated primary care behavioral health in 2006. However, like many FQHCs that specialize in serving underserved populations, the impetus for integration was present for many years prior to its formal inception. The glaring needs of underserved populations, particularly uninsured and Medicaid populations, and the daily burden that lack of integration puts on primary care providers and their teams, makes integration a necessity. For these reasons integration models such as the PCBH have first proliferated through clinics like Access. Like many clinics of its kind, Access’s initial foray into integration required a fortuitous influx of talent and grant funding in order to develop what is now one of the leading integrated care programs nationally.

Talent, in particular, is one of the least discussed aspects of integrated care programs, as most attention is traditionally paid to the underlying service delivery models. However, given that the models of care have been clarified over the course of the last 15 years (Serrano, 2015), it is likely that more focus will be placed on the problems inherent with an underdeveloped workforce to run these models effectively (Miller et al., 2016). Much of the success of Access can be attributed to a steady influx of talented professionals.

The development of Access’s integrated care program began in 2006 with a 0.2 FTE position. The Access population, like many FQHCs, is diverse (40% minority) and primarily consists of the uninsured and Medicaid eligible patients with a small percentage of Medicare and commercial enrollees. Access was able to obtain federal and local grant funding to expand the first author’s time at their clinics, which at the time included two clinics serving approximately 10,000 patients annually. Hiring additional staff quickly became a priority, given the overwhelming psychosocial needs of patients at the clinic who had very little access to specialty mental health care, even with Medicaid insurance.
However, very few clinicians are trained in the primary care behavioral health model, which left a dearth of qualified professionals available for hiring. To meet this workforce need, Access developed a behavioral health training program. Students from varied professional programs, including social work and clinical psychology, interned at Access and received training in the behavioral health model. This pool of talented trainees then served as a well-trained hiring pool of professionals following successful completion of their internships.

By 2015, all but one of the seven staff members of the behavioral health team were former trainees. This core group of professionals had essentially grown up together, developed a common understanding of integrated care, and forged a professional identity firmly embedded in primary care. This shared professional vision resulted in workflows and common clinical pathways such that variations in service delivery between them was minimal. The staff now consists of three social workers and three psychologists who operate interchangeably under the title of Behavioral Health Consultants (BHC), a 0.25 FTE consulting psychiatrist, and two postdoctoral fellows.

The growth of the Access clinic occurred in the context of worsening access to specialty mental health services in Dane County, where Madison is located. This significantly impacted the nature of the integrated care services at Access as it became apparent that primary care providers could not ignore the needs of the severe and persistently mentally ill, populations that are not typically in the purview of primary care treatment. This pushed both the primary care teams and their BHCs to develop expertise to manage patients suffering from not just garden-variety depression but also more complex psychiatric illnesses such as bipolar disorder and schizophrenia. The BHC team was able to assist the primary care teams in identifying which of those patients could be safely cared for in a primary care environment, thus applying flexible but clear boundaries around the primary care service.

By 2015, Access had grown to three primary care clinics serving approximately 15,000 patients annually, and the behavioral health program was seeing one out of every five patients seen by primary care annually. The third clinic was acquired by Access in 2010 through a partnership with the University of Wisconsin Department of Family Medicine residency training program. Central to the success of the PCBH program at Access was the way that it was integrated completely into the life of the organization. From a workforce culture standpoint, the primary care providers consistently viewed and included the BHC staff as clinical peers. The clinic administration likewise treated the BHC team similarly to primary care providers and by 2011 elevated the role of the Behavioral Health director to a senior leadership position as Chief Behavioral Health Officer. The result of this full integration of the team into the lifeblood of the clinic was to elevate the status of behavioral health concerns to a level equal to that of the medical priorities of primary care and to allow both services to develop concurrently. The clinic also undertook efforts to educate all Access employees, from the Chief Executive Officer to registration clerks, on the nature of integrated care and the workings of the PCBH model, indicating how central it viewed the program to clinic operations. This unified approach helped to create a seamless feel of the service provided by Access to its patients, who perceive a unified service. By 2015 Access had developed a mature integrated care program (Serrano, 2014) with minimal staff turnover, a training program that had over 50 students from its inception, several publications (see the references section), and a financial footprint that was in line with the goals of the clinic.

**CORE PROGRAM COMPONENTS**

One of the strengths of the Access program is its well-defined approach to integrated care. The definition of an integrated care program is important because it provides a philosophy that guides the
development of the program over time. No one philosophy has been shown to necessarily be superior to others; however, the presence of a philosophy itself serves as a helpful guide. Here, the core components of the Access program are reviewed.

**Behavioral Health Consultation**

The core of the integrated care program at Access is defined well by the description put forth by Robinson and Reiter (2015) in their seminal text on the PCBH model, otherwise known as the Behavioral Health Consultant (BHC) model. The essential goal of this model is to provide generalist support to primary care providers and their teams across all conceivable behaviorally related presentations in a community sample of primary care patients. This means that the BHC is available to see patients with both mental health conditions and any medically related behavior change concerns such as weight loss, smoke cessation, or insomnia. The model emerged in recognition of the varied patient presentations that entail some behavioral component that primary care providers engage in the majority of their visits. The BHC must therefore be able to support the work of the primary care providers, whether it be assisting an individual with coping more effectively with their diabetes regimen, improving a patient’s motivation to reduce their alcohol intake, or providing skill training for a patient suffering from depression. Because research has demonstrated that over 70% of patient presentations in primary care have a behavioral component, the BHC model attempts to place a generalist provider of behavioral health services alongside the primary care provider.

At Access, this model is operationalized by placing one BHC in each medical pod or central area in the primary care clinic with a flexible schedule of patient follow-ups that allows for immediate or near immediate access for what are called “warm handoffs.” Warm handoffs are patient referrals at the point of care where a primary care provider enlists the support of the BHC on a particular issue. Warm handoffs are able to be integrated into the workflow of the BHC because the BHC does not maintain a traditional caseload, nor does the BHC see patients for the traditional hour-long therapy session. At Access, the BHC schedule of follow-ups includes six slots per day arranged toward the beginning of each clinic session (9 a.m., 9:30 a.m., and 10:30 a.m.; 1 p.m., 1:30 p.m., and 2:30 p.m.). Because typical visits ranged from 15 to 30 minutes and because warm handoffs are more likely to occur when clinic volume is highest, usually toward the end of clinic sessions, BHCs are able to flex their time and meet the demand of same-day patients. Typically 50% of patients seen by the BHC in a given day are warm handoffs. BHCs see between 8 and 12 patients on average per day.

Central to the model of care is the flexibility of the work style and workflow systems. To facilitate this, a BHC must be as visible as possible and have frequent contact throughout the day with the primary care provider and teams. At Access this is facilitated by the placement of BHCs in the medical pods and by providing several flexible ways for primary care providers to contact BHCs wherever they are. This includes including the BHCs in the clinic exam room flag system as well as providing the BHCs with a smartphone-based secure texting service. However, the most common way that warm handoffs are operationalized occurs via in-person communication that is facilitated by the unwritten rule that all members of the team are always interruptible. Therefore, medical providers interrupt BHC consults when needed and BHC providers interrupt medical consults when needed. This reinforces the interdependent nature of the work.

Due to the generalist nature of the model, most patient identification occurs through provider or patient self-identification of a behaviorally related issue. However, protocol-based screening with an instrument that combines the PHQ-9 (depression screen) and GAD-7 (anxiety screen) at all new patient appointments and physicals as well as all patients already identified as having depression on
their problem list also provides additional patient identification. In addition, a newer effort underway uses the Adverse Childhood Experiences survey to identify new patients with trauma history.

Formal test-based assessment is performed infrequently in this setting but does have a role. For example, brief cognitive testing with a Kaufman Brief Intelligence Exam (KBIT) is used in instances when the medical team is concerned about a child who may not have been identified appropriately for services in school. The Millon Clinical Multi-Axial Inventory (MCMI-III) is used in instances when documenting a formal diagnosis of a personality disorder may be beneficial. In general these test-based assessments are performed when there is no possibility of referral to external testing sources and the need is specific and valuable.

The BHC model is at the center of Access’s integrated care program because it provides the most impactful and broad support to the daily work of the primary care teams. In other words, this aspect of the program touches more patient lives and impacts the day-to-day work of Access staff most significantly. From the perspective of the Access clinic all other aspects of the integrated care program should support this broad, population-based effort. As a result the initial years of program development centered on effectively implementing the BHC model by introducing it to the medical teams, building staff competencies and ensuring adequate coverage for all open clinic hours. In addition significant effort was put forth to ensure that foundational administrative aspects of the service were set including documentation and billing and coding standards.

### Consulting Psychiatry

One of the main intervention strategies used at primary care clinics is use of medication, which is natural given the orientation of medicine. The BHC model can help balance the use of medication by providing patients and providers with behavioral alternatives (Serrano & Monden, 2011). However, medication will remain the treatment of choice or necessity for many primary care patients, which makes evidence-based prescribing that much more important.

While a good BHC should have a fundamental knowledge of medications and their common uses, the medication-related questions that primary care providers have may go beyond the expertise of the BHC. To answer this concern and because of the significant prevalence of severe persistent mental illness in the Access population, the BHC team developed a consulting psychiatry service in 2008. This service operates within the overall framework of the BHC model in that the BHC team houses and manages the work of the consulting psychiatrist, an essential element to ensure that the service is not over- or underutilized. The role of the consulting psychiatrist is to provide expert advice to the primary care provider and BHC related to evidence-based practice for prescribing. This takes the form of verbal and written notes that the psychiatrist writes in response to specific queries submitted by the BHC on behalf of the primary care provider. In some instances, a note is written after a one-time consult where the psychiatrist will evaluate the patient face-to-face. The most important criteria for the model to work is for the primary care provider to remain the prescribing provider at all times, thereby protecting the consultation role of the psychiatrist. The result of the model at Access has been that a patient population of 15,000 patients can be covered effectively with the expertise of a 0.25 FTE consulting psychiatrist (Zeidler Schreiter et al., 2013). In addition, the model has encouraged primary care providers to safely and effectively prescribe medication from every class of psychotropic medication currently available, which in turn has created access to thousands of patients who otherwise would have no or very little access to prescribers in the community.

The way in which consulting psychiatry works at Access (Zeidler Schreiter et al., 2013) is an expansion of a consulting psychiatrist role in the Collaborative Care model (Katon et al., 1995). At the point of care the BHC and primary care provider may see a patient with a complex presentation...
or with a pre-existing complex medication regimen. This results in a set of questions or concerns that the BHC then sends in a message delivered through the electronic health record to the BHC in charge of the consulting psychiatry queue. The message is sent using a template in the electronic health record to ensure that messages have a consistency to them that allows the consulting psychiatrist to answer optimally. The BHC in charge of the queue then presents these patients to the consulting psychiatrist during her time at the clinic. It is at this meeting that it is determined which of the patients will be scheduled for a face-to-face consult in the following week (usually the more complex patients) and which of the patients will be addressed with a chart review only. Following either a chart review or a face-to-face consult the psychiatrist produces a note using a template that has become familiar to the medical teams. Primary care providers also attend these meetings when available and can ask direct questions of the psychiatrist. Access has developed a toolkit for clinics interested in implementing this aspect of the model (https://www.hipxchange.org/Access).

Again, workforce talent is an issue with the consulting psychiatry service. Not all psychiatrists are comfortable in the role of a consulting psychiatrist. We have found that to be successful in this consulting role, a psychiatrist needs to be flexible and not be overly concerned with liability issues or have excessively conservative prescribing patterns. The psychiatrist also needs to understand the culture and practices of primary care, which are very different from traditional psychiatric practices, and be willing to teach in an effective, noncondescending way to both the primary care providers and BHCS. Essentially, the psychiatrist needs to embrace the consultative role in much the same way that the BHC does, understand that the key job is to support the role of the primary care provider and understanding that in many cases, even when it is not the optimal place for it, patients will only seek care with their primary care provider.

### Care Management

The care management service at Access is modeled after the care manager role in the IMPACT model (Katon et al., 1995), and modified to fit the available resources at Access. In a traditional IMPACT model the care manager manages a registry on an ongoing basis of patients with depression and makes periodic phone calls to patients, sometimes administering phone-based or in-clinic intervention, as well as assisting with the list of patients reviewed by the consulting psychiatrist. In Access’s iteration, the care manager role is assumed by three Behavioral Health Consultants who are given administrative time to review a registry of patients populated quarterly by patients with PHQ-9 scores of 15 or above, pediatric patients with a diagnosis of ADHD, pediatric patients seen by a BHC in the previous quarter, and any patients seen by the consulting psychiatrist in the previous quarter. The BHCS review these patient lists by performing chart reviews and making phone calls to patients based on their level of engagement, clinical progress, and other factors that would necessitate additional intervention. This methodology combines the best of the IMPACT model, namely the tracking of patients in a registry, with the flexibility needed when a dedicated care manager position cannot be supported financially. One other advantage is that the care manager role is then more closely tied to the clinical practice of the BHC team, serving as a monitoring strategy to ensure that patients are not lost to follow-up and thereby improving the quality of the service.

The concept of care management is based on the understanding that quality improvement requires monitoring processes. The care manager is able to have a more bird’s-eye perspective of care processes compared to clinical staff that are focused on the patient presentation in front of them. The care manager can generate data about subpopulations, such as patients with depression, that the clinic can use to track how well their processes are working. For example, in Access’s iteration a metric that is tracked is the number of charts reviewed that require some form of phone follow-up. This
provides an indication of how many patients are not receiving adequate care at the point of service or simply need more care or contact. Note that in the IMPACT iteration the care manager has the luxury of contacting all the patients according to an algorithm of care that dictates repeat PHQ-9s and interventions like behavioral activation. The Access iteration is an adaptation that better fits our resources and what our patient population is more likely to tolerate. For example, we did not have much luck with administering PHQ-9s over the phone to our patients.

The patient list related to consulting psychiatry is reviewed by a consulting ambulatory care pharmacist who does chart reviews for select quality indicators (Gallimore, Sokhal, Zeidler Schreiter, & Margolis, 2016), including which patients had medication recommendations implemented and which patients did not follow up with their primary care provider following a consult with the consulting psychiatrist. This list serves to monitor the efficacy of the consulting psychiatry service. Other indicators include which patients have had AIMS evaluations (i.e., evaluations for extrapyramidal symptoms often brought on by antipsychotics) performed with the last year and to ensure lab monitoring is up to date.

This creation and management of patient lists is fast becoming an essential element of primary care as practices are compelled to defend and improve their quality metrics. However, the development of management systems for these patient lists is not always turnkey. Access went through several home-grown iterations of systems starting with a simple spreadsheet before settling on a database developed internally (Serrano, Molander, Monden, Grosshans, & Krahn, 2012). Some clinics choose to purchase add-ons to their EHRs or prepackaged systems that run parallel to their EHRs. These often are quite expensive and difficult to maintain, so cost of implementation, life span of the purchased product, and maintenance costs need to be taken into account before deciding how to go about managing and tracking the work of the care manager or care manager role. In addition, because it is likely that a practice will change the patients it wishes to track over time, flexibility becomes a key attribute for any system.

**Health Promotions Clinic**

The prevalence of substance abuse in the population places a substantial burden on primary care practice. In response to this need, Access created an internal consultation service that mirrors what was created with the consulting psychiatry service. The service is led by a family medicine provider who specializes in addiction medicine. This addictionologist provides consultation to primary care providers working with patients with substance abuse issues, often working through the BHCs who again serve as go-betweens for communication in order to reduce the burden on the primary care team. Much like the consulting psychiatry service, the BHC may initiate a referral to the consult service through the EHR messaging system using a template, and this message is routed to the BHC in charge of managing the service. The addictionologist then reviews the referral, which includes the use of the Alcohol Use Disorders Identification Test (AUDIT) and Drug Abuse Screening Test (DAST) tools, and may elect to see the patient face-to-face or provide only a written consultation. One difference in the process is that whereas the consulting psychiatrist does not prescribe for any patients, the addiction medicine provider administers medication-assisted treatment (MAT), including suboxone and injectable naltrexone, to a subset of patients he follows. Otherwise prescribing is managed by the primary care provider.

Typical questions managed by this service include how to safely detox a patient in an outpatient setting, best options for treating alcohol or opioid abuse, or medication interaction questions related to a person’s addiction. Much like the consulting psychiatrist service impact, the health promotions clinic has been effective in raising the comfort level and perceived competence of primary care
providers in managing some aspects of addiction in situations where there are no other alternatives or where patient choice dictates that primary care is the only option. Another side effect of the service is that it has raised the level of discourse in the clinics around the prescribing of controlled substances (i.e., pain medications and psychostimulants). Because more information about patient habits around substance addiction is gathered and there are alternatives for managing addictions, providers are more empowered to set boundaries with patients who are requesting controlled substances while having an existing or historical substance addiction. This has led to more intentional prescribing that takes into account not just a patient’s history but also prods primary care providers to have a medical rationale for prescribing that is specific, documented, and evidence-based.

Training Program

One of the more difficult elements of building integrated care programs in the current era is the lack of ready-trained mental health professionals to run the programs. This problem is compounded in rural states and locations with shortages of any kinds of mental health professionals. Given this context the Access clinic began its integrated care program with the understanding that it would have to grow its own staff through a training program. This training program began with a partnership with the University of Wisconsin Health Psychology Post-Doctoral Fellowship. Beginning in 2007 fellows rotated through the Access primary care clinics complementing their hospital-based experience with primary care experience. Soon thereafter the Access clinics began taking on other types of practicum and internship students from varied programs, averaging 10–15 students per year. Students from disciplines such as Marriage and Family Therapy, Clinical Psychology, Counseling Psychology, and Social Work have come through the Access training program since its inception. Of the seven staff members in 2015, six were recruited through the training program.

The benefits of the training program exceed the presence of a ready workforce. One of the occupational stressors associated with all clinical work, and particularly high-volume clinical work like Behavioral Health Consultation, is burnout. One factor in combating burnout is work diversification, which the training program provides the Access staff with. While the assistance that students may provide to staff varies significantly, the act of performing in a teacher/mentor role ensures that staff have a variety of work activities to dilute the impact of the high-volume clinical work. It also provides a cue to reflect on their clinical work and that of their supervisees. The energy and enthusiasm that students can often bring to their work is also valuable to the well functioning of the Access staff. All staff operate as supervisors regardless of their professional background and all staff communicate with one another regarding the progress of supervisees, while two of the Behavioral Health Consultants act as “leads” in this area to ensure that the training is consistent and well organized.

The training model mirrors the clinical work, as so often is the case in parallel process supervision. Trainees begin with a period of shadowing of staff, learning work habits such as documentation, communication with primary care clinicians and managing flow, and then proceed to begin their clinical work with a period of reverse shadowing where they are followed around by the staff person who directly observes their work in exam rooms. After the supervisor(s) feel comfortable with the clinical acumen of the supervisee, they are allowed to work independently but still usually in tandem with a supervisor. In many instances the supervisor is present in order to allow for billing (a minimum of 16 minutes in most cases). If a supervisor has not spent the requisite time face-to-face with the patient, then billing does not occur. Supervision is planned and sequenced with training modules and checklists; however, there is no planned face-to-face sit-down supervision as in traditional models because this supervision model relies on immediate feedback and training following each patient encounter. In this sense, this style of supervision is significantly more intense than other models and
students tend to progress quickly, which is desired in a model that requires a breadth of skills and abilities. Typically by the end of the 6-month period trainees feel competent and can operate at least semi-independently.

Trainee selection is critical to the growth of an effective training program. The Access staff have learned over years that the best way of selecting trainees is based on an assessment of their temperament and basic predisposition to working in primary care. If a trainee demonstrates a strong desire to learn and an associated flexible temperament and has good core therapeutic skills, they are likely to thrive in Access’s training environment. By contrast, trainees who demonstrate rigidity, whether philosophically or in temperament, and trainees with poor baseline therapeutic skills are likely to struggle. The Access clinic has graduated over 50 students since the inception of the training program and has developed strategies to interview candidates to maximize success. Less than a handful of the students have been identified as “poor fits” by the staff.

**KEY FACTORS IN THE SUCCESS OF THE PROGRAM**

The success of the program at Access is due to a variety of factors related both to the program’s management, clinic characteristics, and the environment in which the clinic resides. These characteristics are discussed in the next sections.

**Vision Alignment With the Mission of Primary Care**

A general critique of early attempts to integrate mental health and primary care services is that many of those attempted to fit specialty mental health into the culturally distinct primary care environment. This resulted in programs that were disconnected from the lifeblood of the mission of primary care, which is a community-based and population-based mission incorporating primary and secondary prevention and intervention. The program at Access, in line with the PCBH model, intentionally aligned itself with that core mission subsuming cultural attributes of the mental health field and promoting itself as a true partner to the primary care clinicians. The over-riding credo of the program has been that the primary care provider is the “first customer.”

An aspect of the core vision was also to maintain model fidelity to achieve consistency of care and to integrate evidence-based components from other models promoted in the integrated care literature. To this end, interested clinics may want to consider adopting language and practices that help them determine whether their efforts are leading to greater integration and greater model fidelity. For example, the Center of Excellence for Integrated Care (http://coeintratedcare.org) has adopted language that differentiates between “programs” and “models” when describing a particular site’s implementation of integrated care. If a site is primarily developing methods and clinical pathways that are untested or unique to the site then that site is implementing a program. If a site is developing methods and clinical pathways that have been tested and disseminated widely then that site is implementing a model. All sites will have unique attributes that are “programmatic” in nature, but the degree to which a site adheres to a model will often correlate highly with their scores on instruments intended to measure the degree or level of integration. One such tool that sites can use is called the MeHAF (http://www.mehaf.org/learning-resources/tools-materials/), a tool with 19 domains covering both clinical service areas and administrative support related to integrated care. Using a tool like this can help guide programs to work closely with evidence-based components and measure progress toward integrated care.
Successful Engagement of Primary Care Providers and Clinic Administration

The program at Access has been primarily successful because it has been championed by primary care clinicians. This effort to obtain their support came through a variety of methods but has principally come as a result of lived experience with the program. The clinicians experienced the benefits of having the support of a BHC and this in turn convinced them that the program was essential to their work. A common refrain among providers at Access is that they would not choose to work anywhere without a PCBH program. Although efforts to gain their support included brief presentations of the model, email communications and joint educational sessions on psychiatric topics, these were not the most potent methods of gaining support. The core strategy, which ultimately proved successful, has been to be as viscerally helpful on a day-to-day basis as possible.

The engagement of administration in the effort was equally important but made more seamless by the pre-existing buy-in of key leaders, including the senior leadership team. While there was a general consensus related to the importance of mental health support for the patient population and strategically for the future of the clinic, what was needed was a clear articulation of what that support would look like and why it was different than simply having mental health providers on-site. The articulation of that vision usually takes multiple contacts with key personnel and much repetition of the core aspects of the model of care, including its population-based mission, its “provider first” mentality, its efficiency and its lack of ability to stand alone as an independent service. One of the ways in which this vision was disseminated and adopted was through the intentional placement of the integrated care program in the hierarchy of the organization. In some cases clinics begin an integrated care program but do not consider where in the organization to house the behavioral health staff and its leadership. In the case of Access, the program has a Chief Behavioral Health Officer who is a part of the senior leadership team and serves as a colleague to the Chief Medical Officer, Chief Financial Officer, Chief Human Resource Officer and the Chief Strategy Officer, who all report to the Chief Executive Officer. This structure facilitates the learning that an organization needs to engage in over time related to what it means to incorporate a behavioral perspective to the care of its patient population. It is this form of lived experience, including many small conversations and interactions that is believed (compared to formal presentations) to actually create cultural change and ultimately the best forms of integration. Creating structures and embedding the behavioral health team and its leadership in those structures can help facilitate this process of engaging the administration of a clinic. Primary care clinicians also heavily influence clinic administration given that they are the most expensive and vital resource most clinics have. Thus, the buy-in of Access primary care providers to the integrated model of care also significantly influenced the administration.

Sustainable Workforce Development

One of the most difficult issues facing integrated care programs is related to the inadequately trained workforce of mental health professionals, most of whom are trained for specialty mental health provision and need to be retrained to work in primary care settings. Given this reality, which is particularly challenging in geographical locations where there is a paucity of training programs and/or mental health professionals, attention needs to be paid not only to the first hire of a program but to future potential staff members. While the Access clinics reside in a medium-sized city, the availability of mental health professionals suitable for this model did not exist at the outset of the program. For that
reason the clinic developed a vision for training students at all levels to work in the model and to use this pipeline as a means of ensuring a stable workforce over time. This vision required a significant commitment because training requires much time and energy and therefore cost. However, the result has been a stable and ready pipeline of well-trained professionals. In many instances, when given the option of either hiring an established mental health professional and retraining them or hiring a student who has graduated from the training program, the clinic chose the student due to perceived superior performance and ease of transitioning to a staff member role. While this methodology may not be suitable or applicable to all settings the importance of having a sustainable approach to the integrated care workforce is essential. Many integrated care programs suffer when staff members transition out of their employment and due to a lack of a workforce development plan often have to restart their integrated care efforts when seeking to replace their lost staff members. Although considering workforce development may cost an organization more money, as is the case with training programs, it is arguable that the cost savings from having a steady supply of human resource is cost-effective in the long run.

Iterative Growth Toward Managing Special Populations

Each clinic has its own challenges based on subpopulations that require additional attention and resources. Managing these subpopulations within the framework of primary care is key to maintaining the population-based emphasis of integrated care programs while attending to the needs of a select portion of a population. This is a delicate balance, because focusing attention on a few patients may mean diverting resources from the primary mission of primary care to the entire community population. Programmatic management is also important because behavioral health providers who are already struggling with a transition to a primary care mentality may find themselves slipping back into a specialty mental health mindset if the program begins to divert resources to specialty populations. The Access program began with a focus on special populations only after the core generalist service was firmly implemented and the BHCs were fully developed in their respective roles.

The concept of iteration is one that is embedded in many practice transformation or change strategy approaches and applies well to growing specialized aspects of integrated care programs. The basic idea is to implement small change efforts, test those efforts, refine the efforts, and then grow the efforts incrementally. This basic approach is how Access approached each of the strategies for managing its special populations.

In the case of care management, the effort began with single provider pilots of managing a patient list with rudimentary technology (spreadsheet) and evolved to more providers and more complex efforts to engage technology (Serrano et al., 2012). At each stage of iteration the question was asked, “Is the effort involved with focusing on this subpopulation detracting from our ability to be available to our primary care providers?” which is a proxy question for, “Are we meeting our population-based mission?” If an effort was too cumbersome or inefficient, the practice was scrapped or reconfigured. For example, in typical IMPACT model protocols, which served as a conceptual basis of the care management scheme, care management requires several patient phone contacts within the first few weeks of SSRI initiation. Based on an assessment of our available resources, it was decided that this degree of patient contact was not sustainable across the patient lists. Instead, a quarterly contact scheme was developed, which was a more manageable workload.

In the case of our Health Promotions clinic targeting the substance abusing subpopulation, efforts began with piloting, pairing a BHC with the addictions specialist. Over time, the effort evolved into a model where the addictions specialist was also available for consultation to other providers through a
Implementation at a FQHC

referral method managed by the BHC lead in that area. Eventually a fellowship grew out of this effort, along with more sophisticated boundaries around who could be cared for effectively in primary care.

In each of these cases, the core primary care population-based mission was preserved by simply asking the question of sustainability and mission at each point of iteration and by being willing as a team and organization to reverse course at any time if the core mission was being threatened.

The “Hive” Approach to Patient Care

One of the most jarring aspects of the PCBH model for specialty trained mental health professionals is the concept of caring for a patient in a team setting versus an individual relationship typical of specialty mental health. This model is taken to its extreme at Access where patients often see more than one BHC in the course of their care, based on the way BHCs rotate clinics. This communal approach has become central to the model of care and has been effective for the functioning of the system of care and for patient care.

The decision was made early on in the development of the PCBH program at Access to rotate the BHCs across the three clinics. The reasoning for this was centered on the advantages this would provide for providing continuous staffing of the clinics. Furthermore, we hypothesized that the clinics would develop more consistently across each practice if the BHC team was sharing the clinics, as opposed to individual BHCs developing their own practices at each clinic. The result has been indeed that the service provided at each of the clinics is consistent and standardized. Beyond this, an interesting and unexpected side effect has been that the team began to develop a “hive” mentality. Because BHCs shared patients and practices, the level of cross-training and communication needed to sustain this forced the BHCs team to work diligently to provide each other feedback and resolve issues quickly as well as hone communication skills. This is especially true of patient care where there can be significant variation in styles and approaches. The fact that the BHC team was sharing these patients forced the team to adopt more consistent methods and to learn from each other’s styles while respecting the differences.

Patient response to the model was similarly positive, although not universally. The majority of patients are willing to trade off the advantages of comfort with one BHC relationship for the convenience of being able to see a BHC whenever they want to. And, due to the efforts by the BHC team to communicate effectively, such transitions for the patient visit to visit are minimized such that the patient has the feeling of being cared for by a group versus an individual. It is also not unusual for patients to provide us feedback on how the group is communicating and working, which provides an interesting vantage point from which to consider clinical work. For example, patients will often provide feedback on how the students are working with them and areas where they need to improve. Of course, the essential element in making this hive approach work is to alert the patient to this at the outset of care. One other side effect of the hive approach is to reinforce the primary care level of care that is PCBH. In other words, the patient understands that what they are engaging in is different than a therapist–patient relationship.

Not all patients respond positively to the hive approach. The minority of patients who express concerns with the approach are dealt with individually and respectfully. In some cases where the concern is raised at the outset of BHC contact, the patient is encouraged to try it out and see how it feels. In many of those instances the patient decides that the approach indeed works well for them. In some cases it is decided that the approach will not work for them and efforts are made to engage the patient with specialty mental health services in the community. In rare cases (maybe one patient per BHC) a patient will be accommodated to see just the one BHC, but this is strategically decided by
the BHC and the team. An example of how this hive approach is explained to patients is (as part of the initial introduction by the BHC):

We work together with your primary care provider to ensure that all your care is coordinated. This also means that it is likely that you will work with one or more of our behavioral health team over time based on when you come into the clinic. Generally speaking, if you come on the same day and same time, the same BHC is available. But in cases where that BHC is not available another team member will work with you and you will not have to repeat yourself or start over because we work very hard to ensure that the whole team is up to date on what we are working on with you.

The advantages of the hive approach cannot be overstated. It is likely singularly responsible for the well functioning of the BHC team and the evolution of the program. It is likely that had each clinic been staffed with its own BHC in a silo that the program would have evolved much more slowly and unevenly across the clinics. BHCs rotate through at least two clinics, often splitting their time evenly between the clinics to minimize the disruption.

**Long-Term Financial Vision**

The number one question that is asked is how Access’s program is financially sustained. Due to the complex and fragmented nature of the health system in the United States, that question is not simple to answer, nor is it generalizable to other settings, including FQHCs in other states. The usual answer is a general statement about the approach that the clinic administration has adopted with respect to the BHC program. This long-term and strategic vision for the role of behavioral health in primary care is what actually sustains the program. All other funding sources and revenue streams are temporary and likely to shift over time due to a variety of factors internal and external to the organization, but what sustains the program is a financial vision for the program’s role within the larger organization.

This role is best described as seeing the PCBH program as a value-creating entity that prepares the clinic for a future where payment is tied more directly to quality outcomes and to the ability of the clinic to care for larger populations efficiently. In other words, the PCBH program is a positioning tool for the clinic to care for more patients more efficiently without relying on the primary care provider to provide all of the care nor to provide all of that care in office visits. That vision is core to the Patient-Centered Medical Home movement and incorporates the tenets of the PCBH model.

While that future value proposition is being worked out, Access does need current funding to sustain the program. However, the experience at Access has been that such funding has come through a variety of sources. Much of that funding has come through federal expansion grants that are unique to FQHCs, but other sources have included private donor grants and revenue generation from fee-for-service billing. Additionally, given that FQHCs are provided “cost-based reimbursement,” which essentially means that its total costs for caring for patients is reimbursed through the Medicaid program (the way this is worked out varies state by state), the costs of the PCBH service are rolled up into those costs, a distinct advantage of the FQHC system. In addition, the presence of the PCBH service has attracted important partners, such as managed care entities and the state Medicaid office to the work being done at the clinic, which serves as a basis for working out reimbursement for the clinic in general. This side benefit of the PCBH service cannot be quantified but is tangible enough to clinic administrators that the program has never been questioned.

Based on experience and the current reimbursement environment the belief is that the best a PCBH program could do is work toward breaking even, which is essentially the best the Access program has been able to do in its life span when only considering fee-for-service payments. The fee-for-service
approach, which relies on face-to-face patient encounters and reimburses more for longer visits than shorter ones, leaves out a significant portion of the productivity of a BHC and therefore falls far short of adequate reimbursement for services. For example, none of the care management activity is reimbursable, nor are the curbside and other consultations that BHCs have with the care teams related to patient care. This said, it is believed that it would be a mistake to ask PCBH programs to be self-sustaining entities within an organizational budget. The role of the PCBH program is more akin to the role that nurses fulfill on the primary care team, and these are never asked to be self-sustaining but are rather assumed to be part of the cost of providing care at a medical home. In the same way, future payment strategies should assume that behavioral health care is a necessary component of care at a medical home.

Of course, this does not mean that program costs should not be monitored. Just as nurse FTEs are calibrated based on cost and productivity needs within a clinic, so too should the financial footprint of the BHC program be calibrated. In the case of the PCBH program at Access this has led to decisions about what provider types to hire and how to utilize students effectively. Access has maintained a 50% rule with its hires, providing for BHC roles for psychologists (which are generally more expensive but also have versatility for program development) and social workers (which are generally less expensive and bring an important perspective to the work). The postdoctoral program has also helped by providing a role for an early career professional at lower cost. All of these factors are weighed when considering the footprint of the service at growth junctures.

All of these factors and others need to be weighed when creating a business model. The key to success here however has been a long-term vision that is not focused on the PCBH program as a singular entity but rather as a strategic component within the strategy of the larger organization.

**KEY OUTCOMES**

Incorporating quality improvement practices into integrated behavioral health work assists in the ability to communicate what the work entails and why it is important. The PCBH model, with its population-based care focus, can be difficult to quantify. There is currently no one metric that can adequately describe the various facets of the work, from direct patient outcomes, patient satisfaction, provider satisfaction, increased access to care, decreased stigma, increased utilization of appropriate care, increased clinic morale, and direct or indirect cost-offsets. Therefore, it is useful to track a series of metrics that can evolve in an organic fashion. Communicating these metrics can create opportunities for funding or for program development and expansion.

Tracking metrics in an organized, ongoing fashion can both inform current practice and communicate to the administration, providers, and outside agencies about the work in a more concrete way. Given that many are still unfamiliar with the PCBH model, it is helpful to have data readily available to assist in describing the work and make it both more understandable and more accessible. The work of a BHC can seem fuzzy to many, that is, unclear what the goals are or even who it is for: the patients or the providers? The answer is, of course, both. The PCP is the primary customer, but all team members are focused on providing comprehensive care to the patients. Therefore metrics in each of the areas, addressing the needs of the various “customers” is essential.

The importance of tracking and measuring outcomes of the work within the PCBH model cannot be overstated. Having data to describe program outcomes is essential for both quality improvement efforts internally and communication with external partners. For example, the team at Access has used the data that has accumulated over time for various research projects to communicate the value of PCBH work to the healthcare community in Dane County, Wisconsin. The process of tracking
metrics, or specific types of data, was incorporated into the program from the beginning, growing more refined and robust over time.

**Population Penetration and Productivity**

While there are many types of data and metrics to consider, some will make more sense in one setting versus another. For the primary care setting, population penetration and productivity can be a place to start. Because the goal of the PCBH model is creating more access for patients (and more accessibility of BHCs for providers) these metrics are obvious cornerstones of a PCBH service. Population penetration is the number of unique patients seen by the BHC service in a given time frame (e.g., a year) divided by the number of unique patients seen by the primary care service, expressed as a percentage. Access has maintained a population penetration of close to 20% over the later years of its program development annually. A standard goal for this metric has not been established; however it is likely that 20% or more would be a suitable goal for a primarily adult population given that the prevalence of mental health conditions in a population is considered to be at least 20% (Satcher, 2000).

Productivity data can be captured in a variety of ways but the most common is patients seen per clinic session or patients seen per BHC. Typically, in a three- to four-hour clinic session, a BHC can be expected to see four to six patients in a well-developed program. While such data can be helpful to assess how effectively a BHC is being used, such data were not used as a BHC performance indicator or in a punitive style with the BHCs. It is important that BHCs be motivated by the population-based goal of the program and use such data to improve population penetration/access versus be concerned about hitting an arbitrary number.

**Clinical Outcomes**

Clinical outcomes are the most obvious type of data to track and consider, but also one of the most challenging to obtain in primary care. This is due to the generalist nature of the PCBH practice. BHCs do not simply see patients with only depression or anxiety, but rather the whole spectrum of mental health issues, subclinical presentations or psychosocial stressors, and medical comorbidities. Therefore, tracking outcomes becomes complex. The research literature bears this out, both in terms of the scarcity of outcome research in PCBH and the limitations in assessment tools specifically developed for use in a primary care practice.

There have been three measures used at Access as a starting point: the PHQ-9, the GAD-7, and antidepressant medication use. The PHQ-9 and GAD-7 measures are reasonable for use in primary care given validation on a broad sample, ease of use, availability in the public domain, and the ability to use them as both screening and outcome measures. However, the measures also have limitations, particularly with a sample of limited literacy. As mentioned previously, Access utilizes a care management program that involves quarterly tracking of high PHQ-9 or GAD-7 scores. This allows for overall monitoring of patients with higher scores, as well as tracking of scores over time. The PHQ-9/GAD-7 are administered at most BHC visits, while medical providers also administer the tools to patients for screening purposes in addition to evaluating treatment progress for patients with a pre-existing diagnosis of depression. Overall, however, the patients who complete these measures have a variety of diagnoses, not limited only to depression or anxiety. This means that these measures capture only a portion of the complexities of presenting issues commonly seen in primary care.

The PHQ-9/GAD-7 can be recorded within electronic health records (EHR), and thought should be given on where to document with an eye toward ease of pulling data reports later. The Access clinics use EPIC software, a well-designed EHR overall but still with limitations. It took trial and error to
figure out reasonable places for documentation that allowed for clinical utility in the patient’s chart (i.e., can providers find the results at a later time?) but also a location that could be mined easily on a quarterly or annual basis to track the results across the population. At present, the care management performed on patients with high PHQ-9 scores, for example, still occurs in a separate software program from the EHR using data extracted from the EHR. Documentation within the EPIC EHR occurs in set of tables called “flowsheets” and can be imported into notes using special phrases. All of this took a significant amount of work over years to develop.

Antidepressant medication use is another metric that has been discussed in the literature. Primary care providers tend to turn to medication first to treat depression and anxiety, which makes sense given their generally biomedical training backgrounds. Serrano and Monden (2011) demonstrated that the Access program was associated with positive changes in evidence-based indicators of depression care. For example, prescribing of antidepressants was observed to have significantly declined for patients newly identified as having depression following the integration of the Behavioral Health program. What this appears to indicate is that once medical providers had an option to offer behavioral health as a treatment, they used it preferentially above medication or that patients selected behavioral treatment more often than prior to the implementation of the PCBH program. Access does not track this measure on a consistent basis but has used this data to begin its evaluation of the program. This points to the reality that some metrics may only be useful on a one-time basis or periodically in the life of a program.

**Process Outcomes**

Process outcomes are another aspect of outcome data that are useful to track, and can give insight into what is happening day to day with the program both internally and externally with other portions of the overall healthcare system. Patient engagement and emergency department (ED) utilization are two examples of process outcomes, or variables that change as a result of the presence of the integration of the BHC team.

Patient engagement has been tracked in a few ways at Access. One is via the care management program, which involves chart reviews of patients in identified populations. A significant portion of the patients on this list have already had contact with providers in the clinic and do not require outreach, thus showing ongoing engagement in the medical system. Further, Serrano and Monden (2011) demonstrated increased patient engagement following integration of the BHC team at Access. Referral rates for outpatient mental health declined sharply as well, as patients were instead seen by the BHC team in the clinic. The flexibility of the PCBH model additionally allows for engagement of patients with more complex mental health issues such as severe and persistent mental illness or comorbid medical issues to be seen within primary care, a population that historically has difficulty with consistent engagement (Fondow et al., 2015). The patients seen the most often at Access by the BHC team represent the most complex patients seen, with more severe and persistent mental health issues, comorbid substance abuse issues, and/or comorbid behavioral health issues, again reflecting the ability of the PCBH model to engage patients effectively while meeting the differing needs of the population (Fondow et al., 2015).

The BHC team at Access is working on another process outcome measure related to the impact of the integration of care on the larger health system locally. A data set of local ED utilization was obtained from the area hospitals in Madison, Wisconsin, and the data analysis was in process at the time of this writing. The data set is unique in that all emergency departments in the city are included, assisted in part by the utilization of a common EHR (EPIC). The data for utilization of the ED has been compared for patients of Access with mental health diagnoses pre- and postintegration of the PCBH
model. A comparison clinic that did not integrate behavioral health services was included as well. For Medicaid patients across clinics, utilization of the ED dropped for the clinics where integration of care occurred. This result suggests greater engagement of the patients in their healthcare home, as opposed to other portions of the healthcare system (i.e., emergency department utilization). Overall, the exploration of this data set has revealed that the complexities of the patient populations make effective comparisons across clinics very difficult, even when it may appear on the surface that equivalent groups are being examined. Patients with Medicaid, for example, may represent fundamentally different groups such as patients temporarily on Medicaid between jobs versus disabled patients. This very complexity is another reason that gathering even basic program metrics is so essential. Metrics enable a way to begin the conversation on implementation/dissemination and program evaluation that eventually can be extended to larger portions of the healthcare system.

Program Outcomes

Program outcomes are another type of metric that are tracked at Access and involves model adherence, population penetration, and specialty population utilization. Tracking these over time can offer important information regarding trends in the patients seen, which translates directly into the daily practice of BHCs.

As suggested by Robinson and Reiter (2015), it can be helpful to track population penetration. This involves obtaining the number of patients seen by medical providers and the number of patients seen by BHC providers for a specified time period, then calculating the percent of the medical patients seen by BHC. Useful time periods for tracking include quarterly and annually. Population penetration at Access has moved from less than 15% to closer to 20%, but the goals for penetration may be different in different settings. Tracking population penetration according to populations of interest can also be helpful. For example, tracking pediatrics versus adult patients can assist in identifying any differences, which can then be addressed with providers. At Access, a gap in penetration was identified, with higher penetration into the adult population. This allowed for discussion of barriers to referrals with pediatric providers. This is one example of the cross-relationship between ongoing clinical work and QI efforts. They both inform the other, ideally creating a symbiotic relationship. As the team at Access does the clinical work, there is periodic check-in to see if the team continues to do what they think they are doing. Then the data is used to refine clinical flows, or to identify weak areas in the practice.

Population penetration numbers track the unique patients seen, not the number of visits, as patients may return for care several times in the time period of question. Therefore, while penetration numbers assist in describing the impact on the population, they do not reflect the whole story. A complementary metric is related to productivity or the number of visits for BHC providers. This is useful to track and identify patterns in upticks or downticks in referrals, which can in turn be communicated back to providers. For example, the team at Access encountered a period of time where the number of visits or productivity increased significantly, while penetration only increased by a few percentage points. This allowed us to identify two trends in the practice. First, there were a few patients who began coming in more often for services. And second, some providers had shifted their practice such that they expected the BHC team to see every patient who had ever seen BHC every time they were in clinic, regardless of reason for visit or how the patient was doing that day. It was found that engaging the providers in conversation about the data was very helpful as it provided a nonjudgmental way to start the discussion and begin to problem-solve issues in referral patterns, bringing referrals back in line with the PCBH model. What were the providers’ thoughts on the numbers and
had they changed their referral practices? The BHC team discussed that patients may not want to see BHC every time they are in clinic, particularly if they are presenting for an acute medical issue. Additionally, the team discussed trying to ping-pong visits between BHC and medical providers, decreasing burden on medical providers and increasing PCP availability for other patients.

The combination of metrics of population penetration and number of visits is extremely useful, and the second is likely more familiar as a measure of productivity. It is clearly useful to track productivity, although it does not capture all facets of the work such as phone calls or curbside consults. It can be useful to track the number of visits to inform staffing decisions. The team at Access also tracks the number of visits across the days of the week and across clinic sites. This has allowed for improved allocation of staffing resources. It also allows for placement of trainees in the clinics at times that will both enable them the opportunity to see the number of patients they need for training but also be the most useful for the team. This seems to be a U-shaped relationship, and the goal is to find the optimal situation where there is enough work for students to provide a strong training experience but also enough time for supervisors to allocate to training.

One way of communicating metrics has been quarterly metrics reports and annual BHC reports. The quarterly reports are used primarily for internal program evaluation purposes, allowing staff to observe the program outcomes over time and watch for any trends or changes. Population penetration, number of visits, and “hot days” are tracked to observe the overall workload on BHCs and the dissemination of the model over time. The annual report is a much larger summary and has evolved over time to include more information. It begins with reviews of the population penetration and visit numbers, and then continues with descriptions of the demographics of patients seen by the BHC team, including language spoken. PHQ-9/GAD-7 data is reviewed as a clinical outcome. Model adherence metrics are examined in line with recommendations from Robinson and Reiter (2015). The program at Access had demonstrated consistently since 2007 that about 50% of patients seen annually by BHCs are seen one time, and 85–90% are seen for four visits or less by the BHC team per year. This is recommended by Robinson and Reiter for model fidelity regarding population-based care. Additionally, the BHC team at Access has observed that the small percentage of patients seen the most often has varied in terms of the maximum number of visits, but remains a small portion of patients overall. This is an example of the flexibility of the model meeting the needs of the population. Additional exploration of this “high utilizing” subset of BHC patients has shown that this group represents patients with more severe and persistent mental health issues, and more comorbidities of substance abuse and/or behavioral health issues (Fondow et al., 2015).

## Dissemination

The integration of the PCBH model at Access has impacted the larger healthcare community in Dane County, Wisconsin. Four healthcare systems locally have chosen to begin integration of behavioral health care using the PCBH model since Access was able to demonstrate the effectiveness of the work. Group Health Cooperative (GHC), the University of Wisconsin (UW) Pediatrics outpatient clinic, Meritor Hospitals and Clinics, and the University Health Services (UHS) for University of Wisconsin–Madison have all consulted with Access on how to begin an integrated program.

In the case of UHS, they have chosen to take portions of the PCBH model and adapt it to their system to meet their unique needs as a provider of student services, particularly related to accessing services and brief interventions for students on campus. The UW Pediatrics PCBH program was started by a graduate of the postdoctoral training program at Access, and has been well received by the providers. This pilot has also been part of a larger effort across the UW system to integrate care.
The PCBH program at GHC began with consultation from one of the Access BHC team members, who worked on-site at GHC to assist in starting the program and eventually left Access to lead the new program at GHC. He worked with other mental health professionals who wished to cross-train, and also with the administration on what the implementation would entail. Another postdoctoral fellow from Access also joined the GHC team. The program has been well received and continues to grow across the multiple GHC clinic locations in Madison. The Meritor hospital/clinic system is in the nascent stages of implementation as of the writing of this text and has consulted with Access as it plans its implementation.

Taken together, this spread of the PCBH model in the county demonstrates the impact that a local clinic can have in becoming a proof-of-concept for area health organizations to invest in integrated care in just under a decade. The data collected as well as the development of a reputation for openness to assist has been key in disseminating the model. For Access, dissemination has been an intentional goal given that a healthcare environment with integrated care as a standard of care is more likely to be supportive of a sustainable financial model for it. Along these lines Access has also participated in state-wide meetings on integrated care including a recent year-long State Health Innovation Project (SHIP) funded by the federal government.

THE FUTURE OF PRIMARY CARE BEHAVIORAL HEALTH AT FEDERALLY QUALIFIED HEALTH CENTERS

The future of primary care behavioral health within community health centers is ripe with ongoing opportunities to promote the health and wellness of the population. The evidence is compelling regarding the benefits of integrated care for patients, providers, and the healthcare system as a whole. However, one must engage in continuous process improvement to identify areas for efficiency while examining the populations we serve to ensure that we are asking the right questions and have pathways developed to meet patient needs as they arise.

The next phase to produce greater impact for population-based care at Access is developing sophisticated population stratification. How do we manage care across clinic staff and varying physical and/or behavioral health conditions? Currently the Access behavioral health team engages in disease-specific care management for depression/anxiety, pediatric ADHD, general pediatrics, and consulting psychiatry. The BHC completes chart reviews and outreach as appropriate for patients with these identified conditions. However, the goal in the future is to offer more enhanced care management by not solely focusing on one disease-specific registry/list but tying together the comorbidities in coordination with the broader healthcare team. This includes involving nurses, support staff (MA/LPN), diabetes educators, registered dieticians, and so forth, working in concert with the BHC to support whole-person health. In addition, the aim is to increase focus on tracking health outcomes as they relate to comorbidities of physical and behavioral health issues (e.g., depression and diabetes, cardiovascular health).

Access is currently engaged in universally surveying its adult patient population for adverse childhood experiences (ACE). We are keenly aware of the impacts of childhood trauma on health conditions later in life. This screening has allowed the opportunity to acknowledge this link directly during patient contacts as well as better inform the healthcare team regarding the impacts of trauma on health. This will be a means of better understanding the population and guiding our future efforts to direct population-based care strategies to the most vulnerable of Access patients.

Community health centers continue to see more and more individuals with complex medical and psychiatric conditions. The reasons for this increase is multifactorial, as there are many barriers to
the patient population served by community health centers in accessing needed behavioral health-care services. Barriers include being uninsured or underinsured, socioeconomic status, transportation issues, stigma, the severity of the illness making navigating traditional mental health resources too challenging, and potentially patient preference with desire to receive care in an environment they find familiar and comfortable. By integrating behavioral health in primary care, health centers are well positioned to screen and detect individuals who would benefit from behavioral health intervention and thus can capitalize on patient motivation in the moment by having a same day BHC visit.

Many systems are fragmented, leading to challenges navigating appropriate access points to receive needed services. There are also distinct silos between medical, mental health, and substance abuse systems. Further fragmenting care is the lack of consistent electronic health records or lack interoperability of systems involving the records of primary care, specialty mental health, and substance abuse services. Thus, having integrated behavioral health and documenting in a shared electronic health record promotes development of comprehensive, patient-centered plans of care. Increasingly individuals with severe mental health issues and substance use disorders are presenting to community health centers partially impacted by siloed substance abuse and mental health specialty services in the community and uncertainty as to how to navigate the system to access needed level of care. While primary care cannot meet all the needs of these individuals, much can be done in the primary care setting to prepare these individuals to navigate the system of care. Examples include addressing motivation and readiness for change and supporting treatment readiness via a motivational enhancement. BHCs can also educate patients on the process for accessing specialty treatment to support successful referral into the community when appropriate. Connecting these silos meaningfully is a necessary element moving forward for the clinic.

Access has successfully used the expertise of a consulting psychiatrist to educate the primary care team regarding psychotropic medication management to help support the needs of its patient population. However, the lack of psychiatrists who take Medicaid and Medicare in the community is a persistent challenge for Access. It hopes to continue to leverage the expertise of the psychiatrist by using her as an educator to share knowledge with the care team to increase Access’s reach into the population. However, quality management of psychotropic prescribing remains a key issue to address systematically. As clinics like Access take on more prescribing responsibility, more sophisticated monitoring systems will be necessary to ensure that patients are receiving standard and safe care. The use of registries is likely to play a bigger role in this respect.

Substance use disorders are commonly seen within primary care, but access to appropriate assessment and treatment is challenging. Access recently received one of hundreds of Human Resources and Services Administration (HRSA) substance abuse expansion grants to better support community health centers in addressing substance abuse issues within the medical home. Specifically this grant is to expand access to medication-assisted treatment (MAT). Community health centers are continually stretching themselves to meet the ever-increasing complex needs of our patient population. However, efforts to provide comprehensive whole person health needs to be intentional and well planned to support development and sustainability of such endeavors. Specifically, Access is faced with the challenge of how a PCBH service responds appropriately to the needs of the addicted population while maintaining fidelity to the model and the staff’s role as generalist practitioners. In an effort to have greater penetration into the population of individuals with substance use disorders or those at risk for development of a substance use disorder we will begin to universally screen for substance use/misuse using the Screening, Brief Intervention, and Referral to Treatment (SBIRT) protocol. The well-established primary care behavioral health team at Access offers the ability to support patients who are identified through the universal screening by providing motivational interviewing, targeted interventions, care plan development, and community referrals as appropriate.
Integrated primary care behavioral health supports addressing both physical and behavioral health conditions with the understanding that primary care is the most common point of entry into care. In addition, the primary care practitioners are in need of increased collaboration to adequately meet the needs of the patients under their care. When considering whole-person health, the issue of chronic pain comes to the forefront as improving outcomes in chronic pain management and increased awareness of judicious opioid prescribing to reduce the impact of inappropriate prescribing and abuse of opioids is both of national and local concern. One of the main goals of Access's work in this area is to collaborate with its primary care colleagues to develop standardized, safe, and effective means to address chronic pain by offering a multidisciplinary approach with an emphasis on patient engagement, functional gains, and support of self-management strategies. The behavioral health team has been an integral part of the development of interdisciplinary team meetings in which care team members are able to review patients with complex diagnoses to determine best next steps in regards to a plan of care. Both patients and providers benefit from a comprehensive team-based approach to care. This too is an evolving area as the clinic develops its boundaries and communicates these to the community and to its patients.

In sum, community health centers with integrated behavioral health care are well positioned to tackle the complex needs of the patients served, but will need to ensure open communication among team members, among community partners, and with patients to ensure understanding of the scope of services provided. And as has been the case thus far in the early history of the Access clinic, the innovations at a community health center can translate to other health systems as well.

REFERENCES

Index

Note: Page numbers in *italics* denote references to Figures and Tables.

3-P Model of Insomnia 470, 475
7-Minute Screen (7MS) 378
36-item Short Form-36v2 Health Survey see SF-36v2 Health Survey
Aatre, G. 37
Abashian, S.W. 615
Abbreviated Mental Test (AMT) 379, 381
Abbreviated multidimensional measures 34–5
AB Cognitive Screen (ABCS) 378–9
Abeloff, M.D. 616, 622
absenteeism 24–5
ACA see Patient Protection and Affordable Care Act (ACA)
Acceptance and Commitment Therapy (ACT) 107, 483, 673–4, 676–81
Access Community Health Centers: addressing substance abuse issues 781; adverse childhood experiences and 780; behavioral health consultation 765–6; care management service 767–8; care manager 767–8; clinical outcomes 776–7; combatting burnout 769; communal approach to patient care 773–4; consulting psychiatry 766–7; core program components 764–70; development of 763–4; engaging primary care providers and administration 771; as health promotions clinic 768–9, 772; “hive” mentality of 773–4; long-term financial vision of 774–5; managing special populations 772–3; Medicaid and 781; medications use 766–7; metrics reports 779; patient engagement 777; population penetration 776, 778; process outcomes 777–8; productivity data 776; programmatic management 772; program outcomes 778–9; protocol-based screening 765–6; quality improvement practices 775; referrals 777, 778; staffing levels 763; substance abuse screening 768–9; success of 770–5; sustainable workforce development 771–2; tracking and measuring outcomes 775–6; trainee selection 770; training program 769–70; vision alignment with mission 770
accountable care organization (ACO) 23, 197, 667
accluturation-related stressors 739
Actigraphy 480
actionable information 146
acute inpatient psychiatric care, need for 61
acute insomnia 473
acute myocardial infarction 617
acute pain 671
acute PTSD 430; see also posttraumatic stress disorder (PTSD)
acute risk 298
acute stress disorder (ASD) 429, 433
Adamis, D. 378
Addenbrooke's Cognitive Examination (ACE-R) 379
addictionology 768
Addiction Severity Index (ASI) 401
ADHD child/adolescent 86
adolescents: alcohol misuse in 391; alcohol screening for 394, 397–8; asthma control test (ACT) 576; asthma control test for 581–2; AUDIT tools 397; CRAFFT screening for 397–8; delayed sleep phase disorder (DSPD) 474; depression in 258; drug abuse screening for 397; illicit drug misuse among 392; integrated care for 739–40; Primary Care Psychology Training Network (PCPTN) 733; suicide and 184
Adolescent Substance Abuse Goal Commitment (ASAGC) 404
adrenocorticotropic hormone (ACTH) release 357
Adult Self Report ADHD Scale 376
Adverse Childhood Events (ACE) Study 766
adverse childhood experiences (ACE) 433–4, 672
Ae Lee, M. 612
affective disorders, prevalence rates of 599
Affordable Care Act see Patient Protection and Affordable Care Act (ACA)
African Americans: depression and 248; eating-related problems 501; illicit drug use 392; PCPTN focus on 736; posttraumatic stress disorder 451, 452; stress and 363; substance abuse screening 396; suicide attempts 183; weekly stress inventory and 362
Agency for Healthcare Research and Quality (AHRQ) 66, 69, 199–200, 389–90
aging process 281
agoraphobia 327, 620
Ahmedani, B.K. 14
AHRQ see Agency for Healthcare Research and Quality (AHRQ)
Alcohol, Smoking, and Substance Involvement Test (ASSIST) 397
Alcohol-Related Problems Survey 394
alcohol use disorders (AUD) 109–10, 718
Alcohol Use Disorders Identification Test see AUDIT
(Alcohol Use Disorders Identification Test)
Alcohol Use Disorders Identification Test (AUDIT) 768
allergic rhinitis 589, 590
Allison, T.G. 618
Almirall, J. 530
Alonso, M.M. 144
alprazolam 620
Alvarez, W. 613
Alzheimer’s disease 185
Alzheimer’s Disease 8 (AD8) 379
AMA Code of Medical Ethics 101
American Academy of Family Physicians National Research Network 391
American Academy of Neurology 372
American Academy of Pediatrics 397
American Academy of Sleep Medicine (AASM) 470, 471
American College of Physicians 471
American Medical Association Code of Ethics 98
American Medical News 167
American Psychological Association (APA) 41, 673, 741: Assessment of Competency Benchmarks Work Group 95; Guidelines for Psychological Practice in Health Care Delivery Systems 25; neuropsychological tests 35; outcomes assessment importance 201; primary care psychology training 108, 112; psychology competencies 96; Society for Health Psychology 112; on suicide 14
American Psychologist 32
Amick, H.R. 391
Anda, R.F. 434
Anders, P. 10
Anders, S.M. 171
Andrescu, C. 612
Andrews, Gavin 154–5
anemia 589, 590
Angst, J. 611
anhedonia 253
anosmia nervosa (AN) 500–1
Anstine, D. 516
antidepressants 12, 283, 620, 777
anti-psychotic medications 60
Antony, M.M. 144
anxiety: case study 346–7; chronic physical conditions and 344–5; decision-making guidelines 327–34; depression comorbidity with 39, 167; determining treatment intervention 333; eating disorders and 509; failure to detect 169; in older adults 326; pharmaceutical cause of 345; sample clinical questions 327; SCL-90-R/BSI and 611–12; self-report measures 328–32; substance withdrawal 345; symptoms 344–5
anxiety disorders: BHL program 718; defined 614; prevalence rates of 319, 599; PTSD and 447; research 754; suicidal behavior and 614; see also generalized anxiety disorder (GAD-7)
Anxiety Scale of the Four-Dimensional Symptoms Questionnaire (4DSQ) 328
Anxiety Sensitivity Index-3 (ASI-3) 328
ADA Code of Ethics 98
APA Ethical Principles and Code of Conduct 100–2
APA Guidelines for Psychological Practice in Health Care Delivery Systems 99–100
Appelbaum, S.A. 39–40
application cluster 104–5
Aragon, M. 640
Arana, G.W. 177
Arata, C.M. 613
Arbisi, P.A. 639, 644, 651
Arcelus, J. 616
Archer, R.P. 147, 152
Arend, P.A. 13, 19
area T-scores 609
area under the curve (AUC) 180, 546
Arizona State University 113
Arnesen, H. 617
arousal 357
art of medicine 63
Ashbury, F.D. 17
Ashton, K. 638, 640, 646, 647
Asian Americans: anxiety and 340; depression and 248; eating-related problems 501; posttraumatic stress disorder 452; suicide risk screening 183
Asmundson, G.J.G. 336
asQ’em methodology 184
assessment see psychological assessment
assessment-based feedback 41
assessment modality, self-report vs. clinician rating 175–6
Assessment of Competency Benchmarks Work Group 95
assessing technology: apps 727; availability/accessibility of 160–1; clinical utility 162–3; computerized adaptive testing capabilities 157; considerations for 159–63; costs of 160; disadvantages of 157; disk-distributed software 156–7; electronic health record systems and 161; faxback 157–8; flexibility 162; implementation of 161; Internet-based 156–7; IVR 158–9; limited utility of 161–2; mail-out/mail-back process 156; paper-and-pencil test forms 156; patient affordability 161; psychometric integrity 160; selection of 156; smartphone abilities 157; telehealth 155–6; testing modalities 159
ASSIST (Alcohol, Smoking, and Substance Involvement Test) 394, 397
Association for Behavioral Health and Wellness 17
Association of Psychology Postdoctoral and Internship Center (APPIC) 112
asthma control: achieving 574–5; Guidelines for the Diagnosis and Management of Asthma 574–5
Asthma Control Questionnaire (ACQ) 578–9
asthma Control Test (ACT): administration of 576–7; for adolescents 581–2; applications of 583–4; brevity of 583–4; clinical validation study 575–6; convergent validity 579; cut-point scores in screening for uncontrolled asthma 578, 580; development of 575–6; discriminant validity 579, 579; electronic medical records (EMRs) 576; internal consistency reliability of 579; MCID of 582–3; psychometric properties 577–82; questionnaire 578–9; responsiveness of 580; risk prediction of 583; scores 577–9; scores across groups 578; scoring of 576; validity of 581
Asthma Impact Survey (AIS) 36, 555
attention (cognitive functioning) 373
attention deficit hyperactivity disorder (ADHD) 375–6, 638
Attikison, C.A. 12–13
AUDIT (Alcohol Use Disorders Identification Test) 37, 85, 109, 182, 391, 393, 395–6
AUDIT-3 395
AUDIT-C 395–6, 720
AUDIT-QF 395
Auxier, A.M. 23, 38
avoidant/restrictive food intake disorder (ARFID) 509
axonal sprouting 688
Baer, B.A. 603
Bagby, R.M. 637–8, 640
Baider, L. 616
Baird, M.A. 102–3
Baker, J. 15
Baldessarini, R.J. 177
Ballenger, J.C. 620
Bamsley, J. 17
Barbaglia, G. 255
Barefoot, J.C. 641
bariatric surgery 638, 646–8, 651–6
Barnett, M.D. 15
Barrett, J.E. 180, 599
Bartlett, D. 620
baseline data 41
baseline risk 297–8
base rates: defined 173; problem of low 176–8; sequential screening and 177–8
Baum, A. 359
Bayliss, M.S. 554
Beard, C. 263, 322, 324, 326
Beck, A.T. 245, 324
Beck Anxiety Inventory (BAI) 35, 148, 321, 323, 328
Beck Anxiety Inventory-Primary Care (BAI-PC) 328, 343–4
Beck Depression Inventory (BDI) 245–6
Beck Depression Inventory-Amend (BDI-IA) 246
Beck Depression Inventory-FastScreen (BDI-FS): consistency with DSM-5 criteria 250–1; consistency with ICD-10-CM criteria 251; described 247–8; psychometric properties 249
Beck Depression Inventory for Primary Care (BDI-PC) 246–7
Beck Depression Inventory-Second Edition (BDI-II): case study 270; consistency with DSM-5 criteria 250–1, 250; consistency with ICD-10-CM criteria 251; described 247; implementation issues 268–70; norms for sex and educational level 248; for pain assessment 419; psychometric properties 248–9; reading level of 148; suicide risk screening 184
Beck Depression Inventory-Short Form (BDI-SF) 246
Beck Hopelessness Scale (BHS) 184
Bedroussian, A. 9
Beers, Clifford 59
Behar, E. 148, 149, 441
behavioral activation 107
behavioral assessment protocols 420
Behavioral/Externalizing Dysfunction (BXD) 636, 640
behavioral health care, defined 57
behavioral healthcare assessment 33–6; see also psychological assessment
behavioral healthcare services: direct spending costs for 4–5; integrating in primary care settings 3
behavioral health consultant core competency tool 124–42
Behavioral Health Consultant model 219
behavioral health disorders: chronic diseases and 9–11; inadequate detection and treatment of 13–16; prevalence in US 6–7; prevalence of 5–8; prevalence rates in primary care settings 7–8
behavioral health in primary care setting 83–5
behavioral health interventions 87, 333
Behavioral Health Laboratory (BHL) 710–11, 717–19, 717–18
Behavioral Health Measure-20 (BHM-20) 305–8
behavioral health personnel (BHPs): levels of care 196–7; outcome metrics 207; process and operational variables 206–7; pursuing Quadruple Aim 206–7; Quadruple Aim relevance to 194
behavioral health practitioner, reluctance to see 12–13
behavioral health professionals, underutilized 78
behavioral health providers: defined 195; treatment philosophies 28
behavioral health services: co-payments for 16; integrating primary care 13
behaviorally induced insufficient sleep syndrome (BISS) 474
behavioral scientist, residency training for 84
behavioral sleep medicine: case studies 487–93; classical and operant conditioning 469–70; diathesis-stress model 470–1; integration with primary care 471–2; two-factor model 469; see also sleep disorders
behavioral sleep medicine (BSM) 468
Behavior and Symptom Identification Scale-24 (BASIS-24) 257
behavior change, patient commitment to 699
behavior change referrals 221–2
Bell, J.A. 547
Benbadis, S.R. 649
Benjamin, A.H. 608
Bennett, R.T. 615
Bennett, S.E. 615
Ben-Porath, Y.S. 417, 637–42, 644, 647–8
Benson, J. 4
benzodiazepines 468
Berger, M.L. 530
Berlin Questionnaire 480
Bernstein, I.H. 534–5
Berriers, G.E. 379
Berry, D.T. 638, 649
Berwick, D.M. 666
Bhattacharai, R. 537
biases 201, 742
binge eating disorder (BED) 500–6, 512
biomedical model of healthcare 3
biomedical model of illness 15
biopsychosocial model 63, 66, 412
biopsychosocial model of pain 411–12, 424, 650
Bishop, S. 414
Björgvinsson, T. 257, 263, 322, 324, 326
Blanchard, E.B. 444
Bland, R. 67
Blessed Information-Memory-Concentration Scale 381
Bliese, P.D. 435, 439, 440
Block, A.R. 420, 648
blood glucose awareness training 509
Blount, A. 12, 19, 65, 69
Blount, F.A. 741
Blue Cross Blue Shield of Michigan 24
Bly, Nelly 59
Björntorp, P. 358–9
Blanchard, E.B. 444
Bland, R. 67
Blessed Information-Memory-Concentration Scale 381
Bliese, P.D. 435, 439, 440
Block, A.R. 420, 648
blood glucose awareness training 509
Blount, A. 12, 19, 65, 69
Blount, F.A. 741
Blue Cross Blue Shield of Michigan 24
Bly, Nelly 59
Board of Educational Affairs 104–5
Bobula, J.A. 171
Boccaccini, M.T. 150
Bodenheimer, T. 666
Bodily Pain (BP) domain scale 533
body scan 698
Bohachick, P. 621
Boleloucky, Z. 606
Bomyea, J. 614
Bongers, P.M. 650
Boon, H.S. 17
borderline personality disorder 416, 447–8
Boone, D.J. 611
Borkovec, T.D. 441
Borson, S. 381
Boulet, V. 622
Bourgatsos, C. 392
Bourque, L.B. 451
Bourier, L.M. 642
Boyle, S.H. 641
Brandt, C.P. 336
Brantley, P.J. 358, 360, 361
Brawman-Mintzer, O. 447
Breckenridge, J. 608
Brems, C. 614, 615
Breting, L.M. 31
Bridges, A.J. 106
Bull’s Eye Plan 692–3, 694, 695, 698
Burnam, M.A. 10, 529
burnout, clinical staff 666, 769
Butcher, J.N. 632
Butler, M. 29
Buysse, D.J. 477
CAGE (Cut down, Annoyed, Guilt, and Eye-opener) screen 182, 394, 396
Cahn, T.S. 608
Calhoun, P.S. 440
CALM technique 692
Cameron, O.G. 612
Campbell, D.G. 448
Campbell-Sills, L. 337
cancer 589, 590, 622; see also oncology
Cannon, William 356
Cape, J. 751
Capital Solution Design 717
CAPS-5 443–6
CAPS-Diagnostic version (CAPS-DX) 443
CAPS-Symptom Status version (CAPS-SX) 443
cardiology 617
care delivery models 76
care management assessment protocols 722
care manager 676
Carey, K.B. 614
Carey, M.P. 614
Carleton, R.N. 255
Carrington, P. 621
Carr, N.W. 537
Carter, Jimmy 61
Caruso, L.S. 470
Carver, J.R. 158
Cathal, A. 4
Cattell, James McKeen 81
causal suicide risk factor 300
Cella, D. 43, 531
Center for Epidemiological Studies-Depression Scale (CES-D): consistency with DSM-5 criteria 258–9, 259; consistency with ICD-10-CM criteria 259, 260; described 251–2; diagnostic accuracy 171; implementation issues 267–70; primary components 252; purpose of 251; reading level of 148; revised 252–3; scale 252; short form 257–8
Center for Epidemiological Studies-Depression Scale-Revised (CES-D-R): consistency with DSM-5 criteria 258–9, 259; consistency with ICD-10-CM criteria 259, 260; described 251–2; diagnostic accuracy 171; implementation issues 267–70; primary components 252; purpose of 251; reading level of 148; revised 252–3; scale 252; short form 257–8
Center for Epidemiological Studies-Depression Scale-Revised 10-item for Adolescents (CESD-R) 257–8
Center for Integrated Health Solutions 105–6
Center of Excellence for Integrated Care 770
Centers for Disease Control and Prevention (CDC): on chronic diseases 9, 10, 554; Healthy Days Core Measure 681, 682; on insufficient sleep 470; on
clinical interview: case example 237–42; communication with primary care provider 237; data collection during 33–4; defined 213; documenting 234; family psychiatric history 231; goals of 213–14; history elements 217; importance of 213; informed consent within 222–4; in medical settings 216–18; in mental health settings 214–16; nonstructured format of 215–16; opening of 229; orienting patient to integrated behavioral health 229–31; past medical history 231; in primary care settings 218–22; rapport building in 214–15; semi-structured 282; social history 231–2; stages of 215; standardized intake guide 232–4; strategies to promote patient responses 680; time constraints 224–5; see also psychological assessment; screening clinical outcomes: studies 753–5; tracking 42 clinical rating scale 175–6 clinical skills 97 clinic huddle 86 Clinician-Administered PTSD Scale (CAPS) 442–6 clinician ratings 373 clinician recognition: aided recognition in 170–1; unaided 169–70; see also screening clinicians: interactions with patients cognitive screening: 7-Minute Screen 378; Short and Sweet Screening Test 381; unaided 170–1; see also screening cognitive impairment: defined 371; dementia and attention deficit hyperactivity disorder 375–6; for persistent pain (PP) 672 cognitive functioning: age-related declines in 371; comorbid depression and 10; financial burden of 9; managing 221–2; types of 9 chronic headaches 673 chronic illness 358 chronic insomnia 468, 473 chronic low back pain 650–1, 673, 674 chronic musculoskeletal pain 419 chronic opioid therapy (COT) 422 chronic pain: BHL program 718; catastrophizing 413–14; catastrophic factors impacting 413–14; improving quality of care with 666–70; quality 650–1; mental health disorders and 415–16; personality disorders and 415–16; prevalence of 584; sexual abuse and 615–16; symptoms of depression 685; see also pain; persistent pain (PP) 672 Chronic Pain-CAT 585 chronic pelvic pain 615–16 chronic stress exposure 507–8 Chudzinski, V. 326 Ciarcia, J.A. 153, 154, 287 Cibula, J.E. 649 Ciechanowski, P.C. 12 circadian process 469, 478, 481 CIT-i Coach 485 Civilian Mississippi Scale (CMS) 451 Civilization Syndrome 358–9, 358 Clark, L.A. 612 classical and operant conditioning 469–70 classic test theory (CTT) 479 Cleary, P.A. 606 Cleary, P.D. 171 clinical barriers to integrated care 28 clinical consultation 105 clinical decision making activities 30 clinical depression see depression Clinical Evaluation Guide (CEG) 517 clinical health psychology 107–9 Clinical Health Psychology Competencies Rating Form 108 opioid pain medication overdose 675; on suicidal behavior 304–5 Centers for Medicare and Medicaid Services (CMS) 4, 23, 68, 724 central nervous system 688, 691 central sensitization to pain 667–9, 671–2, 688 certificate programs 114 Chan, K.T. 554 Chandler, M.J. 378 Chang, G. 182 “Chapters in a book” exercise 700 charting: documentation skills; “Chapters in a book” exercise 700; importance of 226–7; in primary care settings 218–22; process in 383–5; rapid documentation 224–5; see also electronic medical records Chaytor, N. 649 chemical dependency/substance abuse care 57 Chemical Use, Abuse and Dependence Scale (CUAD) 399, 401 Chen, J.K. 376 Cheng, S.K. 484 Cherkin, D.C. 674 childhood trauma 434; see also trauma Chiles, J.A. 25, 182, 183, 608 Chin, W.Y. 169, 554 chlorpromazine 60 Choquette, K.A. 649 chronicle diseases: behavioral health disorders and 9–11; comorbid depression and 10; financial burden of 9; managing 221–2; types of 9 chronic processes 469, 478, 481 Clinician-Administered PTSD Scale (CAPS) 442–6 clinician ratings 373 clinician recognition: aided recognition in 170–1; unaided 169–70; see also screening clinical outcomes: studies 753–5; tracking 42 clinical rating scale 175–6 clinical skills 97 clinic huddle 86 Clinician-Administered PTSD Scale (CAPS) 442–6 clinician ratings 373 clinician recognition: aided recognition in 170–1; unaided 169–70; see also screening clinic personnel: improving work life of 669–70; as source of information 226–7 Clock Drawing Test 370, 377, 378, 381 Coady, E.L. 649 Coffey, P. 615 Cognitive Abilities Screening Instrument (CASI) 379 Cognitive Assessment Screening Test (CAST) 379–80 cognitive behavioral therapy (CBT): for avoidant/ restrictive food intake disorder 509; for eating-related problems 509; for geriatric depression 283–4; for insomnia 468, 481–6, 485; for persistent pain 673 cognitive disorders screening 372 cognitive functioning: age-related declines in 371–2; attention and 373; domains of 373–4; memory and 374; psychiatric-related decline in 372 cognitive fusion 676 cognitive impairment: defined 371; dementia and 280, 376–82; detection of 383; memory and 372–3; screening 184–6; screening concepts 369–70; screening for 383–5; screening vs. assessment 369 cognitive functioning: 7-Minute Screen 378; Abbreviated Mental Test 379, 381; AB Cognitive Screen 378–9; Addenbrooke’s Cognitive Examination 379; Alzheimer’s Disease 8 379; attention deficit hyperactivity disorder 375–6; behavioral observation of 375; Brief Alzheimer Screen 379; Clock Drawing Test 370, 377, 378; Cognitive Abilities Screening Instrument 379; Cognitive Assessment Screening Test 379–80; DemTect 380; Folstein Mini-Mental State Exam (MMSE) 377; General Practitioner Assessment of Cognition 380; in geriatric populations 185–6; Informant Questionnaire on Cognitive Decline in the Elderly 380; instrument selection for 374; Mini-Cog 381; Montreal Cognitive Assessment 382; Post Concussive Symptom Scale (PCCSS) 376; in primary care settings 185; process in 383–5; Rapid Dementia Screening Test 381; Short and Sweet Screening Instrument 380; Short Memory Questionnaire
cutoff values: for BHM-20 306–7; for GAD-7 325–6
Czyz, E.K. 311

Daily Assessment of Symptoms-Anxiety (DAS-A) 329
Daily Planet health center 738
Daily Stress Inventory (DSI) 355, 360–1, 363–4
Dammen, T. 617
Damschroder, L.J. 722, 723–4
DAST (Drug Abuse Screening Test) 394, 397, 768
DAST-10 397
DAST-A 397

data collection 224–5
Davidson, J.R.T. 613
Davis, R. 517
daylight functioning 476–7, 487, 491
Dear, B.F. 323, 335
de Bont, P. 450
deep breathing techniques 109, 748
DeGood, D.E. 413
DeGruy, F. 63
delirium 185, 378
DellaPietra, L. 600
demence, A. 604
Dementia: aging vs 185–6, 280–1; assessment instrument 86; BHL program 718; general cognitive impairment and 376–82; MMPI-2-RF and 642; MMPI-2 with 642; prevalence rates for 185; screening for 280
Dementia of Alzheimer's Type Inventory (DAT) 185
demoralization 417, 633
demoralization Scale 417
DemTect 380
De Nitto, S. 640
Department of Defense (DoD) 105
depression: age effects of 174; assessment for 86; BAI-PC assessment for 346; BHL program 717; chronic diseases and 10; cognitive impairment and 372; comorbid with anxiety 39; comorbid with diabetes 24; effective management of 187; late-life 281; panic with comorbid 611–12; phenomenological characteristics of 174; PIQ-6 scores 588, 589; pregnant women with 14; prevalence rates of 245; PTSD and 448–9; SCL-90-R/BSI and 611–12; scores 753; screening for 38–9, 167, 546; somatic symptoms of 277; symptoms 10; treatment of 12–13; work productivity and 24–5; see also Beck Depression Inventory (BDI); Geriatric Depression Scale (GDS)

Depression Care Managers (DCM) 719–20
Derogatis, L.R. 180–1, 184, 599–600, 606–8, 616, 622
Derogatis, M.F. 184, 607
Derogatis' family of symptom checklists 34–5
Derogatis Psychiatric Rating Scale (DPRS) 601, 622
DeRosa, M.A. 538
detailed inquiry 215
Devinsky, O. 649
DeVol, R. 9
Dewitt, K.N. 620–1
dexamethasone suppression test (DST) 177–8, 611
diabetes: comorbid with major depression 24; eating-related problems and 509; hypoglycemia fears 504, 508–9, 519; PIQ-6 scores 588, 590; SCL-90-R/BSI and 616; stress and 358

diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) 5
Diagnosis and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) 250, 250, 251
diagnostic efficiency 147, 249, 254–7, 262–5
diagnostic identification 170–1
Diagnostic Interview Schedule 255, 282
diaphragmatic breathing 483
diathesis-stress model 470–1
diazepam 620
dickey, B. 43, 529
Dickinson, M. 24–5
differential item functioning (DIF) 336
diffuse patients 218
diffusion of innovation studies 726
direct observation 114
Directory of Training Programs 108
disability, long-term 59

disability profile 411
discriminant validity 323–4, 579, 579; see also validity, screening
disease impact: diagnosis vs. measuring 566–7; measuring symptoms 566; tests 566; see also health-related quality of life (HRQOL)
disease impact measures: Asthma Control Test 574–84; defined 36; dimensionality of 567; headache impact test 569–74; pain impact questionnaire 582–92
disease model 565–6
disordered eating 500–6
disordered eating treatments: Eating Attitudes Test 515–17; Eating Disorder Examination-Questionnaire 510–12; Primary Care Evaluation of Mental Disorders (PRIME-MD) 517–19; Questionnaire on Eating and Weight Patterns-Revised 512–14, 513
disorder-specific measures 35

disorganized attachment style 416
distress and risk assessment method (DRAM) 342
distress states 599–600

DITomasso, R.A. 27
Dix, Dorothea 58
Dizon, J. 484
Dobmeyer, A.C. 110
Doctoral Program in Clinical Psychology 106
Doctoral Psychology Internship Program 106
Doctor of Behavioral Health (DBH) 113
doctor-patient relationship 77, 81, 213
doctors see primary care providers
documentation: of clinical interview, assessment and plan 234; in electronic medical records 405; template 235–6

Dodick, D. 569
Doherty, W.J. 20
Dokmak, A. 381
Donabedian, A. 42

DRAM mZDI 342
Drerup, L.C. 641–2, 650
Drossman, D.A. 618
Drug Abuse Screening Test see DAST (Drug Abuse Screening Test)
drug misuse/dependence 392; see also Drug Abuse Screening Test
drugs: assessment instrument 86; illicit drug use 5; outcomes measures 620

drug use disorder 110
Drum, D.J. 18
Druss, B. 14
dualism 15
dual relationships 98, 101
Dubois, M.F. 530
Dudas, R.B. 379
Dysfunctional Beliefs and Attitudes About Sleep
16-Items Scale (DBAS-16) 478
dysthymic disorder 600
Earleywine, M. 256
Eating Attitudes Test (EAT) 502, 515–17
eating behavior: assessment tools 509; assessment
within primary care setting 499–500;
biopsychosocial influences on 499; case study
520–3; screening tools 502–5; see also disordered
eating treatments; eating-related problems
Eating Disorder Examination-Questionnaire (EDE-Q)
503, 510–12, 511
eating disorder treatment 501–6
eating-related problems: case study 520–3; chronic
stress 507–8; compassion-focused therapy for
674; disordered eating 500–6; emotional eating
507; insufficient sleep 508; low health literacy
506–7; obesity 500; screening tools 502–5, 647;
socioeconomic status 507–8; treatment for 501–6;
see also disordered eating treatments
Eaton, W.W. 252, 258
Edwards, D.W. 153
Eisenstat, R. 378
elderly see geriatric populations
electronic health record (EHR) 195, 195
electronic health record software 29, 161
electronic medical records (EMRs) 195, 195, 237, 404,
474, 576, 727, 781
emotional distress in cancer patients 616–17
emotional eating 505, 507
Emotional/Internalizing Dysfunction (EID) 635–6, 639, 647
emotional life 63
emotion-focused coping 357
Engel, George 63, 81, 217–18
Engel, J. 649
enhanced care access 64
environmental stressors 357
Epidemiologic Catchment Area (ECA) Study 167, 599
epidemiologic screening model 172–3, 173
epilepsy patients 649
Epstein, R.M. 95
Epworth Sleepiness Scale (ESS) 477
ethics: APA guidelines 100–2; resolving dilemmas 102;
students’ perspectives on 98
ethnoviolence 452–3
European Americans 451–2
Everly, G.S. 359
evidence-based care 64
evidence-based interventions 282
executive skills 374
experiential avoidance 676
eye movement desensitization and reprocessing
(EMDR) therapy 450
factorial validity 324–5
Fairburn, C.G. 510, 606–7
Fakhoury, T.A. 649
false negatives, screening 38–9, 173
false positives, screening 38–9, 177, 440–1
family medicine 84
family physicians 98
family psychiatric history 231
Fan Free Clinic 740
faxback technology 157–8, 159
Fayers, P.M. 531
Fear and Distress model of internalizing
psychopathology 637–8
fear avoidance 414–15, 419–20, 504
Fear Avoidance Beliefs Questionnaire (FABQ) 414
Fear Avoidance Components Scale (FACS) 415,
419–20
Fear Questionnaire (FQ) 329
Federally Qualified Health Centers (FQHC) 747, 763;
see also Access Community Health Centers
federally qualified health centers (FQHC) 23, 197
fee-for-service 69, 435–6, 774–5
Feliciano, L. 377
Ferguson, R.J. 343
Fernandez, E. 650
Fernandez, K. 150
Feussner, J.R. 171
fibromyalgia 673
Ficken, J. 151
fidelity, assessing 114
fidelity to service delivery standards 114–15
Fields, S.A. 106
Fiellin, D.A. 182
financial barriers to integrated care 29, 69
Finch, A.E. 41
Finklestein, S. 177
Finn, S.E. 146, 549
Fischer, J. 146, 149, 163
Fisher, D.G. 614, 615
Five-Level Integration Model 20
Five Shot and Rapid Alcohol Problems Screen (RAPS)
396
Flecker, J.M. 280
Fleming, M.F. 182, 650
Flesh Reading Ease formula 148
flexibility 162, 692
floating profile 411
fluid vulnerability theory (FVT) 297–9, 301, 308
Flynn, L.M. 550
Focused Acceptance and Commitment Therapy
676–81
foyer, L.J. 644
FORCAST formula 148
Ford, D.E. 182
Fordyce, Wilbert 420
Forero, C.G. 255
form 90 399, 401–2
Fortin, M. 530
foundational competencies 103
Foundations for Flourishing Class 701, 702
Four Quadrant Clinical Integration Model 21–2
Foy, R. 67–8
fragmentation 57–61
Frank, E. 613
Freedy, J.R. 435, 436, 446, 448
Fris, S. 617
Frueh, B.C. 447
Fu, M.R. 554–5
full intake assessment 228
functional competencies 103
Funderburk, J.S. 479–80
Friis, S. 617
Frueh, B.C. 447
Fu, M.R. 554–5
full intake assessment 228
functional competencies 103
Funderburk, J.S. 479–80
Furher, R.K. 185
GAD-2 326–7
Gaitatzis, A. 649
Gallagher, D. 608
Gallimore, C.E. 768
Gallop, R. 447–8
Garbutt, J.C. 391
Garcia-Campayo, J. 325
Garcini, L.M. 15
Garland, Sheryl 741
Garrett, V.D. 358
Garrison, W. 739
Gartland, D. 644
gastroesophageal reflux disease 588, 590
Gatchel, R.J. 33, 411, 412–13, 415, 417
Gavin, A.R. 612
Geisinger, K.F. 150
gender/gender-keying: anxiety and 325; Beck Depression Inventory and 247; chronic stress and 507; defined 602; emotional distress 174; generalizability and 174; pain assessment and 588; psychological well-being and 617, 619; Short Form Surveys 553; suicidal behavior and 182, 300; when setting normative thresholds 182
general adaptation syndrome (GAS) 356, 357–8
General Health Questionnaire (GHQ) 606
general health status 43
generalizability 174–5
Generalized Anxiety Disorder Inventory 329
Generalized Anxiety Disorder Scale (GAD-7) 37, 85, 148, 329, 748; available norms 322; Chinese version of 322; cross-cultural applicability 336; cutoff scores 325–6; described 320–1; development phase 321; factorial validity 324–5; functional goal 335; implementation barriers 335; internal consistency 322–3; item-to-total score correlations 323; limitation of 345; outcomes assessment 334–6; patient feedback 334; psychometric properties of 322–5; PTSD and 447; sample clinical questions for 327; scoring 320–1; SF-12v2 score 548; Spanish language version of 322, 325; study phase 321; test-retest reliabilities 323; Turkish version of 323
General Practitioner Assessment of Cognition (GPCOG) 380
generic health status measure 36, 529
Geneste, J. 182
George, D. 170–1
Gerber, P.D. 169
Gerhman, P. 449
geriatric depression: antidepressant medications 283; case study 288–9; cognitive behavioral therapy (CBT) for 283–4; cognitive bibliotherapy for 284; outcomes assessment 286–8; psychotherapy for 283; treatment approach for 282–4; treatment monitoring 286; treatment preferences 284–5; see also Geriatric Depression Scale (GDS)
Geriatric Depression Scale (GDS): abbreviated versions of 279; concerns for 277; cutoff values 279, 281–2; for depression screening 146; developing 277–8; limitation of 285, 287–8; with other evaluation data 287; overview 277; protocol use 281–2; suicide risk screening 184; utility of 280; validity of 278
Geriatric Depression Scale-Short Form (GDS-15) 279, 279
geriatric populations: cognitive screening in 185–6; dementia in 86, 185–6; depression comorbidity in 186; general anxiety disorder detection in 326; MMPI-2 with 642; suicidal ideation in 184; see also geriatric depression
Gervais, R.O. 638, 639, 644
Ghushchyan, V. 537
Gift, A.G. 608
Gilbar, O. 616
Gilbody, S. 66, 264
Gillard, N.D. 638
Gill-Rivas, M.V. 604
Glassmire, D.M. 639
Global Deterioration Scale 186
global health measure 421
Global Initiative for Asthma (GINA) 581
Global Mental health (GMH) score 306
Global Sleep Assessment Questionnaire (GSAQ) 478
Glover, G. 149
Gluck, T.M. 617
glycemic control 358
Goetz, R.R. 617
Golding, J.M. 10
Goldsmit, M. 182
Gomez, A. 639
Gonder-Fredrick, L. 616
Gontijo-Guerra, S. 326
Goodie, J.L. 110
Goodpaster, K.P. 647
Gordon, A.J. 182
Gotay, C.C. 616
Grace, M.C. 613
Graduate Psychology Education (GPE) program 733–5
Graham, J.R. 649
Graham, S. 24
Granacher, R.P., Jr. 638
Grant, J.R. 615
Gray, G.Y. 12, 14, 18, 162
Gray, M.J. 444–5
Green, B.L. 613
Greist, J.H. 158
Griffith, E.H. 621
Griger, M.L. 622
Grinenko, D. 516
Grinnell, Rene 754
Grisson, G.R. 38
Gros, D.F. 447
Grossberg, G.T. 377–8
gross domestic product (GDP) 4, 198
Grossman, P. 674
Groth-Marnat, G. 33–4, 144, 149
Group Health Cooperative (GHC) 779–80
Grubaugh, A.L. 451
Grundy, P. 10
Gruss, L. 616
Guidebook of Professional Practices for Behavioral Health and Primary Care Integration: Observations From Exemplary Sites 69
guided imagery 109, 748
Guidelines for Psychological Practice in Health Care Delivery Systems 25
Guidelines for the Diagnosis and Management of Asthma 574–5
Gum, R.L. 13, 19
Gunn, R. 749
Gunn, W.B. 12, 19
Gureje, O. 415
Guthrie, M. 622
Haber, J.D. 16
Haig, R. 17
Hamilton Depression Rating Scale (HDRS) 158, 248, 255, 282, 611
Hamilton Rating Scale for Depression see Hamilton Depression Rating Scale (HDRS)
Hammers, D. 83
Han, B. 442
Hanley, J. 439
Hansen, N.B. 41
Hao, J.Y. 535
Harada, E. 537
Harbin, H. 14
Haroz, E.E. 258
Harp, J.P. 638
Hart, R.P. 185
Harthorn, B.H. 171
Hartman, M. 4
Hartzell, M. 417
Harvard Trauma Questionnaire (HTQ) 452
Hasset, T. 622
Hassles Scale 359, 361
Hatch, A.L. 25
Hathaway, S.R. 632
Hayes Clinic 739, 754
hazardous drinking see alcohol
Headache Impact Test (HIT-6) 555; applications of 574; development of 569–70; Internet-based assessment 569–70; interpreting test results 572–3, 573; psychometric properties 570–2; question pool 569–70; reliability of 571; responsiveness to change 572, 573; scoring of 570; validity of 571–2
headaches 673
health, defined 529
health behavior change 108
Health Brigade, The 740
health care: delivery of 3; evolution of 18
healthcare costs: for behavioral disorder treatment 4–5; of comorbid chronic medical and mental health disorders 10–11; cost savings 24–5, 56; federal spending 55; of integrated health services 5, 24–5; managed care and 82; medical cost-offset 25–6; payment reforms 22; per-member-per-month (PMPM) 55–6; projected 4–5
healthcare coverage, for mentally ill 75
Healthcare Effectiveness Data and Information Set (HEDIS) 195
healthcare liability 168
healthcare systems orientation 617
Health Center Consolidation Act 23
Health Information Technology for Economic and Clinical Health (HITECH) Act 195
Health Insurance Experiment (HIE) 531
health literacy, defined 506–7
Health Professional Shortage Area (HPSA) 736
Health Resources and Services Administration (HRSA) 733
health spending 198
health status: aspects of 43; defined 529; measures of 529; self-assessment of 35–6
health surveillance 531
Healthy Brain and Body: care plan 688, 689–90, 692–3, 694; core processes 676–81, 677; see also Bull’s Eye Plan
Healthy Brain and Body Class: accept difficulties topic 701; action with commitment topic 699; agenda 696; being present topic 693–8; “Chapters in a book” exercise 700; defuse difference topic 700; “driving a bus” metaphor 701; flexible perspectives topic 699–700; introductions 693; “Tell three stories about yourself” exercise 700; topics and skills exercises 697; values topic 698–9; see also Bull’s Eye Plan
Healthy Days Core Measure 681, 682
healthy teams in primary care 703–4
heart disease 588, 590
Heinberg, L.J. 638, 640, 646, 647
Heisel, M.J. 184
Henderson, L. 323
Hendriks, V.M. 614
Henning, K. 615
Hermann, R.C. 43
Hernandez, J. 612
Higgins, E.S. 13
highly distressed patients 617
high stress levels 355; see also stress
Hines, J.L. 376
Hinton, D.E. 674
Hippocrates 356
Hirschman, K. 381
Hispanic community see Latinos
Hispanic Community Health Study/Study of Latinos 257
Hispanic Community Health Study/Study of Latinos 257
Hitchcock, P. 171
HIV infection 617
Ho, F.Y.Y. 484–5
Hodges, J.R. 379
Hodgson, J.L. 79
Hopper, E.W. 171
Hofmann, S.G. 674
Hogg Foundation 11, 16
Hohnmann, A.A. 12–13
Holliday, R. 442
Hollis, J.F. 616
Holman, E.A. 604
Holmes, C.P. 420
Holmes, W.F. 569
homelessness 738
homeostasis 356
Hoogendoorn, W.E. 650
hopelessness 296
Hopkins, T.B. 182
Hopkins Symptom Checklist-25 452, 600
Horowitz, L.M. 184, 603
Horowitz, M.J. 613, 620–1
Horvath, M. 606
Hospital Anxiety and Depression Scale (HADS-A) 324, 329
hospital length of stay (LOS) 25
House, A.T. 25
Howard, K.I. 38
Howland, M. 377
Hsu, K.J. 263
Hubley, S.H. 69
huddle compassion 703
Hudon, C. 530
Hughes, H.M. 615
Human Relations Service in Massachusetts 60
Human Resources and Services Administration (HRSA) substance abuse grants 781
Hunderts, E.M. 95
Hunter, C.L. 11, 12, 33, 110
hypertension 589, 590
hypoglycemia, fear of 504, 508–9, 519
Hypoglycemic Fears Scale-II 504, 509
Hypoglycemic Fears Survey (HFS) 519–20
hypothalamo-pituitary-adrenocortical (HPA) systems 357
hypothetical construct 173
idiopathic insomnia 473
illicit drugs: misuse prevalence rates 392; use of 5
immigrant populations 453
IMPACT model 767, 772
impaired sleep 473
implementation metrics 195
in-bed behaviors 476
incremental validity 147, 647
index event 443
indicated screening 296–7, 304
informant measures 373
Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) 380
information resources 226–8
information source 144–5
information theory 179
informed consent 101, 222–4; see also confidentiality
Inghilleri, M. 640
initial introduction see warm handoffs
insomnia: BISS vs. 474; chronic 468, 473; cognitive behavioral approaches to 469, 471; cognitive behavioral therapy for 746; comorbid 470–1, 473; DSPD vs. 474; idiopathic 473; interventions for 109; perpetuating factors 470, 476; precipitating factors 470, 475; primary 473; PTSD and 449; screening instruments for 86; secondary 473
Insomnia Severity Index (ISI) 479
Institute for Healthcare Improvement (IHI) 19
Institute of Medicine (IOM) 63
insufficient sleep 508
intake guide 232–4
integrated behavioral care (IBC) clinic 750
integrated behavioral health: defined 57; orienting patient to 229–31
integrated care: Accountable Care Organization 23; barriers to 28–9; benefits of 23–6; clinical competencies in 106–11; cost-offset 25–6; cost savings 735; defined 17; examples of 22–3; federally qualified health centers (FQHC) 23; goal of 77; impetus for 18–19; information resources in 226–8; key dimensions of 19–20; patient-centered medical homes 22; patient-centered medical homes and 76–8; qualities of 19–20; service delivery models 20–2; students’ perspectives 96–8; training competencies 741–2, 743; training for 69–70; World Health Organization on 18; see also collaborative care
integrated care model: PTSD in 432; in safety net settings 733; students’ perspectives 97; utilizing screeners 80–1
integrated care movement 63
integrated care settings: benefits of 77; cultural differences in 28; psychologists in 26–8
Integrated care: The Future of Medical and Mental Health Collaboration (Blount) 65
integrated pediatric primary care 739–40
integrated primary healthcare settings see primary care settings
integrated service delivery 20–2
integration: barriers to 68–70; defined 17, 66; evidence based for 66–8
integration barriers 28–9
integrative and collaborative care 99–100
integrative assessment 631
interactive voice response (IVR) 576
interdisciplinary skills 749
internal behavior health consultants (IBHCs) 105, 218–19
internal medicine residency program 749
internal medicine residents 744
International Association for the Study of Pain (IASP) 415, 584
International Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification (ICD-10-CM) 250, 250, 251
International Quality of Life Assessment (IQOLA) Project 150, 532
International Test Commission 150
Internet-based assessment 156–7, 159
Internet-delivered treatment 485
internship/residency 113
Interorganizational Work Group on Competencies for Primary Care Psychology Practice 32, 103
Interpersonal scales of MMPI-2-RF 637, 641
interventions: application cluster 104–5; in behavioral health 107; defined 722–4; in integrated care settings 27; web-based 485
interview guide 232–4
intimate partner violence 86
IQOLA project 538–9
Irons, T.G. 79
irritable bowel syndrome 589, 590
IRT models see item response theory (IRT)
Irvine, A.A. 616
Israel, Y. 182
item response probability functions (IRPF) 567, 568
item response theory (IRT) 157, 255, 340, 421, 479, 567–9
iteration, concept of 772
Ives, J.O. 620
IVR technology 158–9, 159
Jackson, T.A. 381
Jackson, V. 25
Jacobs, D.M. 149
Jacobsen, L.K. 450
Jacobson, N.S. 551–2, 619
Jahoda, M. 529
James, William 59
Jasinski, L.J. 638
Jefferson, J.W. 158
Jeffrey, R.W. 359
Johnson, D. 12
Johnson, J.G. 8, 599–600, 617
Johnson, M.E. 614, 615
Johnson, Susan Bennett 96, 103
Johnston, D.M. 448–9
Johnston, K.M. 376
“Joint Principles: Integrating Behavioral Health Care into the Patient-Centered Medical Home” 64
Joint Principles of the Patient-Centered Medical Home (JPPCMH) 76
Jonas, D.E. 391
Jones, G.N. 360
Jones, Heather 737, 740
Jones, J.R. 88
Jongh, A. de 450
Journal of Psychiatric Research 377
“Just in Time” learning 744
Kaiser Permanente Care Management Institute (KP-CMI) 585
Kaltman, S. 448
Kaltreider, N. 613
Kane, L. 666
Kanner, A.D. 359
Kanzler, K.E. 101
Kapenko, V. 83
Kaplan DeNour, A. 616
Kazmark, P. 378
Katzelnick, D.J. 158
Kaufman Brief Intelligence Exam (KBIT) 766
Kavanagh, H.B. 181, 599
Keefe, F.J. 420
Kellner, R. 612
Kelly, S.J. 615
Kennedy, C.A. 617
Kennedy, John F. 60
Kert, S. 322, 323, 336
Kessler, L.G. 181–2
Kessler, R.C. 12, 14, 376
Khatri, P. 102–3
Kilpatrick, D.G. 613
Kikinis, E. 88
King, T.S. 376
Kiosks 724–5
Kisesuk, T.J. 153
Kishino, N.D. 417
Klag, M.J. 182
Klaus, S.P. 379
Klerman, G.L. 599–600
Kliem, S. 249
Klingman, K.J. 474
Klinkman, M.S. 12
Klonsky, E.D. 296
knowledge-based competencies 106
Kobak, K.A. 158
Koch, R. 611
Koer, B.W. 650
Koeter, M.W. 606
Kolko, D.J. 739
Konik, R. 323
Koos, E. 381
Kopt, S.M. 306
Kornetsky, C. 611
Kosinski, M. 536, 547
Kosten, T.R. 450
Kraemer, H.C. 300
Kraepelin, E. 632
Kramer, P.D. 168
Krishnamurthy, R. 32
Kroenke, K. 261
Kroese, R.L. 75
Kudler, H.S. 613
Kukal, D. 648
Kupfer, D.J. 613
Kuramoto, R. 12
Kwan, B.M. 68
Laderman, M. 13, 15, 17, 19
Lam, C.L. 554
Lam, T.P. 554
Lambert, M.J. 25, 41, 152, 550, 619
Lamson, A.L. 29
Lancman, M. 650
Lang, A.J. 441
language functions 374
Lanoye, Autumn 755
Larkin, K.T. 106
Larner, A.J. 382
Larsen, S. 641
Latinos: anxiety and 340; depression and 248, 257; PCPTN focus on 736, 738–9; posttraumatic stress disorder 451, 452; substance abuse screening 396
Lavery, M. 646
Lazarus, R.S. 356
Leaviss, J. 675
Ledoux, T. 15
Lee, T.T. 639, 644
Lee, R. 379
Leitenberg, H. 615
Levant, R.F. 25, 30
levels of care (LOCs) 146–7, 196–7
Levine, D. 182
Lewis, D.L. 378
Lexicon for Mental Health and Primary Care Integration (Peek) 194
Li, G. 535
Libby, A.M. 537
Liddel, D.L. 184
Liegey-Dougall, A.L. 417
Life Events Checklist (LEC-5) 442–6, 460
Life Experiences Survey 359
Linden, W. 616–17
Lindhorst, T. 612
Lind, J.D. 613
Lind, J.G. 613
Linn, L. 170
Lip, J.D. 183
Locke, D.E. 640, 649
Loeppke, R. 9
Loeppke, R. 9
Loeppke, O. 632
Lohr, M.J. 612
Longitudinal Expert All Data (LEAD) procedure 403
long-term disability 59
Lopez, A.D. 611
Louria, D.B. 617
Lowe, B. 322, 323, 324, 325, 611, 619
Lowery, J.L. 306
Lum, D.L. 531
Luxton, D.D. 160
Lynch, J. 616
McBeth, C.D. 622
McCartney, G. 378
McCord, D.M. 641–2, 650
McCroy, P. 376
McDaniel, L.H. 103, 104
McDaniel, S.H. 17, 19
McDonald, S.D. 440
McDowell, I. 530
McGill Pain Questionnaire (MPQ) 584
McGrath, P.S. 620
Bagwell, E.A. 569
MacGregor, K.L. 479
Machin, D. 531
McHugh, R.K. 148, 149
McIntosh, D.N. 604
Mack, S. 547, 554
Main, C.J. 632
Maisto, S.A. 479
MacLean, L.M. 447–8
McManus, J. 639
Magill, M. 170–1
Magruder, K.M. 171
Magruder-Habib, K. 171
Maguire, T.A. 638
Malik, A.D. 65
Malmstrom, T.K. 377–8
Manea, L. 264, 265
Manly, J.J. 149
Man, J. 650
Maisto, S.A. 479
Majer, A. 617
Majer, S. 170–1
Margruder, K.M. 479
Margruder-Habib, K. 171
Mayer, M.W. 155
Mayer, M.W. 155
Mayer, G.E. 621
Malec, J.J. 179
Maier, A. 617
Majer, S. 170–1
Margruder, K.M. 479
Margruder-Habib, K. 171
Mayer, M.W. 155
Mayer, G.E. 621
mail-out/mail-back process 156, 159
Main, C.J. 650
Maisto, S.A. 479
major depressive disorder (MDD) 177–8, 245, 264, 303, 448, 600; see also depression
Major Depressive Syndrome 261, 265
maladaptive behaviors 487, 490, 507
Malik, A.D. 65
Malmstrom, T.K. 377–8
managed care: effects of psychological assessment on 82–3; guidelines for assessment practice in 83; healthcare costs and 82
Manea, L. 264, 265
Manly, J.J. 149
Mann, J.D. 615
Marek, R.J. 638, 640, 642, 646, 647, 648
Margolis, A.R. 768
Margolis, R.B. 37
Marmar, C. 620–1
Marsh, J.D. 538
Martin, A.B. 4
Martin, M.P. 79
Martinez, E.Z. 618
Martinez, J. 639
Maruish, M.E. 18, 43, 531, 536, 538, 546
Matthews, E.J. 611
Maudsley Dementia Rating Scale 369–70
Maudsley Addiction Profile (MAP) 400, 402
May, A.M. 296
May, S. 25
Meagher, D. 378
Mears, F. 413
measurable outcomes, defined 43
measurement-based care (MBC): defined 713; facilitators of 727–8; implementation barriers 722–6; implementation costs 724; intervention characteristics 722–4; in PCMHI care management 716–17; social marketing approach to 727; staffing levels 726
measurement feedback systems 201
Measurements in the Addictions for Triage and Evaluation (MATE) 400, 402
Medicaid: Access Community Health Centers and 763, 774, 778, 781; cost savings 24, 56; costs of 10; lack of psychologists who accept 781; reimbursement from 23; spending 4; Veterans Health Administration and 735, 756; see also Centers for Medicare and Medicaid Services (CMS)
medical clinic interview 216–18
medical cost-offset 25–6
medical costs see healthcare costs
Medically Underserved Area (MUA) 736
Medical Outcomes Study Short-Form General Health Survey (SF-20) 321, 531–2
medical records 114, 223–4, 227
medical settings: increasing use of psychologists in 81; rapport building in 218
Medical Symptom Validity Test 638
medication assisted treatment (MAT) 768, 781
medications: antidepressants 12, 283, 777; antipsychotic 60; BHC model for 766; chlorpromazine 60; heroin and 675; opioid 422; opioid pain 675; overdosing on 675; for persistent pain 675; phenothiazine 60; psychotropic 12, 14, 76; Thorazine 60
Meehl, P.E. 176
MeHAF tool 770
Meijer, R.R. 604
Meisel, A.W. 614
Melek, S.P. 24
Melisaratos, N. 616
Mellergerard, M. 611
Melnik, K. 88
memory (cognitive functioning) 374
Mendelhall, T.J. 65
Mendelhall, E. 168
Mental Component Summary (MCS) measure 547–8
mental health care: defined 57; in prison system 61, 61
mental health clinicians see clinicians
mental health disorders: chronic pain and 415–16; comorbidity with medical conditions 9–11;
co-occurring prevalence rate 6; cost of co-occurring
199; detection and treatment of 13–14; see also
behavioral health disorders
mental health interview 214–16
Mental Health Inventory-5 (MHI-5) 545–7
Mental Health Inventory-38 (MHI-38) 545
mental health measure 421
mental health professional, role of 81–2
mental health services 13
Mental Health Severity Index 547
Mental Health Survey 183
mental health system; PCP-driven de facto 14–15; as
PCP-driven system 12
Mental Health Systems Act 61
mental hygiene movement 58, 59
mental illness: biopsychosocial approach to 59; short-
term treatments 60; stigma of 15
mentally ill, lack of coverage for 75
Mercier, C. 615
Merrell, J. 640
Meshberg-Cohen, S. 608
Messick, S. 175
Metzger, D.S. 614
Meyer, Adolf 59, 180
Michaels, R. 214–15
Michigan Alcoholism Screening Test (MAST) 394,
396–7
Middlesex Hospital Questionnaire (MHQ) 606
migraine headaches 589, 590
Milanak, M.E. 447
mild cognitive impairment 86
Millard, R.W. 158
Miller, B.F. 13, 19, 65, 69, 741
Miller, M.W. 446
Millman Report 14, 56, 735
mild cognitive impairment 86
Millon Clinical Multiaxial Inventory-III (MCMI-III)
766
Mills, M.E. 614
mind and body dualism 742
mindfulness-based stress reduction (MBSR) 674
mindfulness techniques/intervention 483–4, 617, 692
Mini-Cog 381
Mini-International Neuropsychiatric Interview 255
minimal clinically important difference (MCID)
204–5, 577, 582–3, 590–1
Mini-Mental State Exam (MMSE) 186, 369, 377; see
also Folstein Mini-Mental State Exam (MMSE)
Minnesota Multiphasic Personality Inventory: Clinical
Scales scores 632; criticism of 632; MMPI-2 34, 411,
417, 632–3; MMPI-2-RF 417; negative activation
633; positive activation 633; purpose of 632;
Restandardization Project 632–3; SCL-90-R and 606;
Wiggins Content Scales 606; see also Minnesota
Multiphasic Personality Inventory-2-Restructured
Form (MMPI-2-RF)
Minnen, A. van 450
Minnesota Multiphasic Personality Inventory-2-
Restructured Form (MMPI-2-RF); administration of
642–3; bariatric surgery settings 646–8; behavioral/
externalizing domain 636, 640; in behavioral
medicine settings 645–51; case study 651–6;
for chronic low back pain 650–1; comparison
group 644; described 631; development of 633–5;
eating disorders and 647; emotional/externalizing
domain 635–6, 639; incremental validity 647;
interest domain 637; interpersonal domain 637,
641; interpretation strategy 644–5; interpreting
test results 653–4; limitation of 651; normative
data availability 643; presurgical psychological
evaluations 646–8; psychometrics 641–2; scales of
635–7; scoring of 643; seizure disorders and 649–50;
somatic/cognitive domain 637, 640–1; for spine
surgery evaluations 648; substantive scales 639–41;
Technical Manual 641–2; thought dysfunction
domain 636, 640; validity scales 635, 638
Mior, S.A. 17
Miranda, J. 12–13
missed work days 592
Mitchell, A.J. 377
Mitchell, J.B. 168
MMPI-2 34, 411, 417
MMPI-2-RF 417
models of care 194, 200
modified CBT-I 486
Modified Checklist for Autism in Toddlers (M-CHAT)
756
modified Zung Depression Index (mZDI) 342
Moffett, L.A. 606
Mokken scale 604
Molle, D.W. 378
Monden, K. 777
monoamine oxidase inhibitor (MAOI) 620
Montgomery, W. 537
Montgomery-Asberg Depression Rating Scale
(MADRS) 282
Montreal Cognitive Assessment (MoCA) 377, 378,
382
Moore, J. 170–1
moral treatment movement 57–9
Morgan, J.F. 514, 515
Mori, D.L. 319, 344
Morris, M.C. 448
Mossle, T. 249
MOS Sleep Scale-Revised (MOS Sleep-R) 555
motivational interviewing 108, 219, 748
motor skills 374
Mottram, P.G. 282–3
multiple relationships 101
Multiple Sleep Latency Test 477
Mundt, J.C. 158
Muntaner, C. 252
Murphy, J.M. 179
Murray, C.J. 611
Myers, L. 650
My Mood Monitor (M-3) Checklist 330
Nacul, L.C. 535
Nael, S. 622
Naqvi, S.H. 381
Narayana, S. 37, 38, 39, 151
narcolepsy 477, 478
Nash, J.M. 32, 102–3
Nathan, R.A. 577
National Action Alliance for Suicide Prevention 295
National Committee for Mental Hygiene 59
National Committee for Quality Assurance (NCQA)
22
National Comorbidity Survey Replication (NCS-R)
9–10, 553–4
National Depression Screening Day 167
National Epidemiologic Survey on Alcoholism and
Related Conditions 256–7
National Health and Nutrition Survey (NHANES-I) 254–5
National Health and Wellness Survey (NHWS) 582
National Health Expenditure Accounts Team 4
national health expenditures 4
National Health Interview Survey 9
National Healthy Sleep Awareness Project 470
National Heart, Lung, and Blood Institute (NHLBI) 574–5
National Institute of Mental Health (NIMH) 75, 153, 251, 421, 599
National Institute on Alcohol Abuse and Alcoholism (NIAAA) 395
National Institutes of Health (NIH) 340, 424
National Mental Health Association 15
National Organization for the Reformation of Marijuana Laws listserv (NORML) 256
National PCMHI Evaluation Office 712–13
National Strategy for Suicide Prevention 295
National Survey of Functional Health Status 532
National Survey of Headache Impact (NSHI) 570
National Survey on Drug Use and Health (NSDUH) 5–6, 392
Naudie, D.D. 538
Nease, D.E. 12, 68
Neblett, R. 419
needs analysis 115n2
NEEDS Lifestyle Checklist 685, 687
NEEDS Lifestyle Class 701–3
Negative Affect Scale 256
negative predictive power (NPP) 151–2, 440–1
negative predictive value (NPV) 326
Nelson, P.D. 95
Nesse, R.M. 612
neuropathic pain 671
neuropasticity 688–92
neuropsychological tests 35
Newell, C. 530
Newman, F.L. 153, 154, 163, 287
Newman, M.G. 441
Ng, D.M. 359
Nielson, M. 22
Niemic, R. 37
Nies, A. 620
nociceptive pain 671
Nock, M.K. 183
Nolan, T.W. 666
Noland, R.M. 150
nonepileptic seizures 649–50
nonsuicidal self-injury (NSSI) 616
O’Conner, M. 510
O’Connor, P.G. 182
O’Donnell, M.L. 448
Office of Patient Centered Care and Cultural Transformation (OPCC&CT) 712
Office of the Surgeon General 295
Olson, M. 12
Olmsted, M.P. 517
Omidi, A. 617
oncology 616–17; see also cancer
O’Neill, D. 378
Oordt, M.S. 33, 110
open chart model 742
open-ended questions 97
operational barriers to integrated care 29
operational metrics 196
opiate abuse 86
opioid induced hyperalgesia (OIH) 422
opioid pain medication epidemic 422
Opioid Risk Tool (ORT) 398
Organization for Economic Cooperation and Development (OECD) 198, 199
Oss, M.E. 19
osteoarthritis 588, 589
osteoporosis 588, 589
Other Depressive Syndrome 265
Ottoson, J.O. 611
Ouimette, P. 432, 433, 435, 438
outcome metrics 196
outcomes: defined 42; measurable 43; patient-reported 43
outcomes assessment: acceptability 154; for anxiety treatment monitoring 334–6; for BAI-PC 344; for BSI 618–21; characteristics of 152–3; in clinical drug trials 619–20; defined 201; described 30; for EDE-Q 512; follow-up remeasurement for 162; gender-keyed 619; for geriatric depression 286–8; instrument selection for 152–5; measuring change 204–5; measuring implementation metrics 205; methods and procedures 153; metrics selection 205; practicality of 154–5; PROMIS-Anxiety 342–3; psychometric strengths 153; relevance to target group 153; for SCL-90-R 618–21; for Short Form surveys 550–2
outcomes management 42–3
Overall Anxiety and Severity and Impairment Scale (OASIS) 330, 337–40
overnight polysomnography (PSG) 479
Owby, R. 381
Oxman, T.E. 66
Pagura, J. 448
pain: acute 671; biological pathway of 670–1; biopsychosocial model of 412, 650; central sensitization to 688; defined 415, 584; defused response to 676; fear of 414–15; as fifth vital sign 685; neuropathic 671; nociceptive 671; psychosocial and personality characteristics 412; see also chronic pain; persistent pain (PP)
Pain Anxiety Symptoms Scale (PASS-20) 330
pain assessment: Beck Depression Inventory-Second Edition (BDI-II) 419; behavioral observation 420; case study 423; coping style 413; Fear Avoidance Components Scale (FACS) 419–20; multidimensional approach to 412–16; Pain Catastrophizing Scale 414, 419; Pain Disability
Pain Impact Questionnaire (PIQ-6): acute version 417
pain drawings 417
Pain Disability Questionnaire (PDQ) 418–19
pain disability 86
Pain Catastrophizing Scale 414, 419
Pain Impact Questionnaire–Revised (PIQ-R) 36, 555
Pain Medication Questionnaire (PMQ) 420
pain-prone personality syndrome 411, 417, 424
Pain-Related Self-Statements Scale 414
pain reliever use disorder 392
Pain Severity Rating 587, 591
panic disorder 326, 327, 447, 599, 611–12, 613, 620
Panic Disorder Self-Report Questionnaire (PDSR) 331
Paolaggi, J. 650
paper-and-pencil test forms 156, 340
Park, M. 187
Park, S. 325–6
Park, M. 187
Parkinson, H.A. 336
Parkinson’s disease 379
Patel, V. 168
paternal incest 448
Pathak, D. 612
Patient Aligned Care Team (PACT) 712
Patient-Centered Medical Home Recognition 32–3
patient-centered medical homes 22, 64, 76–8, 102–3
patient-centeredness 216
Patient-Centered Primary Care Collaborative
patient engagement 222, 777
patient feedback 285, 287, 436
Patient Health Questionnaire (PHQ-9): consistency with DSM-5 criteria 265–6, 266; consistency with ICD-10-CM criteria 266–7, 266; defined 85; described 260–1; implementation issues 267–70; for insomnia 479; interpreting test results 303; items and responses to 261–2; norms for sex and age 263; for pain assessment 418; positive screens 304; psychometric properties 262–4; reading level of 148; as screening tool 37; suicide risk screening 302–5; validation studies 269; validity of 305
Patient Health Questionnaire-2: consistency with DSM-5 criteria 265–6, 266; consistency with ICD-10-CM criteria 266–7, 266; defined 262; implementation issues 267–70
Patient Health Questionnaire-8 (PHQ-8) 323
patient help questionnaires: Patient Health Questionnaire (PHQ-9) 260–1; Patient Health Questionnaire-2 262
patient portals 88–9
Patient Protection and Affordable Care Act (ACA) 18–19, 64, 75, 667
patient-provider relationship see doctor-patient relationship
Patient questionnaire (PQ) 517–19
patient-reported outcome measures (PROMs) 201, 530–1
patient-reported outcomes (PRO) 43
patients: care providers impacting 194–5; difficult 86; feedback from 285; functional status assessment 86; on “hive” approach to care 773–4; problems, identification and clarification of 40; reading levels 148–9; satisfaction surveys 36, 146, 681–5, 684
Patterson, M.B. 381
Pattison, P. 448
pay-for-performance 69
payment framework 64
payment models 68–9
Payne-Murphy, J.C. 13
PCBH Big 10 Checklist 667, 668
Peabody, Francis 63
Peckham, C. 666
pediatric primary care 739–40
Pediatric Symptoms Checklist 17 (PSC-17) 756
Peek, C.J. 66
PEG scale 681, 683
Pendlebury, S.T. 379
Penn State Worry Questionnaire (PSWQ) 331, 335
Penn State Worry Questionnaire 3-item version (PSWQ-3) 331
Penn State Worry Questionnaire 8-item version (PSWQ-A) 331
Perceived Stress Scale (PSS) 359–60
Perconte, S.T. 622
Peretz, T. 616
periodic leg movements 477
peripheral nervous system 688, 691
perpetuating factors to insomnia 470, 476
Perrin, Paul 737
Perrine, K. 650
persistent pain (PP) 673; assessment 681–5; brain training 688; central sensitization 667–9; central sensitization and 671–2; compassion-focused therapy for 674–5; defined 671; development of 671; etiology of 670–2; Focused Acceptance and Commitment Therapy 665; lifestyle risk factors 685; medications for 675; mindfulness and flexible perspective to 692; NEEDS Lifestyle Checklist 687; neurological perspectives on 670–2; nonpharmacological therapies for 672–5; patient population 667–9; patient satisfaction surveys 681–5, 684; PEG scale 681, 683; positive psychology and 675; Quadruple Aim and 666; sleep-related disorders and 672; substance use disorders and 672, 674; treatment for 672–5; Triple Aim outcomes with 666; see also chronic pain; pain
Personal Data Sheet 180
personality disorders 415–16, 674
personality/psychopathology 86
Personality Psychopathology–Five (PSY-5) scales 635
personality psychopathology scale 417–18
personalized, proactive patient directed care (PPDC) 712
personal physician 64
Peters, L. 154
Petkus, K.J. 608
Pettersen, S. 13, 64
Peveler, R.C. 606–7
pharmaceuticals 468
pharmacotherapies 282, 471, 619–20
phenelzine 620
phenothiazine 60
Phillips, R.L. 13
Phipps Behavior Chart 180
physical activity guidelines 109
physical functioning (PF) domain scale 533
physical health measure 421
physician-directed medical practice 64
physician-patient relationship 77, 81, 213
Physician Quality Reporting System (PQRS) 196
physicians see primary care providers
Pierucci-Lagha, A. 401
Pietrzak, R.H. 447
Pigeon, W.R. 468, 479
Pignone, M. 39, 171
Pilkonis, P.A. 340, 341
Pinna, K.L.M. 448–9
Pittsburgh Sleep Quality Index (PSQI) 477–8
Pivik, J. 414
Platt, J.J. 614
Polatin, P.B. 415
Polen, M.R. 392
Poleshuck, E.L. 14
Pollard, C.A. 37
polysomnography (PSG) 479
Poppik, M.N. van 650
population-based care, defined 115n1
population-based screening initiatives 715–16
population-based screening, defined 716
populations 715
postsurgical pain 716
posttraumatic stress disorder (PTSD) 86; acute
postdoctoral training 113
postdoctoral fellowship program 106
postconcussive symptom scale (PCSS) 376
primary care assessment instruments 84
primary care: clinical skills set 742; cost-effectiveness
primary care providers: biomedical model of illness
primary care: practice guidelines 715; war-zone PTSD scale 613;
primary career management 98, 104, 742
practice management model 745–6
primary care: practice guidelines 715; war-zone PTSD scale 613;
see also trauma; trauma assessment
Potter, J.S. 674
practice guidelines 98, 104, 742
Practice Management System (PMS) 196
preceptorship model 745–6
preempting event 443
precipitating factors to insomnia 470, 476
predictive validity see validity, screening
predictive validity of assessment 174
predisposing factors to insomnia 470, 476
pregnant women: alcohol misuse screening 200, 390, 391, 396
with depression 14; depression and 612;
eating-related problems 500, 509; GAD-7 screening
tool and 325–6; illicit drug use 392; substance use
screenings for 393
prescriptions: antidepressants 12; misuse of 109–10;
of psychotropic medications 12
Present State Examination (PSE) 606–7
preventable inpatient hospitalizations 755
Previle, M. 326
primary care: clinical skills set 742; cost-effectiveness
of 757; as de facto behavioral healthcare system 11;
defined 3–4, 62; ethical decision-making in 101–2;
history of 62–4; informed consent within 101;
integrating behavioral health 13; intervention skills
key characteristics of 62; training competencies
741–2, 743; as training setting 757; undetected
psychiatric disorder in 169
primary care assessment instruments 84
Primary Care Behavioral Health (PCBH) model 219,
225–6, 667, 775, 779; see also Access Community
Health Centers; Group Health Cooperative (GHC)
Primary Care Behavioral Health Provider
Questionnaire (PPAQ) 115
Primary Care Checklist (PCL) 439–42
Primary Care Evaluation of Mental Disorders
(PRIME-MD) 8, 84–5, 158, 260–1, 503–4,
517–19
Primary Care-Mental Health Integration (PCMHI):
components of 711–12; cumulative new patients
and outpatient encounters for 713; implementation
of 712–14, 714; measurement-based care in 714,
716–17; National PCMHI Evaluation Office 712–13;
psychological assessment within 714–15; screening
process in 715–16, 716; see also Veterans Health
Administration (VHA)
primary care providers: biomedical model of illness
15; burnout 666; clinical interview of 218;
difficulty identifying/diagnosing mental disorders
78; engaging internal behavior consultants 219;
imbalanced wages 62; improving work life of 669–70;
interactions with patients 680; overprescribing
psychotropic medications 14; patient reading
levels 148–9; patient time with 15; prescribing
psychoactive medications 12; relationship building
with other 104; satisfaction surveys 685, 686;
skills for psychological assessments 32–3; specific
skills for successful integration 110–11; treating
behavioral healthcare problems 13–18; treatment
philosophies 28
primary care psychology; see also psychologists
Primary Care Psychology Training Network
(PCPTN): clinical outcomes studies 753–5;
community-based HIV/AIDS outreach program
740; complex care clinic partnership 740–1;
faculty of 737; homelessness and 738; impact of
744; interprofessional training aspect of 749–51; interventions in 743; Latino community and 738–9; master’s theses and dissertations 752; mission of 733, 751; pediatric primary care 739–40; primary care sensitive diagnosis 740–1; real-world studies in safety net setting 752; research contributions by 751–6; safety net clinics 737; students training in 758; team-based learning 750–1; training sites of 734; underserved communities focus 736–40
Primary Care-PTSD Screen (PC-PTSD) 435, 438–9, 457
primary care settings: behavioral health and 83–5; behavioral health disorders prevalence rates in 7–8; cognitive screening in 185; depression, treatment of 12; diagnostic aids in 170–1; documentation for 197; general competencies for working in 102–6; improving treatment of common mental disorders 16; mental illness detection in 13–14; mini-battery in 88; operational barriers 29; psychotropic medications in 12; pull-in 86–7; warm handoff 87
primary care training 741–6
primary insomnia 473
Prins, A. 439
prisons 61, 61
privileges 99
problem-focused coping 357
process, quality of care 42
process metrics 196
Procrustes method 606
professional identity 99
professional organizations 113
professional trainings 113
pro formas 196
program monitoring 205
program outcomes 778–9
progressive muscle relaxation 109, 468
PROMIS-Anxiety 330, 340–3, 346
PROMIS Cooperative Group 340
PROMIS-Sleep Disturbance (SD) 478–9
PROMIS-Related Impairment (SRI) 478–9
propriety score matching 755
provider-related factors of suicide risk assessment 310–11
pseudodementia 186
psychiatric disorders: aided recognition in 170–1; clinical recognition of 168–71; community-based rates of 167; prevalence of 167, 170; prevalence rates of 599; in primary care 76; psychometric principles in screening 173–5; recognition rates 168; screening 167–8; screening tests for 180
psychiatric history 487
psychiatric hospitals 58–9
psychiatric morbidity, rates of accurate identification of 169–70, 170
Psychiatric Research Interview for Substance and Mental Disorders (PRISM) 399, 402–3, 403
psychiatric screenings: caseness criterion 610; in medical settings 180–1; prevalence rates of 180–1
psychoeducation 436
psychogenic nonepileptic seizures 649–50
Psychogenic Nonepileptic Seizures Attitudes (PNES-a) Scale 649
Psychogenic Nonepileptic Seizures Physical Complaints (PNES-pc) Scale 649
psychological assessment: application cluster 104; benefits of 39–40; biopsychosocial approach to 31; clinical decision making activities in 30; clinical interview 33–4; components of 26; comprehensive vs. brief 33; concern-focused assessment 228–9; data collection 33–4; declining use of 83; defined 31; disorder-specific measures 35; documentation 500; early implementations of 81–2; effects of managed care on 82–3; full intake assessment 228; functional 499–500; future of 88–9; goal of 80; integrated care models utilizing 79–80; multidimensional instruments 34–5; for neuropsychological functioning 35; origins of 81; outcomes assessment 30; as outcomes management tool 42–3; patient bias in reporting 175; presurgical 646–8; primary care instruments for 84–5; for problem identification and clarification 40; psychological/psychiatric symptom measures 34–5; psychological testing vs. 31; psychologist skills and competencies 32–3; purpose of 714–15; as screening tool 36–9; terms for conducting 195–6; test selection criteria 80; as treatment monitoring tool 40–2; as treatment planning tool 39–40; see also clinical interview; psychological testing; psychological test instruments
psychological assessment technology see assessment technology
psychological flexibility 676, 678, 679
psychological practice, core competencies for 95
psychological testing: defined 31; history of 87–8; see also psychological assessment; psychological test instruments
psychological test instruments: actionable information 146, 150; adapted versions of 149–50, 151; brevity of 147–8, 151; clinical utility 147, 151; computer-administered measures 149; content 145; continuum of care 146–7, 150; cost-benefit analysis of 143–4; costs of 149, 151, 153–4; ease of use 147, 151; evaluating usefulness of 319–20; feasibility of 149, 151; general considerations for selecting 150–1; instrument selection 144–50; intended purpose relevancy 144; negative predictive power 151–2; normative data availability 146; positive predictive power 151; practicality of 149; psychometric properties 145; psychometric strengths 153; reactivity 149; reading level of 148–9; reading levels 151; for screening and diagnosis 151–2; self-report vs. clinician rating 175–6; source of information 144–5, 150; test-retest reliabilities 145; translated versions of 149–50, 151; validity considerations 145; value 149
psychologists: accessibility of 111; adapting to primary care 111; as administrator/leader 27–8; as clinicians 26–7; commitment to education 105; competencies 95; competencies for 102–3; competency 100; as consultants 27; core skills in assessment and treatment 30; functions of in primary care 78; integrating into primary care clinics 77–8; integrative and collaborative care 99–100; privileges 99; professional identity 99; relationship building 111; as researcher 27; roles of 26–8; skills for successful integration 110–11; as teacher/supervisor 27; as team members 111; training for 112–14
psychology: competencies for 96; competency-based education in 95; as emerging area of practice 95; experience level of 108; foundations shaping practice 98–102; measuring competence...
in 95; scientific foundation of 103–4; students’ perspectives 96–8; training opportunities in 108, 112–14
psychology trainee 743
psychometric integrity of automated test administration 160
psychometric properties: of ACT 577–82; for BAI-PC 343–4; of Beck Depression Inventory-FastScreen (BDI-FS) 249; of BHM-20 307; of CAPS-5 445; of CES-D 254–6; of CESD-R 256–7; cognitive impairment 373; of disease impact measures 567; with DSI 360–1; of GAD-7 322–5; of HIT-6 570; for MMPI-2-RF 641–2; of PHQ-2 264–5; of PHQ-9 262–4, 304; of PIQ-6 586; of PRIME-MD 518; of PROMIS-Anxiety 341–2; of PTSD assessment 451–3; for SCOFF 514; for SCS 308; for SF-12v2 536–7; for SF-36v2 534–6; with WSI 362–3
psychomotor retardation or agitation 253
psychopathological symptomatology 34
psychophysics, signal detection paradigms in 179
psychosis 450–1, 674, 718
psychosocial treatment 11
psychotherapy 283, 306, 620–1
psychotropic medications 12, 14, 76
Ptito, A. 376
PTSD see posttraumatic stress disorder (PTSD)
PTSD Checklist-Civilian (PCL-C) 439
PTSD Checklist for DSM-5 435, 458–9
public educational campaigns on childhood and sexual behavior 60
Pulier, M.L. 155
pull-in 86–7
Quaddflieg, N. 619
Quadruple Aim 193–4, 206–7, 666; see also Triple Aim
quality improvement practices 775
QualityMetric 2009 Norming Study 532, 539
QualityMetric Incorporated 534
quality metrics 196
quality of care 42, 666–70
quality of life 477
Quantified Pain Drawing 417
Questionnaire on Eating and Weight Patterns-Revised (QEWPR) 502, 512–14, 513
Quilty, L.C. 640
Quitkin, F.M. 611
Rabkin, J.G. 617
racial/ethnic groups: future population of 451; illicit drug use 392; PCPTN focus on 736–7; psychosis in 450; safety net clinics serving 746; screening considerations for 451–3
racism, trauma-associated 452–3
Radcliff, Z. 756
Radclenhausen, R. 606
Radloff, Lenore Sawyer 251, 255, 270
Raju, M. 15
Ramirez, G. 39, 171
Rao, Michael 737
Rapid Alcohol Problem Screen (RAPS-4) 182, 394
Rapid Dementia Screening Test (RDST) 381
Rapid Estimate of Adult Literacy in Medicine 149
Rappaport, N. 360
rapport building 214–18, 228, 442
Raskin Rating Scale 255
Ravaris, C.L. 620
RCd scale 417
reactivity 149
reading level 148
Reagan, Ronald 61
real-time charting 98
receiver operating characteristic (ROC) analysis: of ACT 577–8; of BDI-FS 249; of BDI-II 419; of CES-D 257; of DBAS-16 478; defined 178–80; for delirium 186; of GAD-7 326; of MHI-5 546; of PHQ-2 264; of PHQ-9 262; of SCL-90-R 606; for suicidal behavior 184
reconnaissance 215
referrals: at Access Community Health Centers 777; barriers to 778; to behavioral health consultants 219–22, 710; behavior changes prompting 221–2; chronic disease management 221–2; cognitive impairment 370; health behavior change 221–2; for obstructive sleep apnea (OSA) 468–9; reasons for 220–2, 236; risk assessment 220–1; substance use disorders 221; see also warm handoffs
refugee populations 453
Regier, D.A. 11
registered nurses (RNs) as depression care managers 719–20
regression method 179
Reid, M.C. 182
Reifler, D.R. 171
Reiter, J.T. 97–8, 219, 669–70, 741, 765, 778–9
relationship building 104, 111
relationship competency cluster 104
relaxation: skills 109; techniques 748; training 483
reliable change index (RCI) 204–5, 335–6, 552, 619
Remien, R.H. 617
reporting, patient bias in 175
researcher, psychologists as 27
Resick, P.A. 446
residency training: for behavioral scientist 84; in IBC clinic 750
restless leg syndrome (RLS) 477
return on investment (ROI) 205
Reyes, M.A. 379
rheumatoid arthritis 588, 589
rheumatologic pain 673
Rhode Island Change Assessment Questionnaire (URICA) 403–4
Richmond Memorial Health Foundation 739
Rickels, K. 622
Rifkin, L.S. 263
Rish, J.M. 638, 646
risk assessment, crisis interventions and 220–1
risk factor, defined 300
Ritchie, R.P. 185
Robert Wood Johnson Foundation 70
Robin, D.A. 674
Robinson, D.S. 620
Robinson, P.J. 97–8, 219, 669–70, 741, 765, 778–9
ROC analysis see receiver operating characteristic (ROC) analysis
ROC curve see receiver operating characteristic (ROC) analysis
Rocquet, W. 517
Roegner, M. 377–8
Rogentine, D.S. 616
Rogers, E.M. 726
Rogers, K.D. 322–3
Rogers, R. 638
Role-Emotional (RE) domain scale 533
Role-Physical (RP) domain scale 533
Romera, I. 181
Rosen, A. 176
Rosenbaum, R. 620–1
Rosenberg, R. 611
Rosenberg, S.E. 603
Rost, K. 24–5
Roth, M.E. 10
Roy-Byrne, P.P. 612
Rudd, M.D. 301
Rudy, T.E. 413
Ruiz, M.A. 335
Runyan, C.N. 11, 22, 32
Russell, H. 649
Ruz, M.E. 617
Sadock, Elizabeth 753
safety net clinics 737, 746–7
safety net primary care, case study 747–9
St. Louis Mental Status Examination (SLUMS) 377–8
St. Louis University Mental Status examination (SLUMS) 288
Sajatovic, M. 377
Sals, J. 37
SAMHSA see Substance Abuse and Mental Health Services Administration (SAMHSA)
SAMHSA-HRSA Center for Integrated Health Solutions Model 20
Sampson, W.S. 486
Sander, J.W. 649
Sansone, L.A. 167
Sansone, R.A. 167
Saris-Baglama, R.N. 538
satisfaction surveys 36
Sato, A. 673–4
Saunders, B.E. 613
Saunders, W.B. 613
Scanlan, J. 381
Scanlon, J. 381
Schacht, B.D. 341
Schatz, M. 577, 578
Schedule of Recent Experiences 359
Schlegl, S. 619
Schleifer, S.J. 622
Schmidt, S.C. 649
Schoenbaum, M. 14
Schroll, M. 641
Schulberg, H.C. 181, 599
Schurman, R.A. 168
science competency cluster 103–4
SCOFF screening questionnaire 502, 514–15
Screener and Opioid Assessment of Patients with Pain-Reduced (SOAPP-R) 398
screening: abbreviated multidimensional measures 34; accuracy rates of 170–1; by behavioral health personnel 198–200; brief instruments 202; concept of 171–2; conditions for 200; condition-specific instruments 203; cutoff values 178–9; defined 171, 389; for depression 38–9, 167, 546; with disorder-specific instruments 38; effective 37–8; efficacy in 171–2; false negatives 38–9; false positives 38–9; guidelines for effective programs 172; integrated into primary care settings 37–8; integrating with routine check-in 200; methodological issues in 175–80; mini-battery instruments 88; for panic disorder 326; patient bias in reporting 175; positive findings 38; predictive validity 174; as preinterview data gathering 225–6; in primary care setting 85–6; psychological assessment as tool for 36–9; psychometric principles in 173–5; receiver operating characteristic analysis 178–80; role in referral 225; selecting instruments for 202–3; sequential 177–8; for social anxiety disorder 326; stigma of 435; for suicide ideation 297; see also clinical interview; psychological assessment; suicide risk screening; test-retest reliabilities; validity, screening Screening, Brief Interview, Referral to Treatment (SBIRT) model 390, 781 screening instruments: to assess safety issues 226; condition-specific 203 secondary insomnia 473 Sederer, L.I. 43 See et, S. 450 seizure disorders 649–50 Sekel, A.C. 18 self-assessment: of health status 35–6; psychological test instruments 36 self-compassion 704 self-harm ideation 305 self-manage, ability to 744 Self-rating Depression Scale (SDS) 10 self-report biases 201 Seligman, M.E. 675 Sellbom, M. 637–8, 639 Selyne, Hans 356 Semeijn, E.J. 376 Semi-Structured Assessment for Drug Dependence and Alcoholism (SSADDA) 399, 400–1 Semi-structured Assessment for the Genetics of Alcoholism (SSAGA) 400 sensitivity 172–3 Seo, J. 322, 325–6 sequential screening 151, 177–8, 178 serious mental illness (SMI), comorbidity of 10 Serrano, N. 763–4, 777 service delivery models 114–15 severe hypoglycemic episodes 508–9 severe mental illness (SMI) 450 Severity of Dependence Scale 324 Sevilla, J.L. 615 sexual abuse/victimization: chronic pain and 615–16; PTSD and 448; SCL-90-R/BSI and 615–16 SF-10 Health Survey for Children 548 SF-12v2 Health Survey: acute form 532; administration of 538–9; case example 557–9; collateral instruments 555; content-based interpretation 542; criterion-based interpretation 543; described 36; determining change 551–2; developing 529–30; health domain scales 533; implementation of 548–9; IVR technology 158; norm-based interpretation 542; patient feedback 549; problematic areas of functioning 543; psychometric properties 536–7; reliability of 536–7; scoring of 533–4; selection of 538; as short form of SF-36 532; standard form 532; translations of 539; validity of 537; see also Short Form Surveys SF-36v2 Health Survey: acute form 532; administration of 538–9; collateral instruments 555; comorbid conditions effect scores 554; construct validity 535; content-based interpretation 540; criterion-based interpretation 540–2; criterion validity 536; described 36, 529; determining
change 551–2; group-level aggregated data 539; health domain scales 533; implementation of 548–9; norm-based interpretation 539–40; for pain assessment 418, 584; patient feedback 549; problematic areas of functioning 541–2; psychometric properties 534–6; reliability of 534–5; scoring of 533–4; selection of 538; standard form 532; translations of 532, 539; validity of 535–6; see also Short Form Surveys

Shaffer, Carla 737
Shandera-Ochsner, A.L. 638
Shapiro, S. 169, 171
Shea, S.C. 215, 216
Sheehan, B. 381
shell shock symptoms 81
Shen, H. 451
Shidhaye, R. 168
Short and Sweet Screening Instrument (SASSI) 380
Short Form health surveys see SF-12v2 Health Survey; SF-36v2 Health Survey
Short Form Surveys: comorbid conditions 553–5; cost considerations 556; cultural considerations 556; norming study 553; outcomes assessment 550–2; patient acceptance 556; in primary care settings 555–6; staff acceptance 556; see also SF-12v2 Health Survey; SF-36v2 Health Survey
Short Memory Questionnaire (SMQ) 381
Short Test of Mental Status (STMS) 378, 381–2
S-I-G-E-C-A-P assessment 86
S-I-G-E-C-A-P assessment 86
S-I-G-E-C-A-P-S assessment 86
Sikimper, D.R. 171
Silver, R.C. 604
Simon, G.E. 169, 415
Simon, W. 326
single question screening 110, 393, 395
Sinsky, C. 666
Siu, C. 616–17
Siviks, D. 608
Six-Item Cognitive Impairment Test (6CIT) 381
Six-Item Screener (SIS) 380–1
skills-based competencies 106
Slade, M. 149
sleep apnea 588, 589
sleep behaviors 109
sleep deprivation 470, 482
sleep disorders: consequences of 474; electronic medical records (EMRs) 474; insomnia 449; integration barriers 468–9; prevalence rates of 467; types of 467; see also behavioral sleep medicine; sleep problems assessment
sleep education 481
sleep environment 476
sleep history 487, 490
sleep hygiene 109, 482
Sleepio 485
sleep log 482
sleep medicine 467; see also behavioral sleep medicine
sleep problems assessment: 3-P Model of Insomnia 475; in-bed behaviors 476; daytime functioning 476–7; history of problem 475; medical conditions/medications 477; self-report screening measures 477–80; sleep environment 476; sleep history 475–7
sleep-related disorders 672
Sleep Research Society (SRS) 470
sleep restriction 109, 482
sleep-wake patterns 481
Small, S.A. 149
Smallwood, R.F. 674
SMART goal setting 109
smartphone technology 157, 159
Smith, A.W. 359
Smith, C. 252
Smith, G.R. 550
Smith, J.L. 24–5, 442
Smith, R.D. 25, 613
Smith, S. 420
Smith, S.M. 537
Smith, S.R. 147, 152
SMOG Readability formula 148
Smyth, K.A. 377
Snow, K.K. 536
Snyder, D. 616
social anxiety disorder 326, 327, 338
social health domain 421
social history 231–2
social hygiene movement 59
Social Interaction Anxiety Scale (SIAS) 332
Social Phobia Diagnostic Questionnaire (SPDQ) 332
Social Readjustment Rating Scale 359
Society for Health Psychology 112
socioeconomic status 507–8
Sokhal, D. 768
Soloff, P.H. 620
Solomon, A. 441
Somatic/Cognitive Specific Problems Scales 637, 640–1
somatoform disorders 86, 168, 449, 611–12
Somoza, E. 179
Soubhi, H. 530
source of information 144–5, 150, 152, 226–8
Sousa, T.V. 323
Southwick, S.M. 450
Spanswick, C.C. 650
speciality level of care 197
speciality mental health settings 203
specificity 172–3
Sperry, L. 38
Spielman, A.J. 470
spine surgery/spinal cord stimulator evaluations 648
Spira, A. 650
Spiro, L. 280
Spitzer, R.L. 261, 320, 322, 323, 325, 336, 512–13
SSRIs: advent of 12; efficacy of 282–3
staffing ratio 669–70
Stafford, D. 12
standardized intake guide 232–4
standardized performance measures 373
Standards for Educational and Psychological Testing 31, 150
Standish, T.I. 378
Starfield, Barbara 62
Starr, Paul 62
State Innovation Models initiative 68
state psychiatric hospitals 58
State Trait Anxiety Inventory (STAI) 256, 332, 617
Steer, R.A. 614, 622
Stein, D.J. 430, 438
Steinwachs, D.M. 550
Stellato, C.P. 39
stepped-care model of treatment 484–6, 485
Stern, J.D. 616
Index

Stern, Y. 149
Sternbach, R.A. 411
Stevenson, M. 638
Stewart, A.L. 43
Stewart, J.W. 611, 620
stimulus control 109, 481–2
Stoddard, J. J. 16
Stone, A.A. 43, 531
STOP-Bang questionnaire 479–80
stress: acculturation-related 739; chronic illness and 358; as chronic variant of anxiety 613; concept of 356–9; Daily Stress Inventory 355, 360–1, 363–4; daily stressors 359–61; defined 356–7; diabetes mellitus and 358; emotion-focused coping 357; hassles 359; interventions targeting reduction of 355; maladaptive attempt to cope with 359; man-made disasters and 613; measurement of 359–63; Perceived Stress Scale (PSS) 359–60; physiological adaptations 357–8; physiological representation of 357; prevalence rates of 355; problem-focused coping 357; residual levels of 613; role of 231; Weekly Stress Inventory 361–4
stressors 356, 357, 739
Strobl, J. J. 282–3
Stroke Unit Mental Status Examination (SUMSE) 185
Stroahal, K. 12, 19, 182, 183
structure, quality of care 42
Structured Clinical Interview for DSM disorders 255, 282, 399, 415
Stuckey, J. C. 381
students’ perspectives: clinical skills 97; on ethics 98; expectations vs. reality 96–7; on integrated care model 97; on practice management 98
Stungeon, J. 638
subjective outcome measurement 201
Substance Abuse and Mental Health Services Administration (SAMHSA) 4, 6, 20–1, 392, 434
Substance Abuse and Mental Health Services Administration and Health Resources and Services Administration (SAMHSA-HRSA) 105–6
substance-induced anxiety disorder 327
substance misuse 109–10, 393–9
substance use disorders: BHL program 718; CAGE screen 396; cognitive impairment and 372; co-occurring prevalence rate 6; dependence percentages 5; detection and treatment of 13; diagnostic assessment measures for 399–404; pain reliever use disorder 392; persistent pain and 672; prevalence rates of 389; PTSD and 449–50; referrals to behavioral health consultants 221; SCL-90-R/BSI and 614–15; screening tools 393–4; see also behavioral health disorders
substance use screening: Alcohol Use Disorders Identification Test (AUDIT) 395–6; diagnostic assessments 399–400; limitation of 406; purpose of 390; role of 389–92; Single-Question Screening 395; strengths of 406; team-based care guidelines 404–5
subthreshold 85–6
Suchy, Y. 31
suicidal behavior: low base rates 183; risk factors profile 183; screening for 182–4
suicidal behavior disorder 304
suicidal ideation 183, 184, 226, 253, 748
suicidal mode 298, 299
suicidal thoughts 296
suicide: assessment 86, 746; attempts 183; causal risk factors 300; cognitive domain 298; completed 183; fluid vulnerability theory of 297–9; negative screen 301; prevalence rates of 182; primary care as last stop before 295; risk factors profile 300; triggering events 298
Suicide Cognitions Scale (SCS) 308–10
suicidal ideation 297, 305, 307
suicide prevention screening 304
suicide risk assessment: case study 312–13; cultural differences in 311–12; defined 302; patients’ perspectives on 311; provider-related factors 310–11; suicide risk screening vs. 301–2
suicide risk screening: acute risk 298; barriers to 310; baseline risk 297–8; Behavioral Health Measure-20 305–8; within depression screening 296; false negatives 301; implementing methods 310–12; implications for 299–300; indicated screening 296–7; instrument and tools 302–10; Patient Health Questionnaire (PHQ-9) 302–5; SCL-90-R/BSI and 613–14; Suicide Cognitions Scale 308–10; suicide risk assessment vs. 301–2; temporal fluctuations in 298
Sullivan, H. S. 215
Sullivan, M. D. 611
Sullivan, M. J. 414
Sumathipala, A. 168
Summerfieldt, L. J. 144
Suris, A. 442
surveys, patient 36
Sustaining Healthcare Across Integrated Care Efforts (SHAPE) 68
Swedo, S. E. 184
Sweet, J. J. 31
Swets, J. A. 179
sympatho-adrenomedullary (SAM) systems 357
Symptom Checklist-90 (SCL-90-R): Analogue Scale and 616; validity of 606–7; see also Brief Symptom Inventory (BSI)
Symptom Driven Diagnostic System for Primary Care (SDDSP) 171
Symptom Validity (FBS-r) scale 638
systems cluster 104
Talen, M. R. 80
Tarescavage, A. M. 638, 639, 642, 644, 647, 650, 651
Tarsitani, L. 640
Taylor, S. E. 181, 413, 599
Taylor, G. J. 673–4
Tatsuoka, C. 377
TCAs 282–3
team-based learning (TBL) curriculum 750–1
team-based models of care 200, 201
Index 805

teammate compassion 703
technology 152–3; see also assessment technology
Teesen, M. 154
telehealth 155–6
Tellegen, A. 417, 633, 640–1
“Tell three stories about yourself” exercise 700
Terman, J.S. 611
Terman, M. 611
Terrill, A.L. 324
test information functions 568–9
test-retest reliabilities: of ACT 579–80; of BDI-FS 249; of BDI-II 248; of BHM-20 306; of BSI 604; of CAPS-5 444–5; of CES-D 254; of CESD-R 256–8; of CRAFFT 398–9; of DSI 360–3; of EDE-Q 511; of GAD-7 322–3; of GDS 278–80; of GSAQ 478; of HFS-II 520; of HIT-6 571; of ISI 479; of PCL 441; of PHQ-9 262; of PMQ 420; of PRIME-MD 518; of PRISM 403; of PSQI 477; of SCL-90-R 604; of SCOFF 514; of SF-12v2 536–7; of SF-36v2 534–5; of SMQ 381; of SSADDA 401
Thai Mini Neuropsychiatric Interview (MINI) 401
therapeutic assessment 146
Thibodeau, M.A. 336
Thomas, M.L. 638, 649
Thombs, B.D. 167
Thompson, L.W. 608
Thompson, R. 12
Thorazine 60
Thornicroft, G. 149
Thought Dysfunction (THD) 363, 640
Three Word Recall Test (3WR) 378
Thyer, B.A. 612
TIDES see Translating Interventions for Depression in Effective Solutions (TIDES)
Tien, A. 252
time orientation 153
Titchener, J.L. 613
tobacco products: chronic pain and 685; stress and 359
Tong, X. 323, 326
Toomey, T.C. 615
Touradji, P. 149
Tovian, S.M. 31
TRACE (Tolerance, Annoyance, Cut-down, Eye-opener) 396
training see education
transactional model 357
transgender youth 616
Translating Interventions for Depression in Effective Solutions (TIDES) 710–11, 717–18, 719–20
trauma: education deficit in recognizing symptoms 436; impact of 429; impact on physical health 433–4; see also posttraumatic stress disorder (PTSD)
trauma assessment: barriers for patients with 435; implementation of 434–7; population-based, stepped approach 434–5; time constraints 435–6
trauma screening: false positives 440–1; Primary Care Checklist (PCL) 439–42; Primary Care-PTSD Screen (PC-PTSD) 438–9; selection of 438
traumatic brain injury (TBI) 638
traumatic events 86
Treatment Improvement Protocol (TIP) 434
treatment monitoring see treatment planning and monitoring
treatment planning and monitoring: abbreviated multidimensional measures 34; baseline data 41; Beck Depression Inventory-Second Edition (BDI-II) 269; defined 200–1; determining significant change 41–2; for geriatric depression 286; instrumentation for 152; measurement feedback systems 201; measuring change 203–5; multidimensional 234–7; for problem identification 40; process of 41; progress adjustments 41; psychological assessment as tool for 39–40
treatment progress 41
treatment-resistant hypertension (TRH) 537
treatment team 226
tricyclic antidepressant (TCA) 620
Trimble, M. 649
Triple Aim 19, 22, 55, 666; see also Quadruple Aim
Truaax, P. 552, 619
Trude, S. 16
ture negatives, screening 173
ture positives, screening 173
Tryon, R.C. 606
Tsacoumis, S. 157
Tsuji, T. 537
Tuijl, J.P. 381
Tumosa, N. 377–8
Turck, J.A. 83
Turk, D.C. 413, 650
Turner, T. 615
Turner-Bowker, D.M. 538
TWEAK (Tolerance, Worried, Eye-opener, Amnesia, K/Cut Down) 394, 396
two-factor model of sleep 469
unabridged CBT-I 486
unbearability 308
Underwood, L. 751
unitary model 336
University of Massachusetts Medical School Center for Integrated care 114
University of Wisconsin Health Psychology Post-Doctoral Fellowship 769
University of Wisconsin Pediatrics PCBH program 779
unsolvability 308
unstructured assessment 215–16
Uützer, J. 14, 16, 187
Uplifts Scale 359
Urbina, S. 143, 147
Ureno, G. 603
US Department of Veterans Affairs National Center for PTSD 431, 438, 439–40
US Food and Drug Administration (FDA) 530
utility, screening 173
Uttley, L. 675
Valeras, A.B. 80
validity, screening: for HIT-6 571–2; of MMPI-2-RF 638; programmatic nature of 605; of SCL-90-R 606–7; for SF-12v2 537; for SF-36v2 535–6; specificity of 605
Van Dam, D. 450
Van Dam, N.T. 256
Vanderbilt Assessment for Children 756
Vasiliadis, H. 326
VCUHS Pediatrics Clinic 740
Vecchio, T.J. 176
Velayudhan, L. 377, 382
vertical team model of training 745–6
Veterans Health Administration (VHA): as accountable care organization 709; barriers to integrated care implementation 722–6; behavioral health lab 717–19; care management plan 710–11; collaborative care at 710; colocated collaborative care 711, 720–2; co-location program 711; early implementations of psychological assessment 81; Health Promotion/Disease Prevention (HPDP) program 712; history of 710–12; integrated care programs 711–12; kiosks 725; measurement-based care efforts 713, 720–2; mental illness prevalence in population served by 709; Office of Patient Centered Care and Cultural Transformation 712; integrated care programs 711–12; kiosks 725; measurement-based care efforts 713, 720–2; mental illness prevalence in population served by 709; Office of Patient Centered Care and Cultural Transformation 712; Patient Centered Medical Home model 712; personalized, proactive patient directed care (PPPD) approach to care 712; population-based screening initiatives 715–16; posttraumatic stress disorder and 431, 438, 613, 622; Primary Care-Mental Health Integration initiative 711–12; staffing and software support 717; standardized screening instruments 715; see also measurement-based care (MBC); Primary Care-Mental Health Integration (PCMHI)
Veterans Integrated Service Network (VISN) 717
Vietri, J. 537, 554
Vilagut, G. 255
Villasenor, V.S. 603
Vincent, N. 487
Virginia Commonwealth University (VCU) Psychology Department 733, 737, 749; see also Primary Care-Mental Health Integration (PCMHI)
Visual Analog Scale (VAS) 416–17
Vitaliano, P. 381
Vitality (VT) domain scale 533
Voderholzer, U. 357
Vodermaier, A. 616–17
Vogel, M.E. 106
Vollrath, M. 611
Von der Lieth, L. 641
Von Korff, M. 169, 415
Vries, R.M. 604
Wagenaar, H. 529
Waggoner, C.D. 360
Wagley, J. 486
Walker, E.A. 615
Wallace, P. 751
Walsh, K. 487
Walters, S.J. 531
Wan, E.Y. 554
Wang, J. 614
Ward, M.F. 375–6
Ware, J.E., Jr. 43, 531, 535, 536, 547
Warman, M.K. 23
warm handoffs 87, 204, 219–20, 223, 710–11, 765; see also referrals
war-zone PTSD scale (WZ-PTSD) 613
Watson, C. 640
Watson, D. 612, 633
Weathers, F.W. 438, 441, 444, 445, 613
web-based mental health interventions 485
Wechsler Abbreviated Scale of Intelligence II-Second Edition 88
Wechsler Memory Scale-IV (WMS-IV) 369
Weekly Stress Inventory (WSI) 355, 361–4
Weekly Stress Inventory-Short Form (WSI-SF) 363
Weidner, G. 616
Weiss, D.S. 620–1
Weissman, M.M. 167, 599–600, 611
Wells, K.B. 10
Wender Utah Rating Scale (WURS) 375–6
Wersha, R. 638
Werthman, M.J. 35
Westerfeld, J.S. 184
Wetzler, H.P. 531
Wetzler, S. 611
Wharton, R.M. 379
Whelton, P.K. 182
White, M.B. 79
Whitehead, L. 538
Whitehouse, A. 381
White River Junction model 711, 721
Whitlock, E.P. 392
Whittington, C. 751
Whittington, J. 666
whole health care 69, 70
whole-person orientation 22, 64
Wide Range Achievement Test-Fourth Edition 88
Wiesner, M. 608
Wiggins, J.S. 606
Wilber, S.J. 392
Wilkins, K.C. 441
Williams, J.W. 39, 171, 261, 617
Williams, N.J. 468
Williams, P. 179
Wilner, N. 613
Wilson, K. 282–3
Windover, A. 647
Wing, R.R. 359
Winje, D. 613
Winokur, A. 622
Wisdom, J.P. 392
Wise, M.G. 181, 599
Wise, T.N. 180–1, 599
Wiznitzer, M. 606
Wolff, H.G. 357
women: borderline personality disorder and 447–8; suicide attempts 183; see also gender/ gender-keying
Wong, C.J. 37, 38, 39, 151
Woodman, C.L. 620
Woods, J. 14
Woodworth, Robert 180
Word Memory Test 638
workforce development 771–2
work productivity 24–5
work Relative Value Units (wRVUs) 724
World Health Organization 18, 172, 183, 245, 373, 395, 415
World Health Organization Disability Assessment Scale (WHO-DAS II) 324
World Health Organization Quality of Life Scale (WHOQOL) 555
Wygant, D.B. 638, 642, 647
Yager, J. 170
Yano, E.M. 169, 181
Ybarra, M. 252, 258
Yelin, E. 169
Yesavage, Jerome 277, 278
Young, J.L. 621
Young Adult Self-Report (YASR) 606
Youngjohn, J.R. 638

Zabora, J.R. 608
Zargar, F. 617
Zarro, V.J. 182
Zeidler Schreiter, E. 768

Zeng, Q. 322
Zenger, M. 249
Zhong, Q. 322, 323, 325
Zhuang, G. 535–6
Zimmerman, M. 38, 253
Zung, W.K. 10, 170–1
Zung, W.W. 171
Zung Self-Rating Depression Scale (SDS) 148, 170–1
Zyolensky, M.J. 336