



Measuring the Quality of Care for Psychological Health Conditions in the Military Health System

Candidate Quality Measures for Posttraumatic
Stress Disorder and Major Depressive Disorder

Kimberly A. Hepner, Carol P. Roth, Coreen Farris, Elizabeth M. Sloss,
Grant R. Martsof, Harold Alan Pincus, Katherine E. Watkins, Caroline Epley,
Daniel Mandel, Susan D. Hosek, Carrie M. Farmer

For more information on this publication, visit www.rand.org/t/RR464

Library of Congress Cataloging-in-Publication Data is available for this publication.

ISBN: 978-0-8330-8656-3

Published by the RAND Corporation, Santa Monica, Calif.

© Copyright 2015 RAND Corporation

RAND® is a registered trademark.

Limited Print and Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law. This representation of RAND intellectual property is provided for noncommercial use only. Unauthorized posting of this publication online is prohibited. Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Permission is required from RAND to reproduce, or reuse in another form, any of its research documents for commercial use. For information on reprint and linking permissions, please visit www.rand.org/pubs/permissions.html.

The RAND Corporation is a research organization that develops solutions to public policy challenges to help make communities throughout the world safer and more secure, healthier and more prosperous. RAND is nonprofit, nonpartisan, and committed to the public interest.

RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

Support RAND

Make a tax-deductible charitable contribution at
www.rand.org/giving/contribute

www.rand.org

Preface

Providing high-quality treatment and improving outcomes for individuals with psychological health (PH) conditions is a high priority for the military health system (MHS). In the past decade, the number of individuals treated by the MHS for these conditions has grown significantly. However, the extent to which the MHS is currently providing care that is consistent with evidence-based clinical practice guidelines or MHS standards for high-quality care has been difficult to measure across the entire system. There is currently no U.S. Department of Defense-wide system in place to evaluate the quality of PH care provided, assess whether the care is improving outcomes, or identify potential areas for improvement.

To better understand these issues, the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (DCoE) asked the RAND Corporation to develop a framework to help identify and classify a set of measures for monitoring the quality of care provided by the MHS for PH conditions, specifically posttraumatic stress disorder (PTSD) and major depressive disorder (MDD). The goal of the current effort is to select and develop quality measures for PTSD and MDD treatment relevant to the MHS. The task does not include a complete implementation plan; we designed it to provide the foundation for future DCoE–RAND work to pilot and later implement a subset of these measures.

The purpose of this document is to describe a candidate set of quality measures for PTSD and MDD, including the methods used to identify and refine them. This measure set should inform ongoing efforts within DCoE to measure and improve the quality of PH care. For each measure, we provide a brief conceptual description, including the rationale for selecting the measure, the population to which it applies, and what data source might be available for implementing these measures in the MHS. This report should be useful to MHS personnel who provide care for individuals being treated for PTSD or MDD and those responsible for monitoring the quality of that care in focusing their efforts on evidence-based quality measures.

This research was sponsored by DCoE and conducted within the Forces and Resources Policy Center of the RAND National Defense Research Institute, a federally funded research and development center sponsored by the Office of the Secretary of Defense, the Joint Staff, the Unified Combatant Commands, the Navy, the Marine Corps, the defense agencies, and the defense Intelligence Community.

For more information on the RAND Forces and Resources Policy Center, see <http://www.rand.org/nsrd/ndri/centers/frp.html> or contact the director (contact information is provided on the web page).

Contents

Preface	iii
Figures and Tables	vii
Summary	ix
Acknowledgments	xvii
Abbreviations	xix
CHAPTER ONE	
Introduction	1
Defining and Measuring Quality of Care	2
How Quality Measures Are Used	2
Quality Improvement	3
Accountability	3
Research	4
Examples of Improved Performance over Time	4
Overview of Our Process for Quality Measure Development	4
Organization of the Report	5
CHAPTER TWO	
Framework for Classifying Measures of Psychological Health Conditions	7
Dimensions for the Framework	7
Care Continuum	8
Measure Type	9
Matrix Framework for Classifying Psychological Health Measures	11
CHAPTER THREE	
Identification of Measures of Quality of Care for Posttraumatic Stress Disorder and Major Depressive Disorder	13
Measure Selection by Expert Consensus	14
CHAPTER FOUR	
Overview of Candidate Sets of Measures of Care Quality for Posttraumatic Stress Disorder and Major Depressive Disorder	17
Description of Candidate Measures, by Care Continuum Phase	21
Description of Candidate Measures, by Measure Type	23
Description of Candidate Measures Using the Matrix Framework	23
Examination of Need for New Measures	24

Prevention	24
Reintegration	25
Structure Measures	25
CHAPTER FIVE	
Summary and Recommendations	27
Study Limitations	27
Recommendations	28
APPENDIXES	
A. Definitions of Terms Used to Describe Quality Measures	31
B. Descriptions of Measures of Quality of Care for Posttraumatic Stress Disorder	37
C. Descriptions of Measures of Quality of Care for Major Depressive Disorder	103
D. Additional Rationale for Cost Measure	165
E. Candidate Set of Measures of Quality of Care for Posttraumatic Stress Disorder	179
F. Candidate Set of Measures of Quality of Care for Major Depressive Disorder	187
G. Sources of Measures	195
Bibliography	199

Figures and Tables

Figures

S.1.	Questions Addressed by Measures in the Care Continuum.....	xi
S.2.	Questions Addressed by Each Measure Type	xii
2.1.	Questions Addressed by Measures in the Care Continuum.....	9
2.2.	Questions Addressed by Each Measure Type	11

Tables

S.1.	Posttraumatic Stress Disorder and Major Depressive Disorder Measures, by Care Continuum and Type of Measure	xiii
1.1.	Six Aims for Improving Quality of Care, from the Institute of Medicine.....	2
2.1.	Continuum-of-Care Phases.....	8
2.2.	Framework for Classifying Measures for Psychological Health Conditions.....	11
3.1.	National Quality Forum’s Measurement Evaluation Criteria	16
4.1.	Candidate Set of Measures of Quality of Care for Posttraumatic Stress Disorder.....	18
4.2.	Candidate Set of Measures of Quality of Care for Major Depressive Disorder	19
4.3.	Quality Measures for Posttraumatic Stress Disorder and Major Depressive Disorder, by Care Continuum	21
4.4.	Measures of Quality of Care for Posttraumatic Stress Disorder and Major Depressive Disorder, by Measure Type.....	23
4.5.	Measures of Quality of Care for Posttraumatic Stress Disorder and Major Depressive Disorder, by Care Continuum and Measure Type	24
D.1.	TRICARE and Other Insurance Coverage and the System of Care	173
E.1.	Candidate Set of Measures of Quality for Posttraumatic Stress Disorder	180
F.1.	Candidate Set of Measures of Quality of Care for Major Depressive Disorder	188

Summary

In recent years, the number of U.S. service members treated for psychological health (PH) conditions has increased substantially. In particular, at least two PH conditions, posttraumatic stress disorder (PTSD) and major depressive disorder (MDD), have prevalence estimates ranging from 4 to 20 percent for PTSD and 5 to 37 percent for MDD (Institute of Medicine [IOM], 2013, p. 21; Ramchand, Schell, et al., 2010; Schell and Marshall, 2008). Delivering quality care to service members with these conditions is a high-priority goal for the military health system (MHS). Meeting this goal requires understanding the extent to which the care the MHS provides is consistent with evidence-based clinical practice guidelines and its own standards for quality. However, there is currently no MHS-wide system in place to evaluate quality of care for PH conditions, to assess whether the care is improving patient outcomes, or to identify potential areas for improvement.

Purpose and Approach

In order to better understand these issues, the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (DCoE) asked a team of researchers from the RAND Corporation to develop a framework to identify and classify a set of measures for monitoring the quality of care provided by the MHS for two prevalent PH conditions: PTSD and MDD. The goal of this project was to identify, develop, and describe a set of candidate quality measures to assess care for PTSD and MDD.

To accomplish this goal, we performed two tasks:

- We developed a conceptual framework for assessing the quality of care for PH conditions.
- We identified a candidate set of measures for monitoring, assessing, and improving the quality of care for two high-priority PH conditions in the MHS: PTSD and MDD.

This document describes our research approach and the candidate measure sets for PTSD and MDD that we identified. The current task did not include implementation planning but rather was designed to provide the foundation for future DCoE–RAND work to pilot and later implement a subset of these measures.

How Quality Measures Are Used

Quality measures, also called performance measures, indicators, or metrics, provide a way to measure how well health care is being delivered. Such measures provide information about the health care system and highlight areas in which providers can take action to make health care more effective, safe, efficient, and equitable (National Quality Forum [NQF], undated [c]). Quality measure scores are typically presented as the percentage of eligible patients who received recommended care (e.g., percentage of new PTSD patients screened for depression).

According to the Agency for Healthcare Research and Quality (AHRQ) (undated [b]), organizations generally use quality measures for one or more of these three purposes:

- **quality improvement (QI).** Health care organizations can focus on internal or external QI. Internal QI programs measure the quality of care within a health care organization or system. External QI programs measure quality in several health care organizations and compare performance across those organizations (AHRQ, undated [b]).
- **accountability.** An organization can use measures for accountability in any of three ways: public reporting, performance-based payment, and professional certification.
- **research.** Quality measures are used in research studies to measure how frequently evidence-based care is provided to patients and whether care differs across patient subgroups and health care settings, as well as to measure the patient response to interventions.

We illustrate the impact of quality measurement with an example of how care for mental health conditions in commercial health plans has improved steadily for more than a decade, based on two Healthcare Effectiveness Data and Information Set (HEDIS®) quality measures (National Committee for Quality Assurance, 2013a).

Results

Framework for Identifying Candidate Measures

The RAND team conducted a review of existing quality measures and identified 530 measures relevant to PTSD, MDD, or other PH conditions. A group of RAND clinicians and experts in mental health quality measurement reviewed the comprehensive list of quality measures and selected candidate measures based on the extent to which each of the following was true:

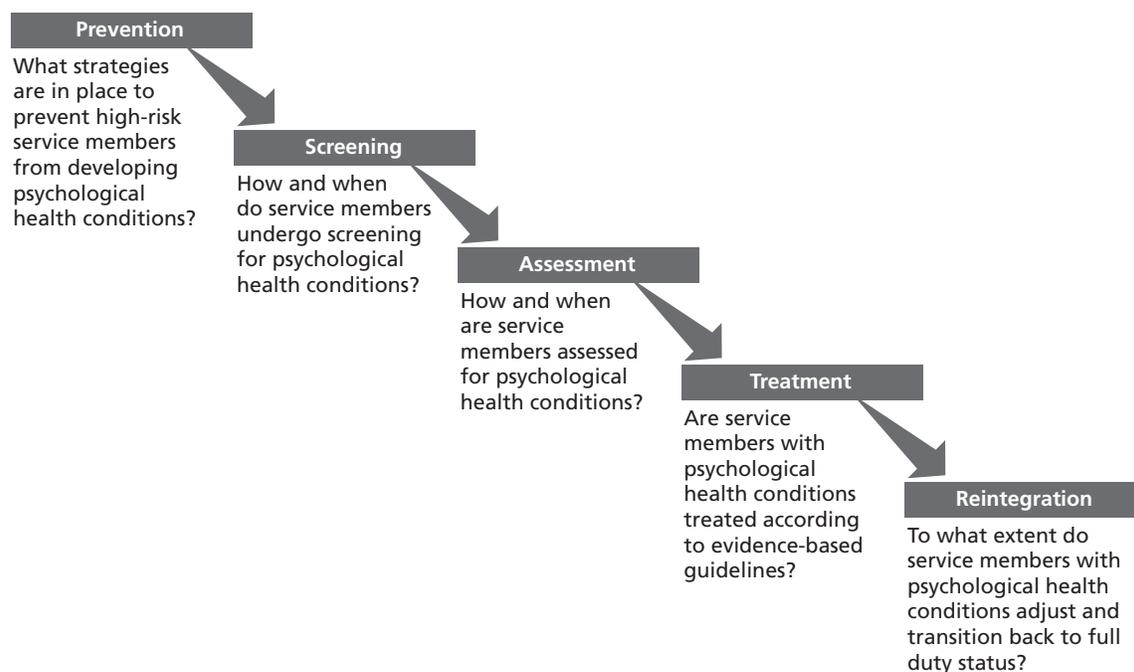
- Adequate scientific evidence or professional consensus exists to support a link between the performance of care specified by the measure and health benefits to patients with PTSD or MDD. This included support from existing clinical practice guidelines, either from a professional organization (e.g., American Psychiatric Association) or U.S. Department of Veterans Affairs (VA) and U.S. Department of Defense (DoD) (validity).
- We would consider a provider or organization with significantly higher rates of adherence to the measure a higher-quality provider or organization (validity).
- The measure had been previously applied to military populations or veterans (feasibility of MHS implementation).

- Either the measure applies to many individuals or there would be serious adverse consequences from not adhering to the indicator (importance).
- There is variation among providers in the provision of the specified care indicating a quality gap and room for performance improvement (importance).
- The information necessary to determine adherence within the MHS is likely to be available in administrative data, in the medical record, or via patient (or family) interview (feasibility of MHS implementation).
- NQF had endorsed the measure. We used that endorsement as a proxy for importance, validity, reliability, feasibility, and usability.

Team members assessed the measures individually by providing a global rating that weighed these criteria, and then the group convened to discuss global ratings and identify the candidate measure sets. In the group discussion, we placed particular emphasis on identifying measures that could be targets for further refinement and field-testing within military treatment facilities (MTFs). DCoE and RAND will complete field-testing the measures during ongoing efforts. We will select measures for field-testing according to (1) high-priority areas, which DCoE will determine, and (2) the feasibility of implementing the measure using existing data sources.

We developed a two-dimensional framework as a tool for gap analysis—that is, to assist us in evaluating whether the existing measures from the environmental scan address important aspects of care for a particular PH condition. We considered multiple dimensions and decided on two. The first dimension relates to the *continuum of care*, which includes five phases that

Figure S.1
Questions Addressed by Measures in the Care Continuum



the DCoE portfolio of research has highlighted (as shown in Figure S.1): *prevention, screening, assessment, treatment, and reintegration.*

The second dimension is the *type* of measure. Figure S.2 shows the five types, which we adapted from the Donabedian model of health care quality (Donabedian, 2003). *Structure measures* assess the settings and available resources in which patients receive health care and the capacity to provide care. *Process measures* assess whether a recommended care process or event takes place during care or the degree to which a procedure or treatment is provided in a manner that reflects fidelity to the evidence base supporting its use. *Outcome measures* describe the outcome of care in terms of patient improvement, recovery, restoration of function, avoidance of deterioration, or survival. *Patient experience measures* assess patients’ perceptions of their providers, the care patients receive, and their health outcomes. *Resource use measures* relate to the resources expended during the care for a patient with a particular condition.

Using these two dimensions, we created a matrix for characterizing the candidate measures (see Table S.1). This matrix serves to highlight areas in which promising candidate measures exist, as well as areas in which measures are sparse or missing completely.

Candidate Measure Sets

We identified 58 candidate quality measures—29 each for PTSD and MDD—and placed these across the care continuum. By design, to provide adequate choice during ongoing efforts, which will include evaluation of measure feasibility, utility, and functionality across MHS settings, we selected a large set of candidate measures. Ten measures in the PTSD set are PTSD-specific, and ten measures in the MDD set are MDD-specific. Nineteen of the measures in each set could apply to care for patients with a range of PH conditions, including PTSD and MDD. As shown in Table S.1, these measures clustered heavily in the treatment phase for both conditions. The most common measures were process measures. The measures most commonly appropriate for use across the entire care continuum were either measures of patient experience or resource use.

Figure S.2
Questions Addressed by Each Measure Type

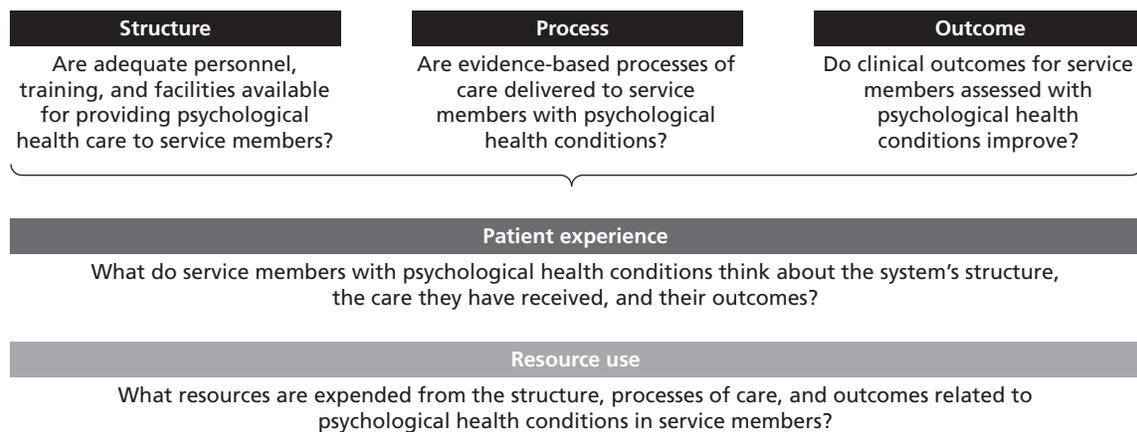


Table S.1
Posttraumatic Stress Disorder and Major Depressive Disorder Measures, by Care Continuum and Type of Measure

Measure Type	Care Continuum											
	Prevention		Screening		Assessment		Treatment		Reintegration		Total	
	PTSD	MDD	PTSD	MDD	PTSD	MDD	PTSD	MDD	PTSD	MDD	PTSD	MDD
Structure	0	0	0	0	0	0	2	2	0	0	2	2
Process	0	0	1	1	4	4	9	9	0	0	14	14
Outcome	0	0	0	0	0	0	5	5	0	0	5	5
Patient experience											6	6
Resource use											2	2
Total											29	29

Examination of Need for New Measures

Our analysis identified two care continuum phases in which there are no existing measures: prevention and reintegration. For the prevention phase of the care continuum, we noted that providers in primary care and specialty mental health care are rarely tasked with mental health prevention. We therefore chose to prioritize components of quality care that are more often the responsibility primarily of providers operating in a treatment setting. For the reintegration phase, we considered but did not select two reintegration measures because it may be particularly challenging for these measures to accurately capture the effect of appropriate reintegration practices rather than differences in baseline severity or time away from duty to receive appropriate treatment. Finally, we noted that there was no existing structure measure to assess the availability of inpatient care and created a new measure to assess inpatient availability.

Our goal was to develop quality measures to assess care provided within MHS treatment settings (e.g., primary care, specialty mental health care). These treatment settings are not tasked with prevention of mental health conditions, which usually occurs at a population level.¹ Typically, service members accessing treatment have already developed PH symptoms. However, the health care system could implement some best practices in prevention (e.g., avoiding immediate mental health services for trauma-exposed individuals *without* acute stress symptoms) (Roberts et al., 2009; Rose et al., 2002). Our belief is that such an event would occur rarely at the level of care for which we are providing recommendations and therefore is not an initial priority area. Instead, we focus on components of quality care that are the responsibility primarily of providers operating in a treatment setting. As administrators and policymakers seek to expand the scope of quality assessment, measurement of the quality of prevention services may become a priority.

¹ Indeed, we acknowledge that major efforts are under way across DoD to address prevention of mental health problems by building resilience. However, developing measures to assess these efforts was beyond the scope of this project.

Study Limitations

The report describes a review of quality measures to identify high-priority measure sets to assess care for PTSD and MDD. The body of this report describes an approach and gives an overview of measure sets, while the appendixes provide conceptual descriptions for each measure. Our approach to developing these measure sets has some limitations. First, we relied on an internal team of RAND experts in mental health and quality measurement to select candidate measures from among multiple measures representing the same measure concept. As part of this process, the experts prioritized measures primarily based on retaining measures that were developed and tested using rigorous methods. In some cases, they also considered whether there was adequate scientific evidence or professional consensus to support a link between the care specified by the measure and the health benefits to patients with PTSD or depression. Although it may have been improved by a formal expert panel process, this process does not require a formal expert panel. Second, we focused on identifying high-priority measure sets that would be relevant for service members; additional measures could be useful to assess care for retirees and family members. Third, complete technical specifications were not readily available for each of the 530 measures, so the review of existing measures focused on available information, the level of detail of which varied across measures. Another limitation of the study is that clinical practice guidelines typically provide general guidance and, even when tailored to individual clinical characteristics, cannot anticipate individual patient preferences (Montori, Brito, and Murad, 2013). Moreover, when developing measures, most developers' accessible data sources do not reliably contain systematic information about patient preferences related to guideline-based recommendations. This limitation applies to all quality measures. Although some efforts have tried to incorporate patient preferences into quality measures (e.g., measures of goal setting and attainment in setting a treatment plan), we rarely identified such measures specifically targeted toward MDD and PTSD.

In addition, although we relied on one framework for categorizing measures, consideration of other frameworks in future work could be helpful. Under the NQF project on Patient Outcomes Measures (NQF, 2011a), a group of mental health experts developed a measurement framework for mental health and substance use. The framework encompasses five characteristics the steering committee considered to be important aspects of measuring the quality of mental health care: inclusion of mental health in broad, cross-cutting measures; consumer, patient, family, and caregiver satisfaction; promotion of healthy behaviors and environment; nontraditional measures (e.g., homelessness); and accountability and care coordination. Alternatively, the IOM aims (IOM, 2001) for improving quality of care—safety, effectiveness, patient-centeredness, timeliness, efficiency, and equitability—might be utilized to assess the completeness of the current set of candidate measures and the direction for future expansion.

Also, the measures have not been specifically assessed for adaptability and feasibility across the range of MHS settings. The measure descriptions provided in Appendixes B and C represent an initial effort to summarize important aspects of the measures and basic operational details that are important to implementation. They do not represent a technical manual but rather provide a foundation for understanding each measure's content and evaluating its potential value as a means for evaluating behavioral health care. This document is not intended as a guide for implementing and using the listed quality measures but rather describes those foundational first steps that must be completed prior to doing so. The implementation of quality measures is a complex and separate process, and many contextual factors, such as costs,

organizational policy and culture, readiness for implementation, and the availability of relevant data influence it (Damschroder et al., 2009; Nicholas et al., 2001). Planning for implementation also includes identification of senior and local clinical champions, solicitation of staff and provider input, provider education, pilot-testing, and evaluation. Developing an implementation strategy will be part of future DCoE–RAND efforts. Once a plan is developed and measures for implementation are selected, the detailed technical specifications to define the application of those measures in that particular setting would be defined while maintaining the basic definitional integrity of the measures as described here.

Recommendations

We offer three recommendations for further work with respect to measures in support of DoD’s continuing efforts to improve quality of care for PH conditions:

Recommendation 1: Select a subset of high-priority, feasible quality measures to pilot-test and implement. This report includes 58 candidate measures to assess care for PTSD and MDD. We do not expect that it will be appropriate or feasible to implement all 58 of these measures. Instead, we present a candidate set of measures to consider for future implementation. For each measure, we have estimated the feasibility, based on our assessment of currently available data sources, of using the measure to assess quality of care. In addition, we acknowledge variability in how care is structured and delivered across service branches and that the delivery of PH care within DoD is constantly evolving, which will affect which measures DoD deems high priority. DoD should consider focusing initial quality measurement efforts on feasible measures that assess high-priority aspects of care at the time of implementation.

Recommendation 2: Consider structured documentation in the medical record to facilitate capturing the data necessary to compute quality measures. Many candidate quality measures rely on important clinical details that are typically documented in the medical record. Although quality measures based on administrative data are typically more feasible to implement (e.g., whether a patient received the recommended number of psychotherapy visits), these data typically do not capture many important aspects of the process of care (e.g., whether the psychotherapy delivered was evidence-based). Integrating structured documentation into the medical record is one method to potentially increase the breadth and value of data that systems routinely capture. The Veterans Health Administration (VHA) is currently pilot-testing these kinds of structured chart notes (Karlin and Cross, 2014). In addition, DCoE is currently developing an Alternative Input Method (AIM) form to capture adherence to a PTSD clinical pathway. Although structured documentation increases feasibility of measure implementation and potential data capture, it has important limitations. It is currently unknown how clinicians use structured documentation and whether these approaches accurately capture the care delivered or are tightly related to clinical outcomes. Natural-language processing (NLP) is another emerging technology that seeks to provide a means for higher-quality data collection without manual medical record review (Shiner et al., 2012). Pilot-testing and evaluating new approaches to structured documentation and data retrieval are important areas of future research.

Recommendation 3: Develop a process for ongoing assessment of quality of care. A single effort to assess quality of care will be of limited use. Instead, assessing the quality of care for PH conditions should occur as part of ongoing QI efforts, through a continuous feedback

loop. As DoD continues to pursue strategies to assess the quality of care, it is essential to consider how these efforts will be sustained and what resources will be required to support ongoing quality measurement.

Acknowledgments

We gratefully acknowledge the support of our project sponsor, Kate McGraw, and staff at the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, especially CDR Angela Williams-Steele and LTC Philip Holcombe. We appreciate the valuable insights we received from Charles Engel, Marcela Horvitz-Lennon, and Daniel Kivlahan. We addressed their constructive critiques as part of RAND's rigorous quality assurance process to improve the quality of this report. We also thank Anna Smith and David Adamson for their assistance in preparation of this report.

Abbreviations

AHRQ	Agency for Healthcare Research and Quality
AIM	Alternative Input Method
APG	Ambulatory Patient Group
AUC	area under the curve
AUDIT	Alcohol Use Disorders Identification Test
AUDIT-C	Alcohol Use Disorders Identification Test–Consumption
BDI	Beck Depression Inventory
BDI-II	Beck Depression Inventory II
BRFSS	Behavioral Risk Factor Surveillance System
BSDS	Bipolar Spectrum Diagnostic Scale
CAGE	cut down, annoyed, guilty, and eye-opener
CAHPS	Consumer Assessment of Healthcare Providers and Systems
CAPS	Clinician-Administered PTSD Scale
CBT	cognitive behavioral therapy
CDC	Centers for Disease Control and Prevention
CESD	Center for Epidemiologic Studies Depression Scale
CGI-BP	Clinical Global Impressions Scale for use in bipolar illness
CMS	Centers for Medicare and Medicaid Services
CPG	clinical practice guideline
CPT	Current Procedural Terminology
DCoE	Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury
DEERS	Defense Enrollment Eligibility Reporting System
DEQ	Distressing Events Questionnaire

DoD	U.S. Department of Defense
DSM-IV	<i>Diagnostic and Statistical Manual of Mental Disorders</i> , 4th ed.
DTS	Davidson Trauma Scale
DUKE-AD	Duke Anxiety–Depression Scale
ECHO	Experience of Care and Health Outcomes
EHR	electronic health record
EMDR	eye movement desensitization and reprocessing
EQ-5D	European Quality of Life, five dimensions
GDS	Geriatric Depression Scale
HEDIS	Healthcare Effectiveness Data and Information Set
HOS	Health Outcomes Survey
HRQOL	health-related quality of life
HRQOL-4	four-item health-related quality-of-life instrument
HRSD-17	Hamilton Rating Scale for Depression, 17-Item
ICC	intraclass correlation
IDS-SR ₃₀	self-rated 30-item version of the Inventory of Depressive Symptomatology
IES-R	Impact of Event Scale—Revised
IOM	Institute of Medicine
IPT	interpersonal psychotherapy
LASC	Los Angeles Symptom Checklist
M-3	My Mood Monitor
MADRS	Montgomery–Åsberg Depression Rating Scale
MDD	major depressive disorder
MDR	Military Health System Data Repository
MHS	military health system
MOSDQ	Medical Outcomes Study Depression Questionnaire
MPSS-SR	Modified PTSD Symptom Scale
M-PTSD	Mississippi Scale for Combat-Related PTSD
M-PTSD-DS	Mississippi Scale for Combat-Related PTSD for Desert Storm War Zone Personnel

MTF	military treatment facility
NCQA	National Committee for Quality Assurance
NHANES	National Health and Nutrition Examination Survey
NQF	National Quality Forum
NQMC	National Quality Measures Clearinghouse
PCL	PTSD Checklist
PCL-C	PTSD Checklist—Civilian Version
PCL-M	PTSD Checklist—Military Version
PCL-S	PTSD Checklist—Specific
PCMH	Patient-Centered Medical Home
PCPI	Physician Consortium for Performance Improvement
PC-PTSD	Primary Care PTSD Screen
PDHA	Post-Deployment Health Assessment
PDHRA	Post-Deployment Health Reassessment
PH	psychological health
PHQ-2	Patient Health Questionnaire (two items)
PHQ-9	Patient Health Questionnaire (nine items)
PMPM	per member per month
PPTSD-R	Purdue PTSD Scale—Revised
PRIME-MD	Primary Care Evaluation of Mental Disorders
PTSD	posttraumatic stress disorder
QALY	quality-adjusted life-year
QI	quality improvement
QIDS-SR ₁₆	16-item Quick Inventory of Depressive Symptomatology
R-CMS	Revised Civilian Mississippi Scale for PTSD
RCT	randomized control trial
RESPECT-Mil	Re-Engineering Systems of Primary Care Treatment in the Military
RVU	relative-value unit
RWP	relative weighted product
SASQ	Single Alcohol Screening Question

SCID	Structured Clinical Interview for <i>Diagnostic and Statistical Manual of Mental Disorders</i> , 4th ed., Axis I Disorders
SDS	Sheehan Disability Scale
SI	suicide ideation
SIT	Stress Inoculation Training
SNRI	serotonin and norepinephrine reuptake inhibitor
SPAN	Startle, Physiological Arousal, Anger, and Numbness
SPTSS	Screen for Posttraumatic Stress Symptoms
SSDS	Self-Stigma of Depression Scale
SSRI	selective serotonin reuptake inhibitor
STAR*D	Sequenced Treatment Alternatives to Relieve Depression
TBI	traumatic brain injury
TF-CBT	Trauma-Focused Cognitive–Behavioral Therapy
TSC-40	Trauma Symptom Checklist—40
TSI	Trauma Symptom Inventory
VA	U.S. Department of Veterans Affairs
VHA	Veterans Health Administration

Introduction

Effective treatment of both physical and psychological health (PH) conditions in service members is an important part of the U.S. Department of Defense's (DoD's) goal of maintaining a healthy mission-ready force (Obama, 2012). Two PH conditions, posttraumatic stress disorder (PTSD) and major depressive disorder (MDD), have prevalence estimates that range from 4 to 20 percent for PTSD and 5 to 37 percent for MDD (Institute of Medicine [IOM], 2013; Ramchand, Schell, et al., 2010; Schell and Marshall, 2008). Providing evidence-based care for individuals with PTSD and MDD, which should improve patient outcomes, is a priority for the military health system (MHS).¹ However, there is currently no MHS-wide system to evaluate whether the care provided to service members is consistent with evidence-based clinical practice guidelines (CPGs) or MHS standards for high-quality care, to assess whether the care provided is improving outcomes, or to identify potential areas for improvement.

At the request of DoD, the RAND Corporation conducted a project to support DoD efforts to measure and improve the quality of care provided to service members who have PH conditions. The goal of the project was to develop a conceptual framework for assessing the quality of care for PH conditions and identify a candidate set of measures for monitoring, assessing, and improving the quality of care for two high-priority PH conditions in the MHS: PTSD and MDD. The effort focused primarily on outpatient care for PTSD and MDD provided in both primary and specialty-care settings. This document describes our conceptual framework for measuring the quality of PH care, our methods for identifying and selecting candidate quality measures, and the candidate measure sets for PTSD and MDD. In the remainder of this chapter, we provide a brief description of how we define *quality care*, as well as an introduction to how quality measures can be useful in assessing and improving the quality of care. We also provide an overview of the process we used to develop the candidate quality measure sets for PTSD and MDD.

¹ Evidence-based practice in psychology has been defined as “the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences” (American Psychological Association Presidential Task Force on Evidence-Based Practice, 2006, p. 280).

Defining and Measuring Quality of Care

The literature has defined *quality of care* in many ways, but the following definition from the IOM provides a sound basis for setting goals for quality improvement (QI). Quality of care is

[t]he degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge. (Committee to Design a Strategy for Quality Review and Assurance in Medicare, 1990, p. 21)

This definition establishes the importance of patient outcomes in deciding what care is high quality and allows for the standards of care to evolve over time as professional knowledge, including evidence from research studies, improves. The seminal report *Crossing the Quality Chasm* (IOM, 2001) presents six aims for improvement of quality of care, which have guided measure development (Table 1.1).

These aims provide broad targets for QI efforts and for development of quality measures that are useful tools in monitoring whether evidence-based care is being provided by health care professionals and hospitals within a system of care.

How Quality Measures Are Used

Quality measures, also called performance measures or metrics, provide a way to measure how well health care is being delivered. Such measures provide information about the health care system and highlight areas in which providers can take action to make health care safer and more equitable (National Quality Forum [NQF], undated [c]). Quality measure scores are typically presented as the percentage of eligible patients who received the recommended care (e.g., percentage of new PTSD patients screened for depression). According to the Agency for Healthcare Research and Quality (AHRQ) (undated [b]), quality measures are generally used by organizations for one or more of three purposes: QI, accountability, and research.

Table 1.1
Six Aims for Improving Quality of Care, from the Institute of Medicine

Aim	Description
Safe	Avoiding injuries to patients from the care that is intended to help them
Effective	Providing services based on scientific knowledge to all who could benefit, and refraining from providing services to those not likely to benefit
Patient-centered	Providing care that is respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient values guide all clinical decisions
Timely	Reducing waits and sometimes-harmful delays for both those who receive and those who give care
Efficient	Avoiding waste, including waste of equipment, supplies, ideas, and energy
Equitable	Providing care that does not vary in quality because of personal characteristics, such as gender, ethnicity, geographic location, and socioeconomic status

Quality Improvement

Health care organizations can focus on internal or external QI. Internal QI programs measure the quality of care within a health care organization or system. Internal QI usually consists of identifying clinical areas with less-than-optimal performance, selecting quality measures related to those areas, and measuring quality of care before and after improvement efforts are implemented (AHRQ, undated [b]). External QI programs measure quality in several health care organizations and compare performance across those organizations (AHRQ, undated [b]). Many agencies in the United States, including state, regional, or national entities or organizations (e.g., the Centers for Medicare and Medicaid Services [CMS]), accreditation and QI organizations (e.g., Joint Commission), and professional organizations (e.g., the American Medical Association's Physician Consortium for Performance Improvement [PCPI]), operate external QI programs. Users of external QI information are usually institutions (e.g., hospitals) or health care providers within the institutions who use the information as performance feedback. Within the MHS, both internal and external QI efforts may be desirable. Internal QI could focus on improving the quality of care within a given military treatment facility (MTF) or clinic, while external QI could focus on measuring and improving quality of care across the MHS. External QI can also help to reduce unjustified variability across MTFs or between the direct- and purchased-care systems.

Accountability

Organizations increasingly use quality measures for accountability purposes, including selecting health care providers, creating financial incentives for health care providers, and maintaining provider performance standards. The organizations and individuals who use quality data for accountability purposes overlap with those that use them for QI purposes, including purchasers and payers of health care (e.g., health insurance plans and CMS), regulatory agencies, accreditation organizations, and even patients (AHRQ, undated [b]). Using quality measures for accountability requires higher standards for validity and reliability than for QI to ensure a fair assessment of and comparison across health care providers. Therefore, more-rigorous adherence to the technical specifications and data-collection methods is necessary.

Organizations use quality measures for accountability purposes in three ways: public reporting, performance-based payment, and professional certification (AHRQ, undated [b]). First, organizations report quality measures publicly (e.g., Medicare's Hospital Compare), and those reports serve as the basis for decisions about health care by employers purchasing health care and by consumers selecting health care providers.² Second, health care systems implement programs to modify payments to physicians and hospitals based on whether they report quality data (i.e., pay-for-reporting programs) and whether they meet performance standards (i.e., pay-for-performance programs). Third, there is increasing use of quality measures in the certification of health care professionals (e.g., the American Board of Medical Specialties [ABMS] Maintenance of Certification process) and provider organizations (e.g., Diabetes Recognition Program offered by the National Committee for Quality Assurance [NCQA] and American Diabetes Association) (AHRQ, undated [b]).

² Hospital Compare is a website designed for consumers to find information about the quality of care provided by hospitals (CMS, 2013). The website was created by CMS and the Hospital Quality Alliance to assist consumers in making decisions about their health care and to support the improvement of care provided in U.S. hospitals.

Research

Quality measures are often used in health services research studies to measure how frequently evidence-based care is provided to different patient subgroups and in a variety of health care settings. They may be employed to measure rates in various population subgroups to document disparities in health care, differences between comparison groups in program evaluations, or changes over time in response to implementation of QI efforts (AHRQ, undated [b]).

Examples of Improved Performance over Time

To illustrate the impact that quality measurement can have on quality of care, we highlight performance data based on two Healthcare Effectiveness Data and Information Set (HEDIS®) quality measures (NCQA, 2013a). For a HEDIS measure targeting major depression, rates of effective acute-phase (i.e., at least 12 weeks) and continuation-phase (i.e., at least six months) treatment with an antidepressant increased from 58.8 percent to 69.1 percent and from 42.1 percent to 53.6 percent, respectively, from 1999 to 2012. For a HEDIS measure focused on hospitalizations for a mental health disorder, rates of follow-up within seven days and 30 days of discharge increased from 47.4 percent to 57.9 percent and from 70.1 percent and 76.0 percent, respectively, from 1999 to 2012. Performance rates for these two measures also improved for Medicaid and Medicare patients over the same time period. These data show care for mental health conditions in commercial and public health plans have improved steadily for more than a decade for patients with mental health conditions since implementation of routine measurement.

Overview of Our Process for Quality Measure Development

We used a systematic process to develop candidate measure sets for both PTSD and MDD, consisting of seven of the steps shown in Box 1.1.

First, we reviewed existing approaches for describing and measuring the quality of care and created a framework for selecting measures to assess the quality of PH care in the MHS context. The framework includes two measurement dimensions: the continuum of care for PH conditions (i.e., prevention, screening, assessment, treatment, and reintegration) and measure type (i.e., structure, process, outcome, patient experience, and resource use). We describe this framework in more detail in Chapter Two. We used this framework to classify existing measures and to adapt or develop new measures as needed.

Box 1.1 Process for Developing Candidate Measure Sets

1. Create a framework for classifying measures.
2. Identify existing measures.
3. Map existing measures to framework.
4. Examine whether new measures should be developed.
5. Develop new measures or adapt measures where needed.
6. Select a high-priority measure set.
7. Prepare conceptual descriptions for the measure set.
8. Operationalize detailed measure specifications.
9. Pilot-test measures within a particular health system.

Second, we identified existing measures by reviewing CPGs and databases of quality measures. We included measures used to assess structure, process, outcome, patient experience, and resource use, focusing primarily on measures assessing care for PTSD and MDD. In addition, we identified measures that assess aspects of care that may not be specific to PTSD and MDD but instead may apply to multiple PH conditions (e.g., assessment for suicide risk).

Third, we mapped identified existing measures to the quality measure framework. Subsequently, we used the framework to take an inventory of existing measures throughout the care continuum (i.e., prevention, screening, assessment, treatment, and reintegration) and across measure types (i.e., structure, process, outcome, patient experience, and resource use) in order to examine whether new measures were needed to assess quality for PTSD and MDD. We then developed or adapted measures to address high-priority areas in which existing measures were lacking. In addition, some measures were refined or adapted to incorporate recent evidence or to increase the likelihood that the measures would be appropriate for the MHS. Subsequently, we reduced the measure sets to focus on high-priority measures for PTSD and MDD. Finally, we prepared detailed descriptions of each candidate measure.

We note that implementation of a quality measure requires developing detailed technical specifications grounded in the data infrastructure available within a particular setting (e.g., *International Classification of Diseases*, ninth revision, clinical modification [ICD-9-CM] diagnosis, procedure codes, database, and variable names). This implementation work and field-testing is an essential next step to prepare these quality measures for use within the MHS (corresponding to steps 8 and 9 above) and is beyond the scope of this project.

This report presents information about each candidate measure that will allow readers to assess each measure's utility for monitoring, evaluation, or research across the MHS or within a particular MTF. For each measure, we provide the following information: measure, numerator, denominator, definitions, measure type, care setting, measure source, rationale for measure inclusion, potential data sources, feasibility, and supporting references. The description of the measure and supporting evidence will allow readers to decide which measures would be appropriate for further development and implementation.

Organization of the Report

Chapter Two describes the conceptual framework we used to select the PTSD and MDD measures. The process of identifying existing measures, mapping existing measures to the framework, developing new and adapted measures, and finalizing high-priority measure sets for PTSD and MDD is described in more detail in Chapter Three. We provide an overview of the measure sets in Chapter Four, and a detailed description of each selected measure in Appendixes A (descriptive terms used to describe the measures), B (PTSD measure descriptions), and C (MDD measure descriptions), as well as an expanded description of a cost measure in Appendix D. Appendixes E and F provide brief summary tables of the complete sets of candidate measures for PTSD and MDD, respectively. Finally, Appendix G lists all the measure sources that we examined.

Framework for Classifying Measures of Psychological Health Conditions

In this section, we describe a two-dimensional framework for classifying measures for PH conditions identified in the environmental scan. Using a framework for organizing existing measures on a specific topic and performing a gap analysis is recommended as the first step in the quality measure development guidelines developed by the Centers for Medicare and Medicaid Services (2014), known as the Measures Management System Blueprint. We developed the framework as a tool for gap analysis—that is, to assist us in evaluating whether the existing measures from the environmental scan address all important aspects of care for a particular PH condition. We used the framework to examine whether all topics in the framework were addressed and, if not, whether new measures should be developed or should be adapted from existing measures to address those topics. We conducted this process for PTSD and MDD, but one could apply this framework to other PH or general medical conditions as well.

Dimensions for the Framework

In designing the framework, we considered several possible dimensions. A previous study (Hermann and Palmer, 2002) developed a framework for PH measures with seven dimensions to assist in selecting measures for a core measure set for PH conditions. Their dimensions were described as follows:

domains (prevention, detection, access, assessment, treatment, continuity, coordination, and safety), clinical populations (diagnostic groups, comorbid conditions, prevalence, morbidity, treatability), vulnerable groups (children, elderly persons, racial and ethnic minorities, rural populations), modalities (medication, electroconvulsive therapy, psychotherapy, other psychosocial interventions), clinical setting (inpatient, ambulatory, intermediate, community, primary care, nursing homes, prisons), level of health care system (population, managed behavioral health care organization, delivery system, facility, provider, patient), and purpose of measurement (internal quality improvement, external quality improvement, consumer selection, purchasing, research).

We considered several of these for our framework. Because our study population was well-defined (i.e., active-duty service members with PTSD or MDD), we eliminated two of these dimensions (i.e., clinical populations and vulnerable groups). Another two dimensions (i.e., level of health care and purpose of measurement) relate to how the measure results will be aggregated and, therefore, were eliminated. All of the three remaining dimensions (i.e., domains, modalities, and clinical setting) are important aspects of care for PH conditions,

and were considered in describing each measure. However, we limited the framework to two dimensions to simplify the structure of the framework. We selected the care continuum, which is similar to “domains” listed above, as one of the framework dimensions. We selected measure type as the second dimension to ensure that as many aspects of care as possible were covered by the measure set. In this chapter, we describe these two dimensions in detail.

Care Continuum

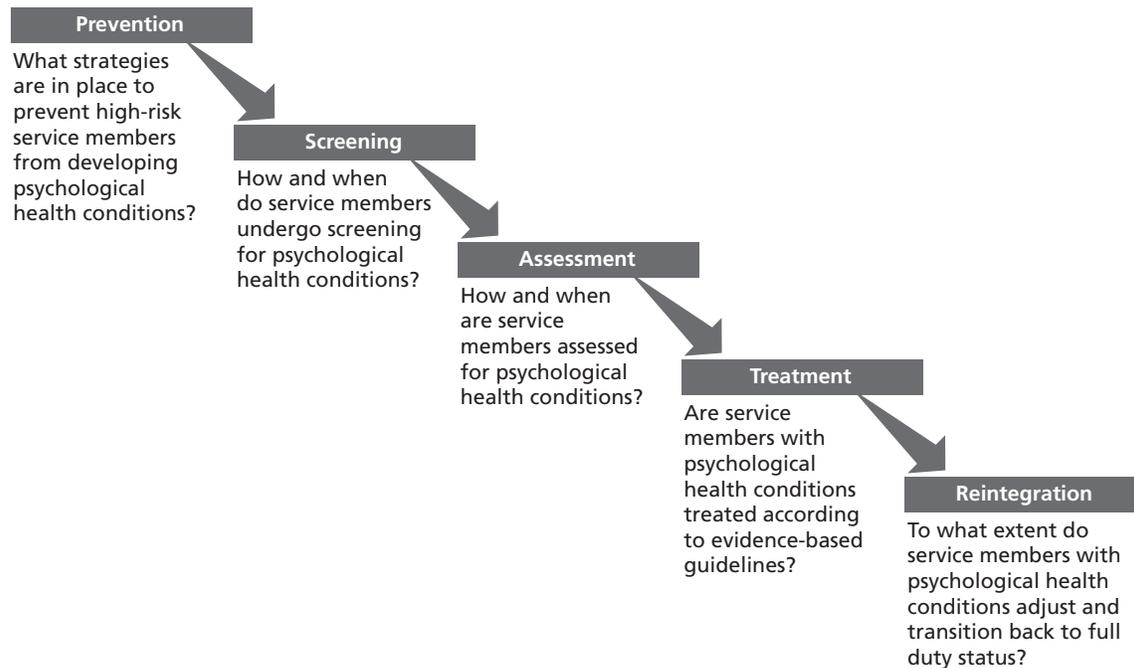
We defined the care continuum as consisting of five phases of care: prevention, screening, assessment, treatment, and reintegration. These phases are an adaptation of the continuum of care for PH adopted by the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (TBI) (DCoE) (i.e., surveillance, prevention, screening and assessment, diagnosis, treatment and recovery, rehabilitation, reintegration) (DCoE, 2011). We adapted this DCoE continuum for the framework by (1) removing surveillance because it was not a focus of the current research, (2) including diagnosis as part of assessment, and (3) including recovery and rehabilitation as part of treatment. The phases of the continuum of care are not well-documented in the literature, though the 2010 *VA [U.S. Department of Veterans Affairs]/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010) includes descriptions of each of the five phases except reintegration. We have defined the five phases of care in Table 2.1.

In practice, these five phases may overlap and not be entirely distinct. Nonetheless, they serve as a useful heuristic for developing and characterizing clinician actions and quality measures. We focus on mental health care delivered within primary care and specialty mental health clinics. As a result, although we use the full continuum as our framework, our selection process prioritized quality measures that assess components of care delivered in these settings. Those measures therefore represent the majority of our measures. Measures related to each phase of care can address important questions about care provided to service members for PH conditions (Figure 2.1). For example, a *screening measure* could assess whether a provider screens a patient for depression, while a *treatment measure* could assess whether a patient with PTSD has a documented treatment plan.

Table 2.1
Continuum-of-Care Phases

Phase	Description
Prevention	Providers try to increase resilience among otherwise-well individuals by using specific health interventions to decrease the likelihood of future illness and to identify people at greater risk for developing a particular PH condition and provide a targeted intervention to prevent it.
Screening	Providers or systems identify health risk factors or disease states in individuals or populations. Providers then systematically follow up on conditions and risks identified by screening.
Assessment	Providers complete medical and psychiatric histories and physical examinations to diagnose patient conditions and identify other factors that may affect the treatment plan (e.g., comorbid conditions, support systems, psychosocial stressors).
Treatment	Providers implement interventions to treat diagnosed conditions. For any condition, this phase would include initial treatment, assessment of response, and adjustment of treatment as needed.
Reintegration	Providers and systems focus their efforts on maintaining well-being and providing assistance and support services to the individual to adjust and transition back to full duty status.

Figure 2.1
Questions Addressed by Measures in the Care Continuum



RAND RR464-2.1

Measure Type

The second dimension of the framework is measure type. *Measure type* is a commonly used term in quality measurement and describes which aspect of health care is the focus of the quality measure (AHRQ, undated [a]). For example, a measure could assess whether there are enough clinicians per clinic, whether providers prescribe the appropriate medication, or whether patients' symptoms improve. NQF, a nonprofit organization that endorses standards for measuring and publicly reporting health care performance in the United States (NQF, undated [a]), uses measure type as one way to characterize each measure included in its measure database. There are three commonly used measure types, which Avedis Donabedian first described in 1966 (and updated in 2003): structure, process, and outcome (Donabedian, 1966, 2003). For our framework, we used Donabedian's descriptors and added two other types of measures: patient experience and resource use. We define the five measure types here:

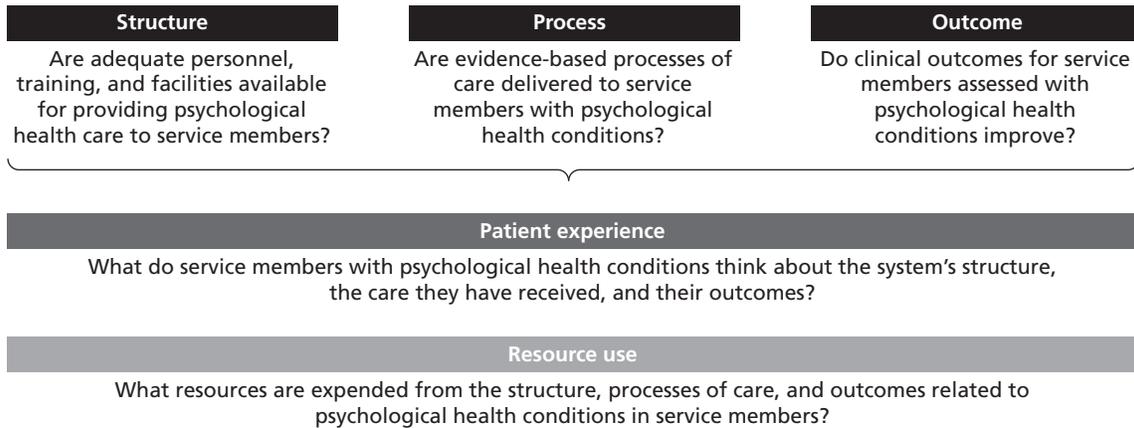
- A *structure measure* is one that assesses the settings in which patients receive health care and capacity to provide care. These measures may focus on the health care organization or characteristics of individual health care providers. Structure measures assess the adequacy of health facilities and equipment, the qualifications of clinicians and their organization within the entity of interest, the administrative structure and operations of programs and institutions providing care, and fiscal or payment organization. An example of a structure measure is “number of inpatient psychiatric beds available per 10,000 active-duty service members.”

- A *process measure* assesses whether a recommended care process or event takes place during the care of a condition or the degree to which a procedure or treatment is provided in a manner that reflects fidelity to the evidence base supporting its use. *Processes* generally refers to receipt of a particular service and often within a particular time frame. An example of a process measure is “percentage of MDD patients who received evidence-based psychotherapy for MDD.”
- An *outcome measure* describes the outcome of care in terms of patient improvement, recovery, restoration of function, avoidance of deterioration, or survival. Measurement of patient outcomes can guide accurate, cost-effective treatment decisions to increase the likelihood of achieving optimal patient outcomes (Harding et al., 2011). Outcome measures typically need risk adjustment to avoid biased comparisons of the measure rates. Without risk adjustment, providers or MTFs may appear to achieve poorer patient outcomes when, in fact, they are treating a sicker patient population. Risk adjustment is a “statistical process for reducing, removing, or clarifying the influences of confounding factors that differ among comparison groups” (CMS and Joint Commission, 2012, p. D-11). We can implement risk adjustment using logistic regression, stratification, and other statistical methods. An example of an outcome measure is “percentage of PTSD patients in a new treatment episode with improvement in functional status within six months of the initial visit.”
- *Patient experience measures* assess patients’ perceptions of their providers, the quality and timeliness of the care they receive, and the perceived impact of their care on their health outcomes. An example of a patient experience measure is “percentage of MDD patients who reported that they always received treatment quickly.”
- *Resource use measures* relate to the resources expended during the care of a patient with the specified condition (Romano, Hussey, and Ritley, 2010). We can measure resource use with a variety of approaches, including (1) utilization measures, (2) standardized resource use measures, and (3) cost measures. We can base measures on either the use of resources during the process of care and recovery or the amount of money spent directly to provide care. An example of a resource use measure is “average costs per member per month (PMPM).”¹ For the purposes of this report, resource use is included as one type of measure, so we do not discuss it separately in the method section. However, recognizing that resource use measures differ from other types of quality measures, we have included a more detailed discussion of the selected cost measure in Appendix D.

These five measure types can address important questions related to PTSD and MDD care for service members, but one could also apply them to other PH conditions. Patient experience and resource use measures can relate to aspects of the structure of the system, the process of care, or the outcomes; thus, we depict them as encompassing each of these domains (Figure 2.2).

¹ *Members* in the abbreviation *PMPM* does not refer strictly to service members. In this context, *members* refers more broadly to all TRICARE-enrolled beneficiaries.

Figure 2.2
Questions Addressed by Each Measure Type



RAND RR464-2.2

Matrix Framework for Classifying Psychological Health Measures

Using the two dimensions described above, care continuum and measure type, we created a matrix framework, with each cell defined by a unique combination of one phase in the continuum of care and one measure type (Table 2.2). Defining the cells in this manner enables us to place each existing measure into the appropriate cell to create a visual map of which measure topics are addressed by one or more existing measures and which measure topics are not addressed. Table 2.2 also highlights that patient experience measures and resource use measures are not typically specific to one phase in the care continuum. Using this framework facilitated the identification and prioritization of quality domains for further measure development.

Table 2.2
Framework for Classifying Measures for Psychological Health Conditions

Measure Type	Care Continuum				
	Prevention	Screening	Assessment	Treatment	Reintegration
Structure					
Process					
Outcome					
Patient experience					
Resource use					

Identification of Measures of Quality of Care for Posttraumatic Stress Disorder and Major Depressive Disorder

This chapter outlines the process by which we searched for and identified existing measures for assessing the quality of care for PTSD and MDD. Between December 2012 and March 2013, we conducted a systematic search of existing quality measures related specifically to PTSD and MDD, as well as measures that applied to patients with a range of PH conditions, including PTSD or MDD (e.g., assessment for suicide risk). The search targeted quality measure databases, measures utilized in related RAND evaluation studies, websites of individuals and organizations identified as stewards of relevant measures, and pertinent reports and CPGs published by VA and DoD. We identified search terms to facilitate database searches. Pertinent measures were included regardless of the level of detail provided (or lack thereof) to define the technical specifications of the measures of interest.

To find measures, we searched quality measure databases maintained by AHRQ (National Quality Measures Clearinghouse, or NQMC; AHRQ, undated [a]) and NQF (undated [a]). We also identified measures related to these conditions developed and tested by researchers from the RAND Corporation and the Altarum Institute in a national evaluation of mental health services delivered by the Veterans Health Administration (Farmer et al., 2010; Sorbero et al., 2010; Watkins, Pincus, Smith, et al., 2011). During the initial search, we adopted an inclusive approach and collected all measures related to adult PH (e.g., mood and anxiety disorders, substance use, suicide) across multiple measure types (e.g., treatment outcomes, cost, patient experience).

We initially scanned the AHRQ and NQF databases using the database search engines with selected search terms to find measures of interest.¹ In order to be sure we had captured all pertinent measures related to PH, we did an additional manual search of downloaded copies of the AHRQ and NQF databases. This search did not capture any additional measures besides what we initially found through the targeted search. We then searched the individual websites of other organizations, including all stewards with at least one PH measure in the AHRQ or NQF database (e.g., HEDIS) and mental health QI consortia (e.g., National Association of State Mental Health Program Directors). Finally, we reviewed pertinent VA and DoD reports and CPGs to identify possible measures related to clinical areas of interest (VA, 2011, 2012a, 2012b; DoD Task Force on the Care, Management, and Transition of Recovering Wounded, Ill, and Injured Members of the Armed Forces, 2012; Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010; U.S. National Institutes

¹ Search terms were *depression, behavioral health, mental health, MDD, PTSD, suicide, post-traumatic stress disorder, post-traumatic stress disorder, trauma, traumatic, antidepressant, psych, psychiatric, psychological, partner violence, recovery, cost, outcome, patient experience, and patient satisfaction.*

of Health, 2013).² Because the focus of this work was on existing quality measures that could be broadly implemented, we did not include a search of the scientific literature for identifying measures. While such a search could reveal the evidence base for existing measures and results of individual quality improvement interventions, it was less likely to be a source of well-documented existing measures for broad application than the specified sources utilized. However, we did consult the literature to address gaps identified when the collected measures were placed in the measure matrix.

The initial compilation of existing quality measures included 530 measures relevant to PTSD, MDD, or PH more broadly (108, 144, and 278 measures, respectively). From this set, we excluded measures assessing PTSD and MDD screening within a limited subpopulation (e.g., patients with diabetes who were screened for MDD). In addition, we removed measures pertaining only to pediatric patients or patients in settings of limited relevance to the service member population (e.g., home health, long-term care facilities). This reflected our focus on identifying high-priority measure sets assessing care for service members. Although some of these same measures could be applicable to family members and retirees, an organization would likely need additional measures to fully address these additional populations. Culling these measures from the set resulted in 416 measures for PTSD, MDD, and PH (101, 144, and 171 measures, respectively).

Measure Selection by Expert Consensus

We next convened from the project team five RAND researchers who are specialists in quality measurement, PH, and military populations to review the preliminary set of measures. The group included two psychiatrists (Harold Alan Pincus and Katherine E. Watkins), two clinical psychologists (Coreen Farris and Kimberly A. Hepner), and one mental health services researcher (Carrie M. Farmer). By consensus, we dropped from the set 61 PH measures that would be inapplicable to most PTSD or MDD patients seen within the MHS (e.g., episodes of physical restraint during inpatient hospitalizations). To identify the best candidate measures from the remaining quality measures, the group considered whether the following were true:

- Adequate scientific evidence or professional consensus exists to support a link between the performance of care specified by the measure and the health benefits to patients with PTSD or depression. This included support from existing clinical practice guidelines, either from a professional organization (e.g., American Psychiatric Association [APA]) or VA or DoD (validity).
- We would consider a provider or organization with significantly higher rates of adherence to the measure a higher-quality provider (validity).
- The measure had been previously applied to military populations, including veterans (feasibility of MHS implementation).
- Either the measure applies to many individuals or there are serious adverse consequences from not adhering to the indicator (importance).
- There is variation among providers in the provision of the specified care indicating a quality gap and room for performance improvement (importance).

² A full list of sources reviewed is included in Appendix G.

- The information necessary to determine adherence within the MHS is likely to be available in administrative data, in the medical record, or via patient (or family) interview (feasibility of MHS implementation).
- NQF had endorsed the measure (NQF-endorsed), an endorsement that we used as a proxy for importance, validity, reliability, feasibility, and usability.

Team members assessed measures individually by providing a global rating that weighed the above criteria, and then the group convened to discuss global ratings and identify the candidate measure sets. Measures identified for further consideration were those that generated a positive rating from at least three of the five team members. There was significant duplication in the identified measures in terms of the care they assessed. For ease of comparison, measures were grouped by focus of care (e.g., assessment, treatment, response to treatment). Where multiple measures existed for a similar concept, preference was given to measures that best met the above-listed criteria, those with more-complete technical specifications, professional organization or NQF endorsement, and those that were more in line with DoD data availability and policy directives. Team members also tried to include measures that covered the continuum of care, all types of measures (e.g., process, outcome), and measures that were important to mental health conditions in addition to PTSD and MDD (e.g., assessment for suicidal ideation). In the group discussion, we placed particular emphasis on identifying measures that could be targets for further refinement and field-testing within MTFs in the near term. In some cases, existing measures were adapted for inclusion in the PTSD and MDD measure sets. For example, this may have been the result of a modification to a measure-related definition, or it may have been the modification of an MDD measure, for example, for its application to PTSD. Also, where no measures existed targeting a particular aspect of care, we constructed new measures.

Consider our selection process a first step toward a formalized process to validate and endorse a set of quality measures. The process we used to identify and select measures is distinct from NQF's process for measure endorsement. The NQF process does not involve a systematic search for measures. Instead, it relies on the publication of a formal request for the submission of validated measures (i.e., a call for measures). It also requires that a measure developer (i.e., measure steward) submit the detailed technical specifications for the measure, documentation of measure reliability and validity, formal harmonization with similar NQF-endorsed measures, and a plan to ensure that the measure is updated on a schedule in keeping with the rate of clinical innovation. Only with these tasks documented can a panel of NQF expert stakeholders consider the measure for endorsement. Those experts rate formal and sequential criteria for measure endorsement, including the importance of the measure, the scientific acceptability of measure properties, usability, and feasibility (Table 3.1). The panel then submits the endorsed measures to additional internal, public, and stakeholder reviews. This formal NQF endorsement process can take up to a year to complete.

Although the process we used to identify and review candidate PTSD and MDD measures took into consideration several of the NQF criteria included in Table 3.1 (e.g., weighing clinical importance during deliberations), we did not attempt to replicate the lengthy process for formal measure endorsement used by NQF. Some of the measures in the MDD measure set are already NQF-endorsed; currently, however, there are no NQF-endorsed measures specific to PTSD care. Other measures from the candidate sets may be submitted for consideration for endorsement by NQF in the future.

Table 3.1
National Quality Forum's Measurement Evaluation Criteria

Stage of Evaluation	Criteria
Conditions to be met prior to measure consideration	<ul style="list-style-type: none"> • The measure is in the public domain or measure steward agreement is signed. • The measure is updated on a schedule commensurate with the rate of clinical innovation. • The measure includes both accountability applications and performance improvement to achieve high-quality, efficient health care. • The measure is fully specified and tested for reliability and validity. • The measure has been harmonized with competing measures.
Measures are evaluated for their suitability based on four sets of standardized criteria (listed in order of importance)	<ul style="list-style-type: none"> • <i>Importance of the measure:</i> The extent to which the specific measure focus is evidence-based, is important to making significant gains in health care quality, and improves health outcomes for a specific high-priority (high-impact) aspect of health care where there is variation in or overall less-than-optimal performance. • <i>Scientific acceptability of measure properties:</i> The extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care, when implemented. • <i>Usability:</i> The extent to which potential audiences are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient health care for individuals or populations. • <i>Feasibility:</i> The extent to which the required data are readily available or could be captured without undue burden and the quality measure can be implemented for performance evaluation.

SOURCE: NQF, undated (b).

Overview of Candidate Sets of Measures of Care Quality for Posttraumatic Stress Disorder and Major Depressive Disorder

As discussed in Chapter Three, the selection of measures for the candidate PTSD and MDD measure sets was an iterative process. The RAND team first identified PTSD measures, MDD measures, and measures that can be applied to patients with a range of PH conditions, including PTSD and MDD, to identify existing measures that could be included in or adapted for the measure sets. Using an expert consensus process, we reviewed these measures and selected measures based on their importance, previous use of the quality measure in a military population, and the feasibility of implementation. We also adapted several MDD measures that are also applicable to the care of PTSD and created similarly constructed PTSD measures. As a result of this process, we selected final sets of 58 measures each for PTSD and MDD. This chapter briefly describes these measures, and complete descriptive information about each measure (including numerator, denominator, and rationale for selection) is available in Appendix B (PTSD) and Appendix C (MDD). This set of measures will be narrowed further in future work to be conducted by RAND researchers, which will identify, with key stakeholders, feasible and high-priority measures for complete technical specification, piloting, and implementation planning.

Tables 4.1 and 4.2, respectively, list the 29 PTSD measures and 29 MDD measures. More-detailed summary tables of each of the measure sets are included for ease of reference in Appendixes E and F. Ten measures in the PTSD set are PTSD-specific, and ten measures in the MDD set are MDD-specific. Nineteen of the measures in each set could apply to care for patients with a range of PH conditions, including PTSD and MDD. These measures are marked with an asterisk (*) added to the measure number in the first column in both Tables 4.1 and 4.2. (To simplify their application across those diagnoses, the asterisked measures have similar specifications for both PTSD and MDD. Because the focus of this project was on PTSD and MDD, it was outside the scope of this project to identify exactly how these broader measures would be applied to other PH conditions.) Tables 4.1 and 4.2 organize the selected measures by position along the continuum of care (screening, assessment, and treatment), as well as by measure type (structure, process, outcome, patient experience, and resource use).¹

¹ As discussed in the next section, we did not identify any appropriate prevention or reintegration measures for PTSD or MDD.

Table 4.1
Candidate Set of Measures of Quality of Care for Posttraumatic Stress Disorder

Phase		Measure Type
Screening		
PTSD-S1*	Percentage of patients screened for PTSD, MDD, and alcohol misuse and, if positive, appropriate follow-up initiated	Process
Assessment		
PTSD-A1	Percentage of PTSD patients in a new treatment episode with assessment of symptoms with the PCL	Process
PTSD-A2	Percentage of PTSD patients in a new treatment episode assessed for depression	Process
PTSD-A3*	Percentage of PTSD patients in a new treatment episode assessed for suicide risk	Process
PTSD-A4*	Percentage of PTSD patients in a new treatment episode assessed for recent substance use	Process
Treatment		
PTSD-T1	Percentage of PTSD patients with symptom assessment with the PCL during a four-month measurement period	Process
PTSD-T2*	Percentage of days when the third available specialty-care appointment is within two days	Structure
PTSD-T3*	Percentage of patient contacts with SI with appropriate follow-up	Process
PTSD-T4*	Percentage of PTSD patients with a documented treatment plan	Process
PTSD-T5	Percentage of PTSD patients with an adequate trial of SSRIs or SNRIs	Process
PTSD-T6	Percentage of PTSD patients newly prescribed an SSRI or SNRI with a follow-up visit within 30 days	Process
PTSD-T7	Percentage of PTSD patients who receive evidence-based psychotherapy for PTSD	Process
PTSD-T8*	Percentage of PTSD patients in a new treatment episode who received any psychotherapy	Process
PTSD-T9*	Percentage of PTSD patients with four psychotherapy visits or two medication-management visits within the first eight weeks	Process
PTSD-T10	Percentage of PTSD patients with response to treatment at six months	Outcome
PTSD-T11	Percentage of PTSD patients with response to treatment at 12 months	Outcome
PTSD-T12	Percentage of PTSD patients in PTSD-symptom remission at six months	Outcome
PTSD-T13	Percentage of PTSD patients in PTSD-symptom remission at 12 months	Outcome
PTSD-T14*	Percentage of PTSD patients in a new treatment episode with improvement in functional status at six months	Outcome
PTSD-T15*	Percentage of psychiatric inpatient hospital discharges of patients with PTSD with follow-up in 30 days or seven days	Process
PTSD-T16*	Number of inpatient psychiatric beds available per 10,000 active-duty service members	Structure

Table 4.1—Continued

Phase		Measure Type
All phases of care		
PTSD-PE1*	Percentage of PTSD patients who report getting treatment quickly	Patient experience
PTSD-PE2*	Percentage of PTSD patients who report being given as much self-management information as they wanted	Patient experience
PTSD-PE3*	Percentage of PTSD patients who report being told about treatment options	Patient experience
PTSD-PE4*	Percentage of PTSD patients who report being helped by counseling or treatment received	Patient experience
PTSD-PE5*	Percentage of PTSD patients who report being better than they were one year ago	Patient experience
PTSD-PE6*	Percentage of PTSD patients who rated counseling and treatment received as 9 or 10 (out of 10)	Patient experience
PTSD-RU1*	Number of psychiatric admissions per 100 patients with PTSD	Resource use
PTSD-RU2*	Average costs PMPM	Resource use

NOTE: * = Measure that could be applied to patients with PH conditions other than PTSD. PCL = PTSD Checklist. SI = suicidal ideation. SSRI = selective serotonin reuptake inhibitor. SNRI = serotonin and norepinephrine reuptake inhibitor.

**Table 4.2
Candidate Set of Measures of Quality of Care for Major Depressive Disorder**

Phase		Measure Type
Screening		
MDD-S1*	Percentage of patients screened for PTSD, MDD, and alcohol misuse and, if positive, appropriate follow-up initiated	Process
Assessment		
MDD-A1	Percentage of MDD patients in a new treatment episode with assessment of symptoms with the PHQ-9	Process
MDD-A2	Percentage of MDD patients in a new treatment episode assessed for manic or hypomanic behaviors	Process
MDD-A3*	Percentage of MDD patients in a new treatment episode assessed for suicide risk	Process
MDD-A4*	Percentage of MDD patients in a new treatment episode assessed for recent substance use	Process
Treatment		
MDD-T1	Percentage of MDD patients with symptom assessment with the PHQ-9 during a four-month measurement period	Process
MDD-T2*	Percentage of days when third available specialty-care appointment is within two days	Structure
MDD-T3*	Percentage of patient contacts with endorsement of SI with appropriate follow-up	Process
MDD-T4*	Percentage of MDD patients with a documented treatment plan	Process

Table 4.2—Continued

Phase		Measure Type
MDD-T5	Percentage of MDD patients newly treated with an antidepressant for 12 weeks or six months	Process
MDD-T6	Percentage of MDD patients newly prescribed an antidepressant with a follow-up visit within 30 days	Process
MDD-T7	Percentage of MDD patients who received evidence-based psychotherapy for MDD	Process
MDD-T8*	Percentage of MDD patients in a new treatment episode who receive any psychotherapy	Process
MDD-T9*	Percentage of MDD patients with four psychotherapy visits or two medication-management visits in the first eight weeks	Process
MDD-T10	Percentage of MDD patients with response to treatment at six months	Outcome
MDD-T11	Percentage of MDD patients with response to treatment at 12 months	Outcome
MDD-T12	Percentage of MDD patients in MDD-symptom remission at six months	Outcome
MDD-T13	Percentage of MDD patients in MDD-symptom remission at 12 months	Outcome
MDD-T14*	Percentage of MDD patients in a new treatment episode with improvement in functional status at six months	Outcome
MDD-T15*	Percentage of psychiatric inpatient hospital discharges of patients with MDD with follow-up in 30 days or seven days	Process
MDD-T16*	Number of inpatient psychiatric beds available per 10,000 active-duty service members	Structure
All phases of care		
MDD-PE1*	Percentage of MDD patients who report getting treatment quickly	Patient experience
MDD-PE2*	Percentage of MDD patients who report being given as much self-management information as they wanted	Patient experience
MDD-PE3*	Percentage of MDD patients who report being told about treatment options	Patient experience
MDD-PE4*	Percentage of MDD patients who report being helped by counseling or treatment received	Patient experience
MDD-PE5*	Percentage of MDD patients who report being better than they were one year ago	Patient experience
MDD-PE6*	Percentage of MDD patients who rated counseling and treatment received as 9 or 10 (out of 10)	Patient experience
MDD-RU1*	Number of psychiatric admissions per 100 patients with MDD	Resource use
MDD-RU2*	Average costs PMPM	Resource use

NOTE: * Measure that could be applied to patients with PH conditions other than MDD. PHQ-9 = Patient Health Questionnaire (nine items).

Description of Candidate Measures, by Care Continuum Phase

In Table 4.3, the measures for PTSD and MDD are broken down by the care continuum phase to which they apply. There are no measures in either measure set for two of the phases, prevention and reintegration. However, each measure set includes one screening measure (including follow-up). The screening measure focuses on screening for PTSD, MDD, and alcohol misuse in the general population of adult patients. This measure emphasizes the need to use standardized tools for screening for these conditions. The measure includes initiating appropriate follow-up for those patients who screened positive for any one of the screened conditions.

There are eight measures related to assessment in the two measure sets. Two PTSD measures assess whether baseline PTSD symptom severity was documented and whether an assessment of comorbidity (depression) was documented. Two MDD measures assess whether baseline MDD-symptom severity was documented and whether an assessment of prior or current symptoms of mania or hypomania was documented. And two measures that appear in both measure sets look for an assessment for comorbid conditions relevant to both PTSD and MDD (assessments for suicide risk and for recent substance use). The two MDD measures addressing assessment for symptoms of prior or current mania or hypomania and assessment for suicide risk are both NQF-endorsed. Both of the symptom severity measures emphasize the need to use a standardized tool to assess a patient's baseline level of severity. In the case of MDD, this is the nine-item PHQ-9 (Kroenke, Spitzer, and Williams, 2001); for PTSD, this tool is the PCL (Blanchard, Hickling, et al., 2003). In response to revisions to the diagnostic criteria for PTSD in the *Diagnostic and Statistical Manual of Mental Disorders* (APA, 2013), the National Center for PTSD is working to revise the PCL (National Center for PTSD, 2014). If scoring criteria change because of this effort, the specifications for measures that include the PCL will need updating. The baseline assessment is a key component of the measurement-based care framework and essential for comparison with later patient assessments to evaluate the patient's response to treatment over time. The use of a single validated tool for this purpose facilitates comparing results across multiple settings of care.

Table 4.3
Quality Measures for Posttraumatic Stress Disorder and Major Depressive Disorder, by Care Continuum

Care Continuum	PTSD	MDD	Total
Prevention	0	0	0
Screening	1	1	2
Assessment	4	4	8
Treatment	16	16	32
Reintegration	0	0	0
Measures that span the entire continuum			
Patient experience	6	6	12
Resource use	2	2	4
Total	29	29	58

There are 32 treatment measures: Eight focus on aspects of PTSD treatment, eight focus on MDD treatment, and eight focus on treatment for a range of PH issues relevant to both PTSD and MDD and are included in both measure sets. The condition-specific treatment measures for PTSD and MDD focus on important aspects of condition-related delivery of measurement-based care and evidence-based treatment options (to the extent possible): periodic assessment of PTSD and MDD symptoms over time with a standardized tool (PCL for PTSD, PHQ-9 for MDD), treatment with appropriate choice of medication and treatment duration (use of SSRIs or SNRIs and an antidepressant for PTSD and MDD), appropriate medication-management follow-up, receipt of evidence-based psychotherapy, and assessment of effectiveness of treatment through documentation of patient progress toward remission as response and remission based on PCL and PHQ-9 scores. Four of the MDD treatment measures are NQF-endorsed (periodic symptom assessment with the PHQ-9, duration of antidepressant therapy, and depression remission at six months and 12 months based on PHQ-9 scores), and two MDD measures are currently under consideration for NQF endorsement (patient response as progress toward remission at six months and 12 months based on PHQ-9 scores). The PTSD measure set includes similarly constructed PTSD measures targeting the same aspects of treatment. The MDD NQF-endorsed measure that focuses on duration of antidepressant therapy is also included in HEDIS 2013 (NCQA, 2012a).

Treatment measures were also included that can apply more broadly to PH conditions other than PTSD and MDD. These include two measures that focus on capacity of available care (timing to third available specialty-care appointment and mental health inpatient bed capacity). Another six measures target appropriate follow-up of significant events (patients with endorsed SI and patients discharged after a mental health inpatient admission), use of a treatment plan to document and manage treatment goals, general measures of receipt of a minimal level of care (receipt of any psychotherapy for a new treatment episode and receipt of care in the first eight weeks of a new treatment episode), and improvement in patient function after six months of care. The measure that monitors follow-up after a mental health inpatient discharge is an NQF-endorsed measure and is included in HEDIS 2013.

In addition to the phase-specific measures, two types of measures span the entire continuum of care: patient experience and resource use. We discuss them here with regard to their content and focus, and they are included in the discussion of measure type below. Each measure set includes six patient experience measures drawn from the NQF-endorsed Experience of Care and Health Outcomes (ECHO) (ECHO Development Team, 2002) survey. The ECHO measures assess patient-reported experience of behavioral health treatment with regard to perceived accessibility, usefulness, provision of treatment options, and perceived impact of the care received. Each measure set also includes two resource use measures (psychiatric inpatient admissions per 100 patients and cost of care PMPM). Both the patient experience and the resource use measures can be widely applied to patients with PH conditions other than PTSD and MDD.

Several candidate measures are appropriate only for patients who are beginning a new course of treatment. Specifically, these patients are either receiving treatment for a particular condition for the first time or returning to seek treatment after a period of receiving no treatment for the condition. A patient in a new treatment episode should receive several processes

of care (e.g., assessment for co-occurring conditions) that may not be appropriate for a patient receiving ongoing treatment. We have defined a new treatment episode as follows:

a PTSD- or MDD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD or MDD is the primary diagnosis

or

an outpatient encounter in which PTSD or MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD or MDD is either the primary or nonprimary diagnosis for six or more months).

When creating a definition such as this, there is always the risk that the definition does not accurately capture the targeted population, either by including any not in a new treatment episode or excluding any who is not. We have chosen to define a new treatment episode with a single encounter in which PTSD or MDD is the *primary* diagnosis with no PTSD or MDD encounters (primary or secondary) in the prior six months. Requiring more than one encounter in the definition risks excluding anyone who started a new course of treatment (as indicated by the provider assigning a diagnosis) but who received inadequate follow-up care.

Description of Candidate Measures, by Measure Type

Table 4.4 breaks down the PTSD and MDD measures by measure type. Almost half of the measures in each set (14 of 29) are process measures. Six measures in each set assess patient experience. Five measures in each set assess outcomes. Two of the measures in each set are structure measures. Lastly, two measures in each set assess resource use.

Description of Candidate Measures Using the Matrix Framework

In Table 4.5, we illustrate how the 58 PTSD and MDD measures were categorized in the two-dimensional framework of care continuum by measure type.

Table 4.4
Measures of Quality of Care for
Posttraumatic Stress Disorder and
Major Depressive Disorder, by
Measure Type

Measure Type	PTSD	MDD	Total
Structure	2	2	4
Process	14	14	28
Outcome	5	5	10
Patient experience	6	6	12
Resource use	2	2	4
Total	29	29	58

Table 4.5
Measures of Quality of Care for Posttraumatic Stress Disorder and Major Depressive Disorder, by Care Continuum and Measure Type

Measure Type	Care Continuum											
	Prevention		Screening		Assessment		Treatment		Reintegration		Total	
	PTSD	MDD	PTSD	MDD	PTSD	MDD	PTSD	MDD	PTSD	MDD	PTSD	MDD
Structure	0	0	0	0	0	0	2	2	0	0	2	2
Process	0	0	1	1	4	4	9	9	0	0	14	14
Outcome	0	0	0	0	0	0	5	5	0	0	5	5
Patient experience											6	6
Resource use											2	2
Total											29	29

As mentioned above and illustrated in the table, we identified no appropriate measures addressing the prevention or reintegration stages of care. The single measure selected for each condition in the screening phase addresses process of care, but we selected no screening-phase measures that focused on structure or outcome. Similarly, the eight measures from the combined PTSD and MDD measure sets related to the assessment phase of care also focus on process of care. Treatment-phase measures that we selected for the candidate sets, however, include measures focusing on structure, process, and outcome (four, 18, and ten measures, respectively, across the two measure sets combined). We selected 12 patient experience measures and four resource use measures across the two measure sets combined. Mapping the candidate measure sets to the framework allowed us to evaluate whether cells with few or no measures suggest high-priority areas for further measure development.

Examination of Need for New Measures

We identified two care continuum phases in which there were no existing measures: prevention and reintegration. We also identified a content gap in measures to assess the availability of inpatient care. We selected some of these high-priority areas to fill with newly developed or adapted measures (e.g., availability of specialty care, inpatient bed capacity).

Prevention

Our goal was to develop quality measures to assess care provided within MHS treatment settings (e.g., primary care, specialty mental health care). These treatment settings do not focus on the prevention of mental health conditions, which is usually attempted at the population level.² Typically, service members accessing treatment have already developed PH symptoms. However, the health care system could implement some best practices in prevention (e.g.,

² Indeed, we acknowledge that major efforts are under way across DoD to address prevention of mental health problems by building resilience. However, developing measures to assess these efforts was beyond the scope of this project.

avoiding immediate mental health services for trauma-exposed individuals *without* acute stress symptoms) (Roberts et al., 2009; Rose et al., 2002). Our belief is that such an event would occur rarely at the level of care for which we are providing recommendations and therefore is not an initial priority area. Instead, we focused on components of quality care that are the responsibility primarily of providers operating in a treatment setting. As administrators and policymakers seek to expand the scope of quality assessment, measurement of the quality of prevention services may become a priority.

Reintegration

Reintegration, the extent to which a service member successfully returns to his or her pre-morbid occupational and social role after completing treatment, is an important outcome of quality care. We identified two candidate outcome measures within this domain. One assessed the number of duty days missed by the service member (Post-Deployment Health Guideline Expert Panel, 2001). The second assessed the proportion of service members treated who returned to full duty (Post-Deployment Health Guideline Expert Panel, 2001). Although we carefully considered both measures, we did not select either for the candidate measure set. Defining the quality of care based on return to duty is problematic for two reasons. Both measures penalize programs that treat more severely symptomatic patients who are unlikely to return to duty relative to programs that treat other patients. For patients with severe symptoms, a more positive outcome may be to transition to a civilian role that best matches a new level of functioning, rather than to return to duty. Finally, the first measure, missed duty days, includes two distinct categories of missed duty days that should not be combined: those due to psychiatric morbidity and those days missed to receive appropriate treatment. Although the two identified measures represent important first steps to measuring reintegration, technical specifications have not been developed for either, and we focused on measures that had some prior foundational work rather than completely new measures. Given the importance of reintegration as an outcome for service members, the complexity of appropriately specifying the measure, and the considerable policy implications, we recommend that developing reintegration measures be undertaken in future work.

Structure Measures

Within the structure component of measure types, we were unable to identify an existing measure corresponding to basic availability of inpatient services. To estimate of the availability of inpatient care, we constructed a measure applicable to the MHS that assessed the average number of inpatient psychiatric beds available per 10,000 active-duty service members (measures PTSD-T16* and MDD-T16*).

We note that there are additional strategies to assess the structure of care that were outside the scope of this project to identify relevant quality measures. For example, researchers and policymakers have been developing measures assessing the extent to which primary care practices have adopted structural characteristics that are consistent with the key principles of the Patient-Centered Medical Home (PCMH) (Martsof et al., 2012; Rittenhouse, Casalino, Gillies, et al., 2008; Rittenhouse, Casalino, Shortell, et al., 2011). PCMH is a system of primary care that aims to improve quality and patient experience and reduce costs through practices' adoption of key structural characteristics, such as care coordinators, disease-management services, electronic health records (EHRs), and enhanced access to care (American Academy of Family Physicians [AAFP] et al., 2007). In addition, other work used a facility survey to assess

whether certain aspects of structure that support delivery of evidence-based care for PH were present. Although these results were not reported as quality measures, they could serve as the foundation for new structure measures with additional development work (Woodroffe et al., 2010). Assessing the capacity of the MHS as a whole to provide coordinated evidence-based care for PTSD and MDD is an important task that should be addressed by future research.

Summary and Recommendations

This document described the process we followed to develop sets of candidate measures for assessing the quality of care provided for PTSD and MDD in the MHS. We reviewed 530 existing quality measures relevant to care for PTSD, MDD, and PH conditions more broadly. We selected some measures from the existing quality measures and developed or adapted others to build the measure sets. As a result of this process, we selected 29 PTSD and 29 MDD measures as candidate measures for near-term implementation by the MHS. Ten measures in the PTSD set are PTSD-specific, and ten measures in the MDD set are MDD-specific. Nineteen of the measures in each set could be applied to patients with a range of PH conditions, including PTSD and MDD.

Study Limitations

Our approach to developing these measure sets has limitations. First, we relied on an internal team of RAND experts in mental health and quality measurement to select candidate measures from among multiple measures representing the same measure concept. As part of this process, the experts prioritized measures primarily based on retaining measures that were developed and tested using rigorous methods. In some cases, they also considered whether there was adequate scientific evidence or professional consensus to support a link between the care specified by the measure and the health benefits to patients with PTSD or depression. Although it may have been improved by a formal expert panel process, an informal selection process does not require a formal expert panel. Second, because our population focus was service members, additional measures are likely needed to assess care for retirees and family members. We note, however, that the measures identified are likely to be highly relevant to these additional populations. Third, complete technical specifications were not readily available for each of the 530 measures, so the review of existing measures focused on the information available from the various data sources in which these measures were described or listed, which varied in the level of detail provided. Another limitation of the study is that patient preferences often are not considered when CPGs are developed (Montori, Brito, and Murad, 2013), and, as a consequence, patient preferences are not generally incorporated into implemented measures based on those guidelines, particularly if those measures are implemented solely with administrative data. This limitation applies to most quality measures.

In addition, although we relied on one framework for categorizing measures, consideration of other frameworks in future work could be helpful. Under the NQF project on Patient Outcomes Measures (NQF, 2011a), a group of mental health experts developed a measurement

framework for mental health and substance use treatment. The framework encompasses five characteristics the steering committee considered to be important aspects of measuring the quality of mental health care: inclusion of mental health in broad, cross-cutting measures; consumer, patient, family, and caregiver satisfaction; promotion of healthy behaviors and environment; nontraditional measures (e.g., homelessness); and accountability and care coordination. Alternatively, the IOM aims (IOM, 2001) for improving quality of care—safety, effectiveness, patient-centeredness, timeliness, efficiency, and equitability—might be utilized to assess the completeness of the current set of candidate measures and the direction for future expansion.

Further, although we relied on the available information to assess the feasibility of using these measures in the MHS, the measures have not been formally assessed for adaptability and feasibility across the range of MHS settings. The measure descriptions provided in Appendixes B and C represent an initial effort to summarize important aspects of the measures and the basic operational details that are important to their potential implementation. They do not make up a technical manual but instead provide a foundation for understanding each measure's content and evaluating its potential value as a means of evaluating the quality of mental health care provided by the MHS. We have included in the measure descriptions some initial assessments of feasibility of implementation based on measure-required data. The actual implementation of quality measures is a complex and separate process (Damschroder et al., 2009; Nicholas et al., 2001). Any implementation of measures requires a comprehensive effort that is sensitive to the patient population to which it will be applied, the care setting, and the care providers within that setting. The feasibility of using a measure will be influenced by the availability of data currently collected or data that could be collected in that setting. Existing systems of documentation of care, site workflow, and provider preferences are some of the factors that would need to be considered and factored into an implementation plan. Planning for implementation also includes identification of senior and local clinical champions, solicitation of staff and provider input, provider education, pilot-testing, and evaluation. Once such a plan is developed and measures for implementation are selected, the detailed technical specifications to define the application of those measures in that particular setting would be defined while maintaining the basic definitional integrity of the measures as described here.

Recommendations

In this section, we provide recommendations to support DoD's continuing efforts to improve quality of care for PH conditions.

Recommendation 1: Select a subset of high-priority, feasible quality measures to pilot-test and implement. This report includes 58 candidate measures to assess care for PTSD and MDD. We do not expect that it will be appropriate or feasible to implement all 58 of these measures. Instead, we present a candidate set of measures to consider for future implementation. For each measure, we have estimated the feasibility of using the measure to assess quality of care performance based on our assessment of currently available data sources in the MHS. Specifically, we indicate whether the measure is green (i.e., can be computed using administrative data or other data that are extractable from existing electronic sources in the MHS), yellow (i.e., required data elements for measure computation may require medical record abstraction but could alternatively be integrated into a data-entry tool and the requisite data elements extracted electronically), or orange (requires medical record abstraction). The majority of the

measures in the PTSD and MDD sets fall into the yellow and orange categories, indicating the need for electronic data extraction or medical record review. In addition, we acknowledge variability in how care is structured and delivered across service branches and that the delivery of PH care within the MHS is constantly evolving, which will affect which measures DoD deems high priority. DoD should consider focusing initial quality measurement efforts on feasible measures that assess high-priority aspects of care at the time of implementation.

Recommendation 2: Consider structured documentation in the medical record to facilitate capturing the data necessary to derive quality measure performance. Many candidate quality measures rely on important clinical details that are typically documented in the medical record. Although quality measures based on administrative data are typically more feasible to implement than those based on the medical record (e.g., whether a patient received the recommended number of psychotherapy visits), these data typically do not capture many important aspects of the process of care (e.g., whether the psychotherapy delivered was evidence-based). Integrating structured documentation into the medical record is one method of increasing the breadth and value of data that are routinely captured. For example, structured documentation could take the form of structured chart notes for an evidence-based psychotherapy, such as cognitive behavioral therapy (CBT). A structured chart note would guide the clinician to input information in specific places and could integrate check boxes or drop-down boxes that would encourage clinicians to document clinical care and capture important qualitative and quantitative data that could be electronically extracted. The Veterans Health Administration (VHA) is currently pilot-testing these kinds of structured chart notes (Karlin and Cross, 2014). In addition, DCoE is currently developing an Alternative Input Method (AIM) form to capture adherence to a PTSD clinical pathway. Although structured documentation increases feasibility of measure implementation and potential data capture, it has important limitations. It is currently unknown how clinicians use structured documentation and whether these approaches accurately capture the care delivered or are tightly linked with clinical outcomes. Natural-language processing (NLP) is another emerging technology that seeks to provide a means for higher-quality data collection without manual medical record review (Shiner et al., 2012). Pilot-testing and evaluating new approaches to structured documentation and data retrieval are important areas for future research.

Recommendation 3: Develop a process for ongoing assessment of quality of care. A single effort to assess quality of care will be of limited use. Instead, assessing the quality of care for PH conditions should be an ongoing process of assessing quality, taking steps to address identified areas for QI, and reassessing quality routinely. As DoD continues to pursue strategies to assess the quality of care, it is essential to consider how these efforts will be sustained and what resources will be required to support ongoing quality measurement.

Definitions of Terms Used to Describe Quality Measures

In this appendix, we describe the terms we use in Appendixes B and C, which provide brief conceptual descriptions of each of the candidate quality measures for PTSD and MDD. These measures are intended to evaluate care for adults 18 years of age or older who are service members. They target primarily outpatient care for PTSD and MDD and apply to both primary and specialty-care settings. The purpose of these descriptions is to provide a summary for each measure, including a description of the focus of the measure, information describing the measure's basic operationalization, and the rationale supporting its inclusion in the measure set. The content of the descriptions of each measure includes the types of information described in this appendix.

Measure Name

The measure name includes the number assigned to the measure, the PH condition to which the measure applies, and a brief title reflecting the measure focus. The number format includes reference to the targeted psychological condition and the stage of care to which it refers. For example, in the measure called PTSD-T16*, *PTSD* indicates a measure contained in the PTSD measure set. *MDD* would refer to inclusion in the MDD set. The letter or letters preceding the measure number refer to the stage of care within the continuum targeted or the type of measure (for those types of measures that extend across all stages of care within the continuum): S = screening, A = assessment, T = treatment, PE = patient experience, and RU = resource use. In addition, an asterisk denotes a measure that can be more broadly applied to psychological conditions other than PTSD and MDD. So PTSD-T16* is a measure of the quality of care for PTSD that assesses care in the treatment phase but could also be more broadly applied to other PH areas besides PTSD.

Measure

This is a brief description of the structure, recommended care, outcome, patient experience, or resource targeted by the measure.

Numerator

The numerator of a quality measure indicates the number of individuals or events from the denominator in which the structure, recommended care, outcome, patient experience, or resource targeted by the measure was received or occurred (e.g., a survey indicating a satisfaction-with-overall-care rating of 9 or higher, a follow-up visit after hospital discharge for a mental health condition).

Denominator

The denominator of a quality measure indicates the number of service members in the population to which the measure applies (e.g., patients with PTSD in a new treatment episode).

Definition

Definition sections provide some detail of how to define a specific element in the measure (e.g., new treatment episode, standardized tool), either in the numerator or denominator. These definitions are broad but are intended to define the basic elements contained in the measure under discussion. The details of these definitions will need to be specified prior to implementation and field-testing of each measure.

Care Continuum

The phases of the continuum of care are the following:

- **Prevention** is the phase during which efforts are made to increase resilience among otherwise-well individuals through the use of specific health interventions to decrease the likelihood of future illness or to identify individuals at greater risk for developing a particular PH condition and provide a targeted intervention to prevent it.
- **Screening** is the phase during which health risk factors or disease states are identified in an individual or in a population. It is generally expected that conditions and risks identified by screening are systematically followed up.
- **Assessment** is the phase during which medical and psychiatric history and physical examinations are completed in order to diagnose patient conditions and identify other factors that may affect the treatment plan (e.g., comorbid conditions, support systems, psychosocial stressors).
- **Treatment** is the phase during which interventions are implemented to treat a diagnosed condition. For any condition, this phase would include initial treatment, assessment of response, and adjustment of treatment as needed.
- **Reintegration** is the phase during which efforts are directed toward maintenance of well-being and the provision of assistance and support services to the individual to adjust and transition back to full duty status.

Measure Type

The measure type characterizes each measure in terms of the aspect of health care delivery that it describes. The measure types are the following:

- A *structure measure* is one that assesses the settings in which health care is provided and the capacity to provide care. These measures may focus on the health care organization or characteristics of individual health care providers. Structure measures assess the adequacy of health facilities and equipment, the qualifications of clinicians and their organizations within the entity of interest, the administrative structure and operations of programs and institutions providing care, and fiscal or payment organization. An example of a structure measure is “number of inpatient psychiatric beds available per 10,000 active-duty service members.”
- A *process measure* assesses whether a recommended care process or event takes place during the care of a condition or the degree to which a procedure or treatment is provided in a manner that reflects fidelity to the evidence base supporting its use. *Processes* generally refers to receipt of a particular service and often within a particular time frame. An example of a process measure is “percentage of MDD patients who received evidence-based psychotherapy for MDD.”
- An *outcome measure* describes the outcome of care in terms of patient improvement, recovery, restoration of function, avoidance of deterioration, or survival. Outcome measures are a key component of measurement-based care, in which repeated assessments of patient outcomes guide accurate, cost-effective treatment decisions to increase the likelihood of achieving optimal patient outcomes (Harding et al., 2011). Outcome measures typically need risk adjustment to avoid biased comparisons of the measure rates. Without risk adjustment, providers or MTFs may appear to achieve poorer patient outcomes when, in fact, they are treating a sicker patient population. Risk adjustment is a “statistical process for reducing, removing, or clarifying the influences of confounding factors that differ among comparison groups” (CMS and Joint Commission, 2012, p. D-11). Risk adjustment can be implemented through the use of logistic regression, stratification, and other statistical methods. An example of an outcome measure is “percentage of PTSD patients in a new treatment episode with improvement in functional status within six months of the initial visit.”
- *Patient experience measures* assess patients’ perceptions of the quality and timeliness of the care they receive and the perceived impact of their care on their health outcomes. An example of a patient experience measure is “percentage of MDD patients who reported that they received treatment quickly.”
- *Resource use measures* relate to the resources expended during the care of a condition (Romano, Hussey, and Ritley, 2010). Resource use can be measured using a variety of approaches, including (1) utilization measures, (2) standardized resource use measures, and (3) cost measures. Measures can be based on either the use of resources during the process of care and recovery or the amount of money spent directly to provide care. An example of a resource use measure is “average costs PMPM.” For the purposes of this report, resource use is included as one type of measure and, therefore, is not discussed separately in the method section. However, recognizing that resource use measures differ

from other types of quality measures, we have included a more detailed discussion of the selected cost measure in Appendix D.

Care Setting

This is the care setting to which the measure applies: outpatient (including urgent or emergency and ambulatory care), inpatient (acute hospital care), or residential.

Measure Source

This is the source from which the measure was identified. In some cases, the measure may have been modified from its original form; when this occurred, the source reference is preceded by “Adapted from the following:” to indicate this change. In some cases, a measure is identified as “new” when no related existing measures were found and a new measure was developed.

Rationale for Measure Inclusion

This section summarizes the empirical evidence supporting the measure and recommendations from CPGs relevant to the measure. In addition, when a measure was adapted from its original, the rationale section includes a description of the necessary changes. Finally, any issues relevant to the need for case-mix adjustment are discussed.

Potential Data Source

These are data sources that can be used to collect data for the numerator and denominator of the measure. For example, data sources can include administrative claims (e.g., ICD-9 and Current Procedural Terminology [CPT] codes), electronic clinical data (e.g., laboratory test results), medical records (paper or EHR), patient surveys (patient self-reported data), and management data (e.g., plan enrollment, clinical staffing information). The particular data source used to implement a measure depends on data availability and data format at the site of measurement.

Feasibility

Basic guidance as to the degree of difficulty associated with operationalizing and applying the measure. Feasibility is ultimately determined by the availability of data at the study site and other data documentation practices. The data options described here are general and may not

apply in all settings. A color-coded system has been created to categorize the general level of difficulty associated with data collection for each measure:

- Green indicates that the measure can be computed using administrative data or other data that are extractable from existing electronic sources (e.g., CPT codes, pharmacy data).
- Yellow indicates that the required data elements for measure computation may require medical record abstraction but could alternatively be integrated into a data-entry tool and the requisite data elements extracted electronically (e.g., screening tool that can be scored that the patient completes at a clinic visit on a computer data portal).
- Orange indicates that the measure computation (numerator or denominator) requires medical record abstraction (e.g., justification for discontinuing or not offering a pharmacologic treatment).

Descriptions of Measures of Quality of Care for Posttraumatic Stress Disorder

Measure Set: Posttraumatic Stress Disorder—PTSD-S1*: Screening and Follow-Up for Common Psychological Health Conditions

Measure

Percentage of patients screened for PTSD, MDD, and alcohol misuse and, if positive, appropriate follow-up initiated¹

Numerator

Patients in the denominator who are

(a) screened annually for the following PH conditions using a standardized tool

- PTSD
- MDD
- alcohol misuse

(b) and, if screened positive, appropriate follow-up was initiated within seven days for any of these:

- PTSD
- MDD
- alcohol misuse

Denominator

(a) All patients

(b) Patients who screened positive for PTSD, MDD, or alcohol misuse

Definitions

Screening

Screening for PTSD. Many standardized tools are available for screening for PTSD. The four-item Primary Care PTSD Screen (PC-PTSD) is the most widely used (IOM, 2012; Prins et al., 2003). Screening tools may be as short as a single item (Gore et al., 2008), or a longer tool may be used that screens for the condition and assesses additional symptoms that can facilitate

¹ The documentation for this measure is similar to that specified for MDD-S1*.

the determination of the presence or absence of a diagnosis of PTSD (e.g., PCL). Examples of validated PTSD tools for adults include the following:

- Davidson Trauma Scale (DTS) (Zlotnick et al., 1996)
- Distressing Events Questionnaire (DEQ) (Kubany et al., 2000)
- five-item primary care anxiety screener (Means-Christensen et al., 2006)
- Impact of Event Scale—Revised (IES-R) (Tehrani, Cox, and Cox, 2002)
- Los Angeles Symptom Checklist (LASC) (King et al., 1995)
- My Mood Monitor (M-3) checklist (Gaynes et al., 2010)
- Mississippi Scale for Combat-Related PTSD (M-PTSD) and M-PTSD for Desert Storm War Zone Personnel (M-PTSD-DS) (Keane, Caddell, and Taylor, 1988)
- Modified PTSD Symptom Scale (MPSS-SR) (Falsetti et al., 1993)
- Penn Inventory for PTSD (Hammarberg, 1992)
- Posttraumatic Adjustment Scale (O'Donnell, Creamer, Parslow, et al., 2008)
- Post-Deployment Health Assessment (PDHA) and Reassessment (PDHRA) (incorporate the PC-PTSD) (Bailey, 1998)
- PC-PTSD (Prins et al., 2003)
- PTSD Brief Screen (Leskin and Westrup, 1999)
- PCL—Civilian Version (PCL-C) and PCL—Military Version (PCL-M) (Blanchard, Hickling, et al., 2003; Weathers, Huska, and Keane, 1991)
- Purdue PTSD Scale—Revised (PPTSD-R) (Vrana and Lauterbach, 1991)
- Revised Civilian Mississippi Scale for PTSD (R-CMS) (Norris and Perilla, 1996)
- Screen for Posttraumatic Stress Symptoms (SPTSS) (Carlson, 2002)
- Short PTSD Rating Interview (Connor and Davidson, 2001)
- Short Screening Scale for *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. (DSM-IV) PTSD (Breslau et al., 1999)
- single-item PTSD Screener (Gore et al., 2008)
- Startle, Physiological Arousal, Anger, and Numbness (SPAN) (J. Davidson, 2002)
- Trauma Screening Questionnaire (Brewin et al., 2002)
- Trauma Symptom Checklist—40 (TSC-40) (Briere, 1996)
- Trauma Symptom Inventory (TSI) (Briere et al., 1995)
- two-item and six-item PCLs (Lang and Stein, 2005).

Positive screen for PTSD. What constitutes a positive screen depends on the standardized screening tool used. The PC-PTSD, for example, is a four-item screen that was designed for use in primary care and other health care settings and is currently used to screen for PTSD in veterans at VA (Prins et al., 2003). 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.

Screening for MDD. There are several standardized tools used for screening for depression. Screening tools may be quite short with one or two items (e.g., the two-item Patient Health Questionnaire [PHQ-2]), or a longer tool may be used that screens and assesses addi-

tional symptoms that can facilitate the determination of the presence or absence of a diagnosis of MDD (e.g., PHQ-9). Examples of validated MDD tools for adults include the following:

- Beck Depression Inventory (BDI, BDI-II) (Beck, Steer, and Brown, 1996; Beck, 1988)
- Center for Epidemiologic Studies Depression Scale (CESD) (Devins et al., 1988)
- Geriatric Depression Scale (GDS) (Lyness et al., 1997)
- Duke Anxiety–Depression Scale (DUKE-AD) (Parkerson, Broadhead, and Tse, 1996)
- Hamilton Rating Scale for Depression, 17-Item (HRSD-17) (Hamilton, 1960)
- Hamilton Rating Scale for Depression, 21-Item (HRSD-21) (Hamilton, 1967)
- Hamilton Rating Scale for Depression, 24-Item (HRSD-24) (J. B. W. Williams, 1988)
- self-rated 30-item version of the Inventory of Depressive Symptomatology (IDS-SR₃₀) (Rush, Gullion, et al., 1996)
- Montgomery–Åsberg Depression Rating Scale (MADRS) (Svanborg and Åsberg, 2001)
- Medical Outcomes Study Depression Questionnaire (MOSDQ) (Burnam et al., 1988)
- PHQ-2 (Kroenke, Spitzer, and Williams, 2003)
- PHQ-9 (Spitzer, Kroenke, and Williams, 1999)
- Primary Care Evaluation of Mental Disorders (PRIME-MD) (Spitzer, Kroenke, and Williams, 1999)
- 16-item Quick Inventory of Depressive Symptomatology (QIDS-SR₁₆) (Rush, Trivedi, Ibrahim, et al., 2003)
- Self-Stigma of Depression Scale (SSDS) (Barney et al., 2010)
- Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) MDD module (First et al., 1996)
- Zung Self-Rating Depression Scale (Gabrys and Peters, 1985).

Positive screen for MDD. What constitutes a positive screen depends on the standardized screening tool used. For example, the PHQ-2 score ranges from 0 to 6, with a positive screen designated as a score greater than 2 (Kroenke, Spitzer, and Williams, 2003).

Screening for alcohol misuse. There are several screening tools used for screening for misuse of alcohol. The following are recommended as brief, feasible screening tools (Bradley and Berger, 2013):

- Alcohol Use Disorders Identification Test (AUDIT)—Consumption (AUDIT-C) (Bradley, DeBenedetti, et al., 2007)
- AUDIT (Babor and Grant, 1989)
- Single Alcohol Screening Question (SASQ) (Seale et al., 2006).

Positive screen for alcohol misuse. What constitutes a positive screen depends on the standardized screening tool used. For example, the AUDIT-C score ranges from 0 to 12 with a positive screen designated as a score of at least 5 within VA medical centers. Although previous work has indicated that two or three points for women and four points for men may balance sensitivity in most settings (Bradley and Berger, 2013), VA adopted the higher cut point to decrease the burden of false positives on providers (Lapham et al., 2012).

Follow-Up for Positive Screen

Received appropriate follow-up. The elicitation of a positive screen for PTSD, MDD, or alcohol misuse requires further assessment of the patient to determine whether the patient has the condition in question. This requires one or more of the following:

- additional evaluation: This could be the administration of a standardized tool designed to assess for the presence or absence of the condition (e.g., PHQ-9 for depression, PCL for PTSD, or AUDIT for alcohol misuse or abuse) or a structured diagnostic interview that resulted in the determination of the presence or absence of the condition.
- referral to a practitioner qualified to diagnose and treat the condition in question
- pharmacologic intervention
- other interventions or follow-up appropriate to the treatment of the condition in question.

Care Continuum

Screening

Measure Type

Process

Care Setting

Outpatient

Measure Sources

Adapted from the following:

NQF, “NQF #0418 Screening for Clinical Depression,” last update September 23, 2011b.

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

VA, *2012 Performance and Accountability Report: Honoring All Who Have Served*, 2012a. As of May 7, 2014: <http://www.va.gov/budget/report/>

Rationale for Measure Inclusion

Among active-duty service members, PTSD, MDD, and alcohol use disorders are among the most common psychiatric conditions, and all have been linked to combat deployments (Kessler, Sonnega, et al., 1995; Schell and Marshall, 2008). Moreover, these conditions frequently go undiagnosed, and all are associated with functional disability that results in significant personal and societal costs (Allison-Aipa et al., 2010; Gahm and Lucenko, 2008; Gerrity, Corson, and Dobscha, 2007; Thomas et al., 2010; E. C. Williams et al., 2012). Screening, defined as the examination of an otherwise-healthy population to identify people who are at higher risk for a condition (Morrison, 1992), is recommended when the identified condition is an important health problem, the screening instrument is valid, and adequate resources are available to ensure that someone who screens positive receives a thorough assessment and, if needed, follow-up care (IOM, 2012; Rona, Hyams, and Wessely, 2005). Given the prevalence of PTSD, MDD, and alcohol use disorders among service members, and significant disability associated with each, population screening should be considered (Post-Deployment Health Guideline Expert Panel, 2001). Note that, although screening for common postdeployment

conditions is recommended, screening should always be conducted in concert with adequate clinical resources to ensure that every patient who screens positive for a condition receives a timely and thorough diagnostic assessment and that adequate staffing is in place to provide those who are diagnosed with a condition with guideline-appropriate treatment (U.S. Preventive Services Task Force, 2009).

Posttraumatic Stress Disorder

VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress (Management of Post-Traumatic Stress Working Group, 2010) rates the evidence as only fair in support of the claim that screening for PTSD improves health outcomes and that the cost-benefit ratio is favorable. Similarly, the IOM cautions that, although it is generally accepted that early screening for psychological disorders and subsequent treatment would improve outcomes, the evidence to support this claim is not yet strong (IOM, 2012). Nonetheless, *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* currently recommends that all new patients be screened for PTSD and then subsequently screened annually, and it assigns the strength of this recommendation a grade of B. This corresponds to a recommendation that clinicians provide the service to eligible patients and a finding of at least fair evidence that the intervention improves health outcomes. Of the many screening and assessment tools available to assess PTSD symptoms, the VA/DoD guideline for care recommends one of the following four validated scales: PC-PTSD (Prins et al., 2003), PTSD Brief Screen (Leskin and Westrup, 1999), Short Screening Scale for DSM-IV PTSD (Breslau et al., 1999), or the PCL-M (Blanchard, Jones-Alexander, et al., 1996; Weathers, Huska, and Keane, 1991). Note that the PC-PTSD is among the most commonly used of the screening instruments and has been embedded in DoD's PDHA and PDHRA (IOM, 2012).

Major Depressive Disorder

Clinical Practice Guideline: Management of Major Depressive Disorder (MDD) (Management of MDD Working Group, 2009) recommends that the PHQ-2 (Kroenke, Spitzer, and Williams, 2003) be completed annually by all patients seen in primary care settings. This recommendation is provided a grade of A, indicating a strong recommendation with good evidence that the treatment improves important health outcomes (Management of MDD Working Group, 2009). As in PTSD, it is critical that screening be paired with the resources to follow up any positive screen with a thorough assessment that includes both a diagnostic assessment and a suicide-risk evaluation.

Alcohol Misuse

Clinical Practice Guideline: Management of Substance Use Disorders (Management of Substance Use Disorders Working Group, 2009) recommends routine annual screening for alcohol abuse for all general health and mental health patients. The authors give the level of evidence a grade of A, which corresponds to a strong recommendation with good evidence that the treatment improves important health outcomes. Providers are encouraged to choose one of the two following brief screens: AUDIT-C (Bradley, DeBenedetti, et al., 2007) or the SASQ (Seale et al., 2006). Note that the commonly used cut down, annoyed, guilty, and eye-opener (CAGE) screener is not considered an appropriate screen for past-year risky or hazardous drinking because the items focus on lifetime use and consequences (Management of Substance Use Disorders Working Group, 2009; Ewing, 1984).

This measure was based on an NQF-endorsed measure (0418), which recommends annual screening for depression (NQF, undated [a]); screening measures for substance use and comorbid psychiatric conditions included in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011); and a postdeployment health care expert panel recommending screening for common postdeployment conditions (Post-Deployment Health Guideline Expert Panel, 2001).

Potential Data Sources***Numerator***

For (a) and (b): patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominators

For (a): administrative claims. For (b): patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

The (a) denominator for this measure (patients seen in outpatient care) can be identified most feasibly through administrative claims. The (b) denominator (patients who screened positive for screened conditions) would require electronic access to the screening result or medical record review. The numerator also has two parts: (a) screening with a standardized tool and, if positive, (b) appropriate follow-up. Screening is most feasible when using a brief, standardized tool that is quantitative and has a clear definition of a positive versus negative screen. For this compound measure, brief screens for the three conditions can be combined for efficiency of screening. Quantitative scores from standardized tools that are entered into an EHR with extractable data fields or by patients via a web-based portal can facilitate identifying the numerator for screening from the electronic clinical data. The confirmation of appropriate follow-up for a positive screen is more complex in that multiple actions can satisfy this part of the measure. However, the administration of a brief screening tool followed by a more thorough assessment tool for cases that are positive (e.g., positive PHQ-2 followed by PHQ-9 or the use of PHQ-9 for screening and assessment) satisfies the follow-up requirement, and the data may be very feasible to access electronically. Other actions that satisfy the implementation of appropriate follow-up (e.g., referral) would likely require medical record review.

Feasibility Code, Screening

Green to yellow

Feasibility Code, Appropriate Follow-Up

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-A1: Baseline Symptom Assessment with Posttraumatic Stress Disorder Checklist

Measure

Percentage of PTSD patients in a new treatment episode with assessment of symptoms with the PCL

Numerator

Patients in the denominator who have an assessment of PTSD symptoms within the first 30 days of a new treatment episode using the PCL

Denominator

Patients with PTSD in a new treatment episode

Definitions

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

PCL. The PCL is a 17-item self-report measure of symptoms of PTSD (Blanchard, Jones-Alexander, et al., 1996). Slight variants of the scale exist, all of which are scored identically. The PCL-C is the general version (civilian), and scale items refer to a “stressful experience from the past.” The PCL-M is the military version, and items refer to a “stressful military experience.” The PCL-S (PCL—Specific) limits responses to one particular stressful event by requiring the respondent to nominate a single stressful event; subsequent items refer to “the stressful experience.”

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

Measurement of PTSD symptoms at the start of care using a standardized instrument allows clinicians to track treatment response quantitatively and, when necessary because of treatment nonresponse, to adjust the treatment plan. In order to make a determination of symptom improvement (or nonresponse) at a future time point, a baseline assessment of symptoms is necessary.

Fontana and Rosenheck (1994) have recommended that any symptom-tracking instrument assess PTSD symptoms across multiple domains of functioning. Full coverage of symptoms and types of functional impairment may improve the sensitivity of the instrument to treatment response. Indeed, Greenberg, Rosenheck, and Fontana (2003) showed that the assessment of PTSD symptoms using one of two standardized scales was related to treatment outcomes.

The VA/DoD CPGs (Management of Post-Traumatic Stress Working Group, 2010) cite evidence supporting thorough assessment of PTSD symptoms for patients in both primary and mental health specialty-care settings (Lagomasino, Daly, and Stoudemire, 1999; E. R. Williams and Shepherd, 2000) but do not specifically recommend that a standardized assessment instrument be used in all cases. However, the VHA Mental Health Program Evaluation Consultation Group has recommended that standardized PTSD symptom-assessment instruments be used consistently across all VHA care services for patients with PTSD. The consultation group noted that medical record chart data are unreliable for tracking PTSD symptoms and outcomes and therefore recommended that a quality indicator measuring the assessment of PTSD with a standardized instrument be developed. This measure, developed by RAND researchers for the VHA, was a response to that call (Farmer et al., 2010).

We recommend that the PCL (Blanchard, Jones-Alexander, et al., 1996) be the standardized measurement tool for this purpose for a variety of reasons. First, the PCL has been validated with active-duty service members (Bliese et al., 2008). Importantly, the psychometric properties of the PCL are strong, with good internal consistency ($\alpha = 0.94\text{--}0.97$) and reliable scale scores across short test-retest periods ($r = 0.88\text{--}0.96$) (Blanchard, Jones-Alexander, et al., 1996; Ruggiero et al., 2003; Weathers, Litz, et al., 1993). Note that long-measurement-period reliability is neither expected nor desirable for a measure designed to be sensitive to symptom change over time. In fact, to the contrary, it is important that measures employed for this purpose are sensitive to symptom change in response to treatment—a criterion passed by the PCL (Monson et al., 2008). Validity of the PCL as an indicator of DSM-IV diagnosis and symptom strength is strong, with the PCL being nearly collinear ($r = 0.93$) with the Clinician-Administered PTSD Scale (CAPS), the gold-standard measure for psychiatric diagnosis (Blanchard, Jones-Alexander, et al., 1996). Moreover, there is good convergent validity with other validated measures of PTSD ($r = 0.77\text{--}0.93$) (Blanchard, Jones-Alexander, et al., 1996; Weathers, Litz, et al., 1993). Of note is the fact that the diagnostic criteria for PTSD were updated in the 2013 revision of *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. (DSM-V) (APA, 2013). At the time of this report, the National Center for PTSD was working to revise the PCL to reflect these changes (National Center for PTSD, 2014). It will be important to track the outcome of its efforts in order to determine the ongoing suitability of the instrument and the need to update items or scoring protocols. This measure was adapted from the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011). What constitutes a break in care was changed from five months to six months to match the time frame that is more generally used. Additionally, the requirement

for a six-month break in PTSD-related medication was deleted because of the risk of eliminating patients from the denominator who were being treated for conditions other than PTSD because PTSD-related medications may be used to treat other conditions. Although that evaluation accepted using any one of many standardized tools available for this assessment, we are recommending the use of the PCL to establish an objective, baseline score. This score becomes an essential part of the clinical data that will be used to monitor the patient's response to treatment over time. Because of the popularity of the PCL, it is recommended as the standardized tool to be utilized with this measure.

Potential Data Sources***Numerator***

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure (a new treatment episode of PTSD) can be derived from administrative claims data (most easily) or from medical records. However, the determination of the assessment of PTSD symptoms using the PCL could most feasibly be accessed if the PCL were recorded in EHR-extractable data fields or collected from patients via a web-based portal. Without electronic access to these scores, the more labor-intensive medical record review would be required.

Feasibility Code

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-A2: Assessment for Depression

Measure

Percentage of PTSD patients in a new treatment episode assessed for depression

Numerator

Patients in the denominator who are assessed for comorbid depression within 30 days of the new treatment episode

Denominator

PTSD patients in a new treatment episode

Definitions

Assess for depression. Any documentation of the presence or absence of depression (including a depression diagnosis) or any assessment of mood, either by formal assessment (standardized tool) or by interview. Standardized tools for screening and assessment include the following:

- BDI and BDI-II (Beck, Steer, and Brown, 1996; Beck, 1988)
- CESD (five-, ten-, or 20-item version) (Radloff, 1977)
- HRSD (Hamilton, 1960, 1967; Radloff, 1977)
- IDS-SR₃₀ (Rush, Gullion, et al., 1996)
- MADRS (Montgomery and Asberg, 1979)
- MOSDQ (Burnam et al., 1988)
- PHQ-2 (Kroenke, Spitzer, and Williams, 2003)
- PHQ-9 (Spitzer, Kroenke, and Williams, 1999)
- QIDS-SR₁₆ (Rush, Trivedi, Ibrahim, et al., 2003)
- SSDS (Barney et al., 2010)
- SCID MDD module (First et al., 1996).

Informal assessment includes documentation of the presence or absence of depressive symptoms (e.g., sad mood, suicidal thoughts, hopelessness).

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

There is considerable comorbidity between PTSD and MDD (Erickson et al., 2001; O'Donnell, Creamer, and Pattison, 2004; Perlman et al., 2011). A 2008 study found that, among service members with probable PTSD, two-thirds also screened positive for depression (Schell and Marshall, 2008). According to *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress*, “co-morbid medical and psychiatric conditions are important to recognize, because they can modify clinical determinations of prognosis, patient or provider treatment priorities, selection of interventions, and the setting where PTSD care will be provided” (Management of Post-Traumatic Stress Working Group, 2010).

Although the VA/DoD guidelines recommend assessing a range of psychiatric comorbidities (Management of Post-Traumatic Stress Working Group, 2010), this indicator has been developed to address only depression. Depression is the most prevalent psychiatric comorbidity found in populations with PTSD, and standardized instruments for assessing depression are available to facilitate reliable and valid assessments (Gahm and Lucenko, 2008; IOM, 2012).

This measure was adapted from the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011). What constitutes a break in care was changed from five months to six months to match the time frame that is more generally used. Additionally, the requirement for a six-month break in PTSD-related medication was deleted because of the risk of eliminating patients from the denominator who were being treated for conditions other than PTSD because PTSD-related medications may be used to treat other conditions.

Potential Data Sources**Numerator**

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure applies to PTSD patients in a new treatment episode and can be identified with administrative claims data. The performance of the screen for depression could most feasibly be accessed if the score on a standardized depression instrument (e.g., PHQ-9) is recorded in EHR-extractable data fields or collected from patients via a web-based portal. Other means of documenting screening for depression (e.g., clinical notation of their

presence or absence, use of a paper record) would require medical record review, which makes that option the most labor-intensive.

Feasibility Code

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-A3*: Assessment for Suicide Risk

Measure

Percentage of patients in a new treatment episode for PTSD assessed for suicide risk²

Numerator

Patients in the denominator who are assessed for current suicide risk during the same visit in which a new treatment episode began or in the 14 days prior

Denominator

Patients in a new treatment episode for PTSD

Definitions

Assessment of suicide risk. Suicide risk assessment must include questions about the following:

- SI
- patient’s intent to initiate a suicide attempt

and, if either is present,

- patient’s plans for a suicide attempt
- whether the patient has means for completing suicide.

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

SI. SI includes any reference to the patient not wanting to live anymore, comments about killing oneself or doing oneself serious harm, passing thoughts of death, or similar thoughts. Absence of SI is documentation of specific denial of SI (e.g., “no suicidal thoughts,” “denies SI”). Using the PHQ-9, which includes an item assessing SI, would count for assessment of SI, but a PHQ of fewer items would not.

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

² The documentation for this measure is similar to that specified for MDD-A3*.

Measure Sources

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NQF, “NQF #0104 Adult Major Depressive Disorder: Suicide Risk Assessment,” last updated March 15, 2013i.

Rationale for Measure Inclusion

Given the increased risk of attempted and completed suicide associated with most psychiatric conditions (Cavanagh et al., 2003; Kelly and Mann, 1996), it is important for providers to assess SI among new or returning patients and, when present, to implement a safety plan and begin quality mental health services (Ramchand, Acosta, et al., 2011). Case-control studies show that one-half to three-quarters of all suicides can be attributed to psychiatric disorders, typically mood and anxiety disorders (Cavanagh et al., 2003). MDD, the most strongly related disorder, increases the risk for death by suicide by 20 times relative to the general population (Cavanagh et al., 2003; Harris and Barraclough, 1997). Among anxiety disorders, PTSD is the most tightly linked with SI (Kessler, Borges, and Walters, 1999; Sareen et al., 2005). The demographic profile of active-duty service members (younger and more likely to be male than the civilian population) also matches the demographic risk factors for completed suicide (Goldsmith et al., 2002; McKeown, Cuffe, and Schulz, 2006). For these reasons, it is important that every patient with a new treatment episode for a PH condition be assessed for suicidal risk.

This measure is based on the NQF-endorsed measure 0104, which recommends screening for suicide risk for any patient with a new treatment episode of MDD (NQF, undated [a]). The NQF measure has been expanded to be applied to any patient with a new treatment episode for PTSD, which is consistent with *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide*, as well as VA/DoD guidelines for the treatment of substance use disorders, PTSD, MDD, and psychosis (Management of MDD Working Group, 2009; Management of Bipolar Disorder Working Group, 2010; Management of Substance Use Disorders Working Group, 2009; Assessment and Management of Risk for Suicide Working Group, 2013). The measure used in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011) looked for an annual assessment of SI, whereas this measure looks for a suicide risk assessment at the time of a new treatment episode. Assessing SI is a routine part of the mental status exam conducted in psychiatry, and APA recommends that it be used as part of standard practice (Work Group on Psychiatric Evaluation, 2006). This recommendation received a grade of I, which indicates that it was “recommended with substantial clinical confidence.” This measure’s required components for a suicide risk assessment (ideation, intent, plans, and means) is consistent with recommendations in *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide* (Assessment and Management of Risk for Suicide Working Group, 2013).

The VA/DoD guideline also recommends that treatment providers consider nonmodifiable risk factors for suicide (e.g., younger age, male gender, family history of suicide, same-sex orientation) and modifiable risk factors (e.g., unstable housing, financial problems, psychiatric disorders) in order to determine whether the relative risk of a completed suicide is low, intermediate, or high (Assessment and Management of Risk for Suicide Working Group, 2013).

These risk factors and ultimate risk status are not included in the current measure. Determining acute risk status for suicide (low, intermediate, or high) requires complex clinical judgment; integrating all risks into a single acute risk category would be difficult to perform reliably or consistently with the clinician responsible for the patient's clinical care. It is our judgment that instantiating these guidelines into a quality measure will require a record of the *clinician's judgment* of the patient's risk category. Such a record of the clinician's judgment is not currently a field in the EHR and would therefore require medical record review. That being said, we suspect that even a medical record review would reveal that not all guideline-specified risk factors are documented in the record. However, we also believe that, as these recently released guidelines are promulgated, it is possible that a field will be added to the EHR requiring providers to indicate—when a patient is positive for SI—whether the acute risk of an attempt is low, intermediate, or high. Were this to occur, the quality measure in place should be updated to include this field in the criteria for passing the measure.

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure can be defined with administrative claims data or by medical record review. The numerator for this measure would typically require medical record review. Although it would be possible to identify cases in the numerator that were negative for suicide risk (e.g., negative response to PHQ-9 item addressing SI) using electronic clinical data (e.g., if the PHQ-9 item responses were recorded in extractable fields of an electronic medical record or by a patient via a data portal), any case in which such a specific tool was not used or the screen was positive (and, therefore, required further assessment of intent and plan) would require medical record review.

Feasibility Code

Orange

Measure Set: Posttraumatic Stress Disorder—PTSD-A4*: Assessment of Recent Substance Use

Measure

Percentage of PTSD patients in a new treatment episode assessed for recent substance use³

Numerator

Patients in the denominator who have an assessment of recent substance abuse, including type, quantity, and frequency, within the first 30 days of a new treatment episode for PTSD

Denominator

Patients in a new treatment episode for PTSD

Definitions

Assessment of use. Documentation of no recent alcohol and no recent drug use *or* documentation of recent alcohol or drug use, including type, quantity, and frequency for all substances used. An appropriate screening tool may be used.

- type: An assessment of alcohol, marijuana, cocaine, heroin or other opiates, amphetamine or methamphetamine, or note indicating that the patient denied all other substance use
- quantity (for alcohol only): Any evidence of a quantity assessment, including number of drinks per day, number of drinks per week, any note about binge drinking (at least five drinks in one drinking episode for men, at least four drinks in one drinking episode for women)
- frequency: Note about daily, monthly, weekly, or occasional use.

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

Recent use. Use in the past three months

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

³ The documentation for this measure is similar to that specified for MDD-A4*.

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NQF, “NQF #0110 Bipolar Disorder and Major Depression: Appraisal for Alcohol or Chemical Substance Use,” last updated February 6, 2013c.

Rationale for Measure Inclusion

Mental health patients who are currently using alcohol or other drugs do not respond to treatment as well as patients who are not using alcohol or drugs (Le Fauve et al., 2004). Moreover, the impairment associated with their mental health conditions appears to be more severe and chronic than for patients without concurrent substance use (Kessler, 2004). *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* and the 2009 *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* both recommend that current substance use patterns of patients with these disorders be assessed in order to identify substance abuse or dependency, including alcohol, nicotine, and prescribed and illicit drugs (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010).

This measure was modified from its use in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011) in a variety of ways. First, what constitutes a break in care was changed from five months to six months to match the time frame that is more generally used. Additionally, the requirement for a six-month break in PTSD-related medication was deleted because of the risk of eliminating patients from the denominator who were being treated for conditions other than PTSD because most PTSD-related medications may also be used to treat other conditions.

There is a similar measure endorsed by NQF (0110), which recommends assessing comorbid alcohol and substance use in patients with bipolar or unipolar depression (NQF, undated [a]). We have expanded the NQF measure to apply to any patient with a new treatment episode for PTSD because of the prevalence of comorbid alcohol and drug misuse associated with other PH conditions (Kessler, 2004).

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims

Feasibility

The denominator for this measure applies to any patient with a new treatment episode for PTSD and can be identified with administrative claims data. Determining whether an assessment of recent substance use had occurred could be completed using patient self-reported survey data or medical record review. Portions of this assessment could be accessible from electronic clinical data if, for example, a web-based patient portal were developed and used to collect screening information such as this or documented in extractable fields of an EHR. More

likely, manual medical record review would be used to determine whether clinical documentation of an assessment of recent substance use were present and whether it covered the required aspects of assessment.

Feasibility Code

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T1: Periodic Symptom Assessment with the PTSD Checklist

Measure

Percentage of PTSD patients with symptom assessment with the PCL during the four-month measurement period

Numerator

Patients in the denominator who have the PCL administered at least once during the four-month measurement period

Denominator

Patients with the diagnosis of PTSD and an encounter within each four-month period

Definitions

Four-month measurement period with an encounter. Time window in which the PTSD patient is either seen at an office visit or contacted via another method (phone, email) during a four-month time period defined by dates of service that fall into that time period (e.g., June 1, 2012, to September 30, 2012)

PCL. The PCL is a 17-item self-report measure of symptoms of PTSD (Blanchard, Jones-Alexander, et al., 1996). Slight variants of the scale exist, all of which are scored identically. The PCL-C is the general (civilian) version, and scale items refer to a “stressful experience from the past.” The PCL-M is the military version, and items refer to a “stressful military experience.” The PCL-S (specific) limits responses to one particular stressful event by requiring the respondent to nominate a single stressful event; subsequent items refer to “the stressful experience.”

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

Adapted from the following:

NQF, “NQF #0712 Depression Utilization of the PHQ-9 Tool,” last updated April 24, 2013k.

Rationale for Measure Inclusion

This measure is based on clinical care recommendations in *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010). The guideline recommends “regular follow-up with monitoring and documentation of symptom status” in the treatment of PTSD in both primary care and specialty mental health settings. In discussing the regularity of monitoring, the guideline recommends that patients be

assessed at every treatment visit and encourages clinicians to consider a validated measure, such as the PCL. Comprehensive reassessments and evaluations should occur

every three months after initiating treatment for PTSD, in order to monitor changes in clinical status and revise the intervention plan accordingly. The interval of three months is suggested because many controlled trials of first line therapies for PTSD recommended in this guideline demonstrate clinically significant changes during this time frame. (Management of Post-Traumatic Stress Working Group, 2010, p. 94)

There is an increasing emphasis on the need to deliver care that is evidence-based and effective. Harding and colleagues (2011) make the case for measurement-based care as the standard for psychiatric practice to align with physical health care. Standardized, repeated measurement of PTSD symptoms allows clinicians to track individual patient response to treatment and allows administrators and organizations to monitor the treatment outcomes of larger patient groups. Greenberg, Rosenheck, and Fontana (2003) have shown that standardized assessment of PTSD symptoms is related to PTSD treatment outcomes. Elsewhere, Fontana and Rosenheck (1994, p. 407) addressed the importance of using standardized instruments to assess PTSD symptoms across “multiple domains of functioning, while at the same time minimizing the overall length of the data collection protocols.”

This measure was adapted from the NQF measure for MDD that dictates monitoring response to treatment over time using the PHQ-9 (NQF, undated [a]). The suggested frequency of reassessment is at least once during every four-month interval that includes a patient encounter. We applied this model of use of the PHQ-9 to the regular use of the PCL to monitor response to treatment for PTSD that is objective and quantitative and that can be used to assess ongoing treatment response. This indicator reflects the NQF recommendations for MDD as applied to any PTSD-diagnosed patient in a new treatment episode.

For a variety of reasons, we recommend that the PCL (Blanchard, Jones-Alexander, et al., 1996) be the standardized measurement tool for this purpose. First, the PCL has been validated with active-duty service members (Bliese et al., 2008). Importantly, the psychometric properties of the PCL are strong, with good internal consistency ($\alpha = 0.94\text{--}0.97$) and reliable scale scores across short test-retest periods ($r = 0.88\text{--}0.96$) (Blanchard, Jones-Alexander, et al., 1996; Ruggiero et al., 2003; Weathers, Litz, et al., 1993). Note that long-measurement-period reliability is neither expected nor desirable for a measure designed to be sensitive to symptom change over time. In fact, to the contrary, it is important to establish that measures employed for this purpose are sensitive to symptom change in response to treatment—a criterion passed by the PCL (Monson et al., 2008). Validity of the PCL as an indicator of DSM-IV diagnosis and symptom strength is strong, with the PCL being nearly collinear ($r = 0.93$) with the CAPS, the gold-standard measure for psychiatric diagnosis (Blanchard, Jones-Alexander, et al., 1996). Moreover, there is good convergent validity with other validated measures of PTSD ($r = 0.77\text{--}0.93$) (Blanchard, Jones-Alexander, et al., 1996; Weathers, Litz, et al., 1993). Of note is the fact that the diagnostic criteria for PTSD were updated in the 2013 revision of the DSM (APA, 2013). At the time of this report, the National Center for PTSD was working to revise the PCL to reflect these changes (National Center for PTSD, 2014). It will be important to track the outcome of its efforts in order to determine the ongoing suitability of the instrument and the need to update items or scoring protocols.

Potential Data Sources***Numerator***

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure (patients with PTSD and encounters during the measurement period) can be calculated from administrative claims data, depending on what types of encounters are accessible. However, the determination of the use of the PCL to assess severity and treatment response requires either medical record abstraction or the documentation of a PCL score that is data accessible. Examples of data accessibility would be an EHR with an extractable data field designated for the documentation of the PCL score or a patient data portal used for completion of the PCL tool. Lacking electronic access to these data elements would necessitate the use of the more labor-intensive medical record review.

Feasibility Code

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T2*: Availability of Specialty-Care Visits

Measure

Percentage of days when third available specialty-care appointment is within two days⁴

Numerator

Days in the measurement period on which the third available specialty-care visit is within two days (48 hours)

Denominator

Days in the measurement period

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

Third available appointment. Using a paper or electronic scheduling assistant, for a dummy patient in a given measurement point, the new or return specialty-care appointment that is third in a sequential list of available specialty-care appointments

Care Continuum

Treatment

Measure Type

Structure

Care Setting

Outpatient

Measure Source

Institute for Healthcare Improvement, “Third Next Available Appointment,” undated; referenced 2011. As of September 18, 2013:
<http://www.ihl.org/knowledge/Pages/Measures/ThirdNextAvailableAppointment.aspx>

Rationale for Measure Inclusion

Long wait times for care can interfere with treatment engagement (Festinger et al., 1995; Gallucci, Swartz, and Hackerman, 2005; MacDonald, Brown, and Ellis, 2000). If a first appointment is scheduled weeks after an initial call, a patient who was prepared to address his or her substance use or psychiatric symptoms at the time he or she contacted a facility may no longer be motivated to engage in treatment. For this reason, many treatment facilities strive to ensure that specialty-care appointments are scheduled to occur quickly after patients call to request care.

One measurement strategy to gauge the availability of care is to record the next available appointment at a given measurement point. For example, a dummy patient can be entered

⁴ The documentation for this measure is similar to that specified for MDD-T2*.

into a scheduling system, and the first appointment available for that patient is recorded. This strategy to measure treatment availability is not ideal. Same-day cancellations and other unexpected events can create openings in a clinic's schedule that arbitrarily and uncharacteristically make it appear that specialty-care appointments are readily available. Moreover, a real patient may be unable to attend a first available appointment that occurs within hours of his or her call. Thus, we recommend that the third available appointment be used instead as a more accurate and sensitive measure of care availability (Institute for Healthcare Improvement, undated; Oldham, 2001). A third-available-appointment measure of treatment availability reduces the likelihood of chance and unexpected occurrences from the measure (Institute for Healthcare Improvement, undated).

To make the distinction between first-available and third-available measures more concrete, imagine two very different clinics. Clinic A is overstaffed, fills only 70 percent of its available appointments and can nearly always provide a new patient with an intake appointment on the day that the patient prefers. Clinic B operates with significant staff shortages, filling 100 percent of its available appointments, and new patients experience long delays. Both clinics typically have one patient who cancels his or her appointment on a given day. Now imagine that an administrator wanted to determine wait times for specialty care and decided to use the first available appointment as an indicator. Clinic A would have 30 percent of its appointments unfilled on the first day and would be coded with a wait time of one day or less. However, clinic B would also have an appointment available on the first day (a patient canceled that day) and would be coded as having a wait time of one day or less. That is, using a first-available-appointment indicator, both clinics would pass and would appear identical. However, from the patient's perspective, these two clinics are dramatically different. When a patient calls clinic A, he or she would be offered many different appointment times from which to choose; a patient who calls clinic B would be offered a same-day appointment that the patient is unlikely to be able to attend (because of travel times, work schedules, or child-care interference). When the patient declines the same-day cancellation appointment, the clinic's 100-percent schedule rate would mean that the next unscheduled appointment may be weeks or even months away. In some ways, same-day cancellations are noise in the system that makes it difficult for the administrator to observe real wait times. However, if *three* appointments were available within two days, it is likely that the clinic consistently maintains some immediate availability for new patients. Clinics with full booking and long wait times are much less likely to have three appointments available within two days and are therefore more likely to be identified using a third-available-appointment measure.

Potential Data Sources

Numerator

Management data

Denominator

Management data

Feasibility

Data collection can be completed electronically or manually. For manual data collection, the schedule is searched forward from the index day to count to the day of the third avail-

able appointment. An electronic scheduling system may allow searches to be conducted and recorded automatically, which would make the measure quite feasible to implement.

Feasibility Code

Green to yellow

Measure Set: Posttraumatic Stress Disorder—PTSD-T3*: Appropriate Follow-Up for Endorsed Suicidal Ideation

Measure

Percentage of patient contacts with SI with appropriate follow-up⁵

Numerator

Documentation of appropriate follow-up for the SI, intent, or behavior

Denominator

Outpatient visits or contacts in which the PTSD patient endorsed SI, intent, or behavior

Definitions

Appropriate follow-up. Appropriate follow-up is specific to the patient’s presentation with regard to SI with or without intent and suicidal behavior.

For positive SI and *negative* for intent:

- provision of resource list (outpatient): patient given a list of resources to call or visit if in danger
- appointment for follow-up

For positive suicidal ideation and *positive* for intent:

- provision of a family intervention or patient referral for hospitalization

For suicidal behavior:

- patient referral for hospitalization

Family intervention. Documented provider discussion with patient and family members regarding the patient’s SI and strategies to keep the patient safe

SI. Any reference to the patient not wanting to live anymore, comments about killing oneself or doing oneself serious harm, and thoughts of death as a solution

Suicidal behavior. Any attempt to kill oneself. It includes attempted suicide and suicidal gestures (a suicidal action unlikely to be being fatal).

Suicidal intent. Documentation indicating imminent threat of suicide: Patient has a specific plan to kill him- or herself (e.g., location, timing) or has selected and has access to the means for suicide (e.g., pills, firearms).

Care Continuum

Treatment

Measure Type

Process

⁵ The documentation for this measure is similar to that specified for MDD-T3*.

Care Setting

Outpatient

Measure Source

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

Given that SI may predict suicidal behavior, it is important that providers with patients who endorse these thoughts provide immediate and appropriate follow-up care to reduce their patients' risk.

For patients who are actively suicidal, inpatient psychiatric hospitalization is a common prevention measure to ensure their safety. Although hospitalization typically prevents suicide during the stay, hospitalization alone has not been demonstrated to reduce the risk of suicide following discharge (Goldsmith et al., 2002; Work Group on Suicidal Behaviors, 2003). Rather, specific interventions that are conducted during the inpatient stay are the key (Brown et al., 2005; Linehan et al., 2006). Unfortunately, hospital stays are often too short to allow any specific intervention to be delivered (Goldsmith et al., 2002). Nonetheless, without other strategies to keep a patient safe who poses a short-term danger to him- or herself, hospitalization may be an appropriate strategy (Assessment and Management of Risk for Suicide Working Group, 2013; Work Group on Suicidal Behaviors, 2003).

An alternative is to develop a safety plan with a suicidal patient and his or her family and support network. These plans, often written up and signed as “contracts” between the patient and the provider, are widely used by mental health providers (Miller, Jacobs, and Gutheil, 1998). No evidence exists to support their effectiveness (Goldsmith et al., 2002), but detecting a treatment effect in programs targeting low-base-rate behaviors, such as suicide, is difficult (Work Group on Suicidal Behaviors, 2003). One component of safety plans, means restriction, does hold promise (Ramchand, Acosta, et al., 2011). *Means restriction* refers to any strategy that removes a suicidal patient's access to lethal means. This typically refers to removal of firearms from the patient's residence or access to firearms while on duty but also includes public health initiatives, such as packaging medications that are lethal when overdosed in blister packs or engineering shower rods to fail if an individual attempts to use one to hang him- or herself (Ramchand, Acosta, et al., 2011). Safety plans, particularly when they involve the patient's family, can and should include means-restriction plans. Given that firearms are the most common route to suicide among service members, DoD providers may wish to pay particular attention to developing plans with the patient and family to restrict firearm access (Hilton et al., 2009; Blue Ribbon Work Group on Suicide Prevention in the Veteran Population, 2008). Note that safety plans, which put specific suicide risk reduction strategies into place, are distinct from no-suicide contracts, in which the patient simply promises not to engage in suicidal behavior. No-suicide contracts are not recommended because of the lack of supportive empirical evidence and concern that providers may not closely monitor suicidal patients who sign such contracts (Assessment and Management of Risk for Suicide Working Group, 2013).

This measure is based on *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide* (Assessment and Management of Risk for Suicide Work-

ing Group, 2013). The recommendations are also consistent with *Clinical Practice Guideline: Management of Substance Use Disorders*, *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress*, and *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)*, which state that, when a patient is a threat to him- or herself or others, a plan should be implemented to ensure safety until the patient can be further evaluated and treated by a mental health professional (Bongar, 2002; Management of Bipolar Disorder Working Group, 2010; Management of MDD Working Group, 2009; Management of Substance Use Disorders Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). The APA CPGs also recommend thorough assessment of suicidality during intake evaluations (Work Group on Suicidal Behaviors, 2003). The indicator was developed by RAND researchers incorporating consultation with suicide experts and VA clinical leadership for the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011).

It is important to note that *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide* (Assessment and Management of Risk for Suicide Working Group, 2013) specifies that the recommended course of treatment be tied to a clinical judgment of whether the acute risk for suicide is low, intermediate, or high. This recommendation is not instantiated in the current measure definition. Decisions about acute risk status require the clinical provider to integrate data about SI, thoughts, planning, impulse control, previous attempts, persistence of ideation, and the strength of intent to act into a single risk status judgment. That acute risk status judgment (low, intermediate, high) is then mapped onto several possible clinical responses. When acute risk status is low, the provider can choose to consult with a behavioral health provider or address the safety issues and treat the presenting problems. When acute risk status is intermediate, the recommendations are to limit access to lethal means, conduct a complete behavioral evaluation (or refer to a behavioral health provider to do so), and determine an appropriate referral. The appropriate referral is left to the judgment of the clinician, who must select the “least restrictive level of care necessary to ensure safety.” When acute risk status is high, the guidelines recommend maintenance of direct observational control of the patient and transfer to an emergency care setting for hospitalization. As these guidelines are promulgated, it is possible that fields will be added to the electronic record to capture more-complex decisions, such as assignment to an acute risk category. Moreover, choices for the “least restrictive level of care necessary to ensure safety” may be further operationalized into an “if-then” decisionmaking tool to guide provider action. If these two steps occur, it will be important to update the quality measure to more precisely match the VA/DoD guideline (Assessment and Management of Risk for Suicide Working Group, 2013).

Potential Data Sources

Numerator

Medical record (EHR, paper record)

Denominator

Medical record (EHR, paper record)

Feasibility

The data source for this measure is the patient's medical record. Because of the complexity of the screening and assessment for SI and intent, and the application of an appropriate follow-up, this measure requires review of medical record documentation.

Feasibility Code

Orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T4*: Documented Treatment Plan

Measure

Percentage of PTSD patients with a documented treatment plan⁶

Numerator

Number of patients in the denominator with a treatment plan in the medical record

Denominator

Patients with PTSD

Definition

Treatment plan. This is a written plan that provides a means for the systematic documentation of a patient's care and progress and includes the following:

- a list of the patient's problems
- a measurable goal or set of goals related to each of the listed problems
- a specific plan of treatment to achieve each goal.

Care Continuum

Treatment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Rationale for Measure Inclusion

Treatment plans are believed to improve care coordination, interdisciplinary communication, and evidence-based care and provide a basis from which to track appropriate provision of care and treatment response. Many of the elements of measurement-based care, such as baseline assessment, monitoring over time, changes to the treatment plan when patient does not respond, and tracking to avoid relapse during the maintenance phase, all depend on there being a treatment plan in place. By making the initial course of treatment explicit, the provider is supported in ensuring that an adequate treatment trial is completed, that patient response can be assessed early and often, and that patients who are not improving can be provided with

⁶ The documentation for this measure is similar to that specified for MDD-T4*.

alternative treatment in a timely manner. In many ways, the documentation itself is the means to other ends, but, ultimately, it may be a vital element of quality care.

This measure is consistent with recommendations in *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010) and *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009), the American Psychiatric Association (APA) practice guideline for MDD (Work Group on Major Depressive Disorder, 2010), and the IOM's *Treatment for Posttraumatic Stress Disorder in Military and Vetern [sic] Populations: Initial Assessment* (2012), all of which include recommendations regarding the presence and content of patient treatment plans.

Potential Data Sources**Numerator**

Medical record (EHR, paper record)

Denominator

Administrative claims

Feasibility

This denominator for this measure (patients with PTSD) can be identified from administrative claims data. However, the assessment for the presence and adequacy of a treatment plan requires medical record review.

Feasibility Code

Orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T5: Adequate Trial of Selective Serotonin Reuptake Inhibitors and Serotonin and Norepinephrine Reuptake Inhibitors

Measure

Percentage of PTSD patients with an adequate trial of SSRIs and SNRIs

Numerator

Patients in the denominator who receive an adequate trial of SSRIs or SNRIs

(a) for at least 60 days

(b) for at least 60 days *or* have documented reasons for discontinuing SSRI or SNRI treatment in less than 60 days of the start of the SSRI or SNRI trial

Denominator

Patients with PTSD with a new prescription for an SSRI or SNRI

Definitions

Adequate SSRI or SNRI trial. Two or more consecutive 30-day prescriptions or one 90-day prescription for an SSRI or SNRI

New prescription for SSRI or SNRI. Prescription given for an SSRI or SNRI for someone for whom no prescription for any SSRI or SNRI was filled in the prior 90 days

SNRI.

- desvenlafaxine
- duloxetine
- venlafaxine
- venlafaxine extended release

SSRI.

- citalopram
- escitalopram
- fluoxetine
- fluvoxamine
- paroxetine
- sertraline

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

This measure is adapted from the VA Mental Health Program (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011). In that evaluation, this measure was applied to PTSD patients with a new treatment episode. This measure applies to all PTSD patients who are newly treated with an SSRI or SNRI, as long as there was no treatment with the same class of drug in the prior 90 days. This measure has been modified to include two options for data collection: one using exclusively administrative data and the other using medical record data to supplement the administrative data.

This indicator is based on recommendations in the *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010). The guideline strongly recommends SSRIs or SNRIs as monotherapy treatment options for PTSD. The CPG authors gave the strength of the evidence supporting this recommendation a grade of A, which is reserved for recommendations based on “good evidence that the intervention improves important health outcomes” with the added requirement that “benefits substantially outweigh harm” (Management of Post-Traumatic Stress Working Group, 2010, p. 7). Clinically, a grade of A indicates a strong recommendation for clinicians to provide the treatment to eligible patients.

A trial of an SSRI or SNRI should be optimized before shifting to a new treatment strategy. The VA/DoD CPG recommends that side effects and outcomes be monitored for a minimum of eight weeks before a clinician proceeds to a new treatment trial for a nonresponsive patient (Management of Post-Traumatic Stress Working Group, 2010). The grade for this timing recommendation is C, which indicates that there exists “fair” evidence to conclude that the recommendation “can improve health outcomes” but that the “balance of benefits to harms is too close to justify a general recommendation” (Management of Post-Traumatic Stress Working Group, 2010). Given the low grade of evidence supporting the timing for this measure, it will be important to continue to validate this measure to ensure that the threshold provides a maximized opportunity for an SSRI or SNRI to begin to reduce symptoms while minimizing the length of the time spent on unsuccessful medication trials.

Empirical support, from randomized control trials (RCTs) and meta-analyses of those trials, exists to justify the use of SSRIs and SNRIs as first-line agents for the treatment of PTSD (Brady et al., 2000; J. Davidson, Rothbaum, van der Kolk, et al., 2001; Foa, Davidson, and Frances, 1999; Jonas et al., 2013; Stein, Ipser, and Seedat, 2009). A 2012 review of PTSD pharmacotherapy indicated that the largest and most trials showing efficacy have been with the SSRIs (Ipser and Stein, 2012). Venlafaxine, an SNRI, has had positive results in two trials with more than 800 participants with non-combat-related PTSD (J. Davidson, Baldwin, et al., 2006; J. Davidson, Rothbaum, Tucker, et al., 2006). PTSD practice guidelines from the International Society for Traumatic Stress Studies and APA echo the recommendations of the VA/DoD CPG (APA, 2004; Benedek et al., 2009; Foa, Keane, and Friedman, 2000). In contrast, a 2008 IOM report concluded that there was insufficient evidence to categorize SSRIs as

an effective treatment for PTSD (IOM, 2008). Note, however, that a subsequent IOM report on treatment of PTSD among service members stated that there “are several effective pharmacotherapies for treating PTSD, particularly SSRIs” (IOM, 2012, p. 273).

Potential Data Sources

Numerators

- (a) Administrative claims and pharmacy data
- (b) Administrative claims and pharmacy data *and* medical record data

Denominator

Administrative claims and pharmacy data

Feasibility

This measure can be implemented as an administrative data measure or as a measure that combines administrative data and medical record review. The denominator in either case can be derived from administrative claims and pharmacy data. Two choices of numerator are presented here. Numerator (a) is calculated from administrative claims and pharmacy data alone, making it highly feasible, but it lacks information about valid reasons that an initiated medication trial may have been terminated early. Numerator (b) includes medical record review to supplement administrative data with medication-termination information. Using both data sources provides more-complete data but decreases feasibility because of the effort related to medical record review.

Feasibility Code, Numerator (a)

Green

Feasibility Code, Numerator (b)

Orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T6: Follow-Up of New Prescription for a Selective Serotonin Reuptake Inhibitor or Serotonin and Norepinephrine Reuptake Inhibitor

Measure

Percentage of PTSD patients newly prescribed an SSRI or SNRI with a follow-up visit *within 30 days*

Numerator

Patients in the denominator who have a follow-up visit within 30 days of a new prescription for an SSRI or SNRI

Denominator

Patients with PTSD with a new prescription for an SSRI or SNRI

Definitions

Follow-up visit. Visit with the SSRI- or SNRI-prescribing provider coded as a medication-management visit

New prescription for SSRI or SNRI. Prescription given for an SSRI or SNRI for someone for whom no prescription for an SSRI or SNRI had been filled in the prior 90 days

SNRI.

- desvenlafaxine
- duloxetine
- venlafaxine
- venlafaxine extended release

SSRI.

- citalopram
- escitalopram
- fluoxetine
- fluvoxamine
- paroxetine
- sertraline

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

New measure

Rationale for Measure Inclusion

Although there is clear evidence that antidepressant medications are associated with symptom reduction (Fournier et al., 2010), one-third of patients will discontinue treatment within a month of receiving their prescriptions (Simon, 2002). For this reason, it is important for providers to maintain contact with patients in order to assess side effects and barriers to medication adherence and treatment engagement. Providers who follow up with patients have the opportunity to work collaboratively with them to troubleshoot strategies to maintain medication adherence and treatment engagement.

This is a newly developed measure that will require validation. We believe that the 30-day follow-up window represents an adequate trial to allow the provider to make a determination of initial response and evaluate side effects experienced by the patient (Management of Post-Traumatic Stress Working Group, 2010). The follow-up visit provides an opportunity to titrate dosage, substitute a different SSRI or SNRI, or discontinue pharmacological treatment. Although the RAND team selected a 30-day window for the first follow-up, we note that this time period was selected based on clinical judgment. Research has not yet been conducted to determine the precise threshold for the time period. Validation research will be necessary in order to determine the time frame that jointly maximizes the time available for the provider and patient to schedule a visit and ensures that the time frame is no longer than the period after which treatment engagement suffers.

Finally, we draw attention to the different time frames specified for this measure and the T9 measures (PTSD and MDD). This measure requires two medication-management visits (prescribing visit and follow-up medication-management visit) within *30 days*, while the T9 measure allows *eight weeks* in which to complete the second medication-management visit. The reason for this difference is that the T9 measure assesses the *minimally* appropriate level of care for mental health patients, while this measure sets a higher threshold for ideal care.

Potential Data Sources***Numerator***

Administrative claims

Denominator

Administrative claims and pharmacy data

Feasibility

This measure can be implemented using administrative claims data and pharmacy data, making it very feasible to operationalize.

Feasibility Code

Green

Measure Set: Posttraumatic Stress Disorder—PTSD-T7: Evidence-Based Psychotherapy

Measure

Percentage of PTSD patients who receive evidence-based psychotherapy for PTSD

Numerators

(a) Patients in the denominator during the measurement period who received any evidence-based psychotherapy visits

(b) The total number of evidence-based psychotherapy visits received during the measurement period

Denominators

(a) All PTSD patients

(b) Patients with PTSD diagnosis who are receiving any psychotherapy

Definitions

Evidence-based psychotherapy. For PTSD, evidence-based psychotherapies include Trauma-Focused Cognitive–Behavioral Therapy (TF-CBT) and Stress Inoculation Training (SIT).

Measurement period. Period of time during which PTSD care is evaluated (e.g., monthly, quarterly, annually)

Psychotherapy. One or more psychotherapy encounters in which PTSD is the primary diagnosis

SIT. A general anxiety-management treatment that has been applied to PTSD. Techniques overlap with some CBT techniques but typically are non–trauma focused. Techniques include relaxation training, breathing retraining, assertiveness training, positive thinking, and thought stopping (Foa, Rothbaum, et al., 1991; Meichenbaum, 1974).

TF-CBT. Any structured psychotherapy that uses behavioral (e.g., prolonged exposure), cognitive (e.g., modification of maladaptive thoughts about trauma), or a combination of behavioral and cognitive techniques to reduce PTSD symptoms (Beck, Emery, and Greenberg, 2005; Foa, Hembree, Cahill, et al., 2005; Resick, Galovski, et al., 2008). Specific techniques include the following:

- *exposure* to habituate the patient to trauma reminders and reduce anxiety. Exposure may occur during the session, or the session may include a discussion with exposure conducted as homework. Examples of exposure techniques include in vivo exposure, imaginal exposure, flooding, prolonged exposure, directed exposure, interoceptive exposure, and eye movement desensitization and reprocessing (EMDR).
- *discussion of thoughts related to the traumatic event or events and cognitive restructuring* (e.g., discussing the connection between thoughts and feelings, treating thoughts as hypotheses about the world, evaluating the evidence supporting or disproving a certain thought, finding cognitive distortions in thoughts, creating rational and helpful alternative thoughts for a negative thought, Socratic questioning)

- *education about the cognitive behavioral model of PTSD* (e.g., discussion of the role of avoidance in maintaining PTSD symptoms, discussion of the role of maladaptive thoughts in maintaining PTSD symptoms)
- *anxiety-management techniques* (e.g., relaxation training, breathing training)
- *emotion regulation* (e.g., acceptance of negative emotions, tolerating the present moment, developing a nonjudgmental stance)
- *collaboratively determined homework* for the patient to practice skills learned in session, discussion of homework assigned during previous session, and problem-solving regarding homework noncompliance
- *activity monitoring and scheduling* (e.g., patient monitors his or her activities for a specific period of time, discussing the connection between rewarding activities and mood, determining activities that the patient finds pleasurable or provide him or her with a sense of mastery, scheduling these activities as homework)
- *training in coping skills*, such as assertiveness and problem-solving
- *relapse prevention* (e.g., discussion about how to deal with symptom recurrence, anniversaries of the trauma, traumas that may occur in the future)
- *education about CBT for PTSD* (e.g., providing the patient with a rationale for treatment, i.e., why a certain CBT technique works to reduce PTSD symptoms).

Total number of evidence-based psychotherapy visits. Total number of visits during the measurement period with the same provider as the first evidence-based psychotherapy visit

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

This measure is based on clinical care recommendations in *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010). The CPG authors identify TF-CBT and SIT as the two broad classes of evidence-based psychotherapy for PTSD. They give the strength of the evidence a grade of A, indicating that there is good evidence to support the claim that the intervention improved outcomes. This measure comes from the VHA Mental Health Program Evaluation (Farmer et al., 2010;

Watkins, Pincus, Paddock, et al., 2011) and has been updated from the source to include newer recommended evidence-based PTSD therapies.

Selection of these two classes of psychotherapy as the first-line behavioral treatments is consistent with other systematic reviews, including a Cochrane review that concluded that TF-CBT, stress management (a class that includes SIT), and EMDR are effective in the treatment of PTSD (Bisson and Andrew, 2007). Note that we have classified EMDR as a variant of TF-CBT, given evidence that calls into question the contribution of the eye-movement component of the treatment above and beyond the imaginal exposure component (P. Davidson and Parker, 2001; Spates et al., 2009). Civilian guidelines echo the VA/DoD CPGs. APA's *Practice Guideline for the Treatment of Patients with Acute Stress Disorder and Posttraumatic Stress Disorder* (APA, 2004) includes the recommendation that CBT be considered for acute and chronic PTSD and that other appropriate treatments include TF-CBT variants (e.g., EMDR, imagery rehearsal) and stress inoculation. An AHRQ report on treatment for PTSD confirms these conclusions (Jonas et al., 2013).

TF-CBT refers to a broad range of psychological interventions based on learning theory, cognitive theory, emotional processing theory, and fear-conditioning models (see definition above). Treatment includes a variety of techniques most commonly involving exposure to trauma stimuli or cognitive restructuring. These treatments are structured, typically time limited (eight to 12 sessions), and often manualized (e.g., Beck, Emery, and Greenberg, 2005; Foa, Hembree, and Rothbaum, 2007; Resick, Galovski, et al., 2008). Prolonged exposure (Foa, Hembree, and Rothbaum, 2007), a treatment protocol that transitions from imaginal exposure to in vivo exposure, has been demonstrated to reduce PTSD in a variety of populations (for review, see Cahill et al., 2009). For prolonged exposure and other exposure treatments, symptom improvement is rapid, and effect sizes are large and maintained over time (Foa, Hembree, Cahill, et al., 2009; Powers et al., 2010; Resick, Nishith, et al., 2002). In one long-term follow-up, PTSD remitted in 80 percent of treated patients, and remission was maintained for five to ten years (Resick, Williams, et al., 2012). In comparative-effectiveness trials, exposure is superior to supportive counseling, relaxation training, treatment as usual, psychotherapy without an exposure element, and combinations of pharmacology, counseling, and group therapy (Asukai et al., 2010; Boudewyns and Hyer, 1990; Bryant, Moulds, and Nixon, 2003; Marks et al., 1998; Nacasch et al., 2011; Schnurr et al., 2007; Taylor et al., 2003; Vaughan, 1994). Exposure treatments are comparable in efficacy to SIT and cognitive therapy techniques alone (for meta-analytic review, see Powers et al., 2010). Adding SIT to exposure therapy produced little added benefit (Foa, Dancu, et al., 1999; Foa, Hembree, Cahill, et al., 2009). Cognitive techniques alone (without exposure) are also effective in reducing PTSD symptoms (Cottraux et al., 2008; Marks et al., 1998; Resick, Galovski, et al., 2008; Tarrrier et al., 1999).

SIT was originally developed for a broad class of anxiety disorders (Meichenbaum, 1974) and later modified to treat PTSD among rape victims (Kilpatrick, Veronen, and Resick, 1982). The treatment does not focus as explicitly on trauma memories and includes relaxation training, education on positive thinking and positive self-talk, thought-stopping strategies, and assertiveness training (Foa, Davidson, and Frances, 1999). Comparative-effectiveness studies find SIT to be equally effective to prolonged exposure and more effective than waiting-list control (Foa, Dancu, et al., 1999; Foa, Rothbaum, et al., 1991).

Potential Data Sources

Numerator

Medical record (EHR, paper record)

Denominator

Administrative claims

Feasibility

The denominators for this measure (patients with PTSD and those patients receiving any psychotherapy) can be identified with administrative claims data or medical record review. The numerators require medical record review to determine the therapy approach used to treat the patient's PTSD and assess whether therapy was evidence-based.

Feasibility Code

Orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T8*: Psychotherapy for New Treatment Episode

Measure

Percentage of PTSD patients in a new treatment episode who received any psychotherapy⁷

Numerator

Patients in the denominator receiving any psychotherapy within four months after starting a new treatment episode

Denominator

Patients in a new treatment episode for PTSD

Definitions

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

Psychotherapy. One or more psychotherapy encounters with PTSD as the primary or nonprimary diagnosis. If the initial visit triggering the new treatment episode is a psychotherapy-related encounter, there must be at least one additional psychotherapy encounter to pass

Care Continuum

Treatment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Adapted from the following:

Sorbero, Melony E., Thomas E. Mandle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

This measure is consistent with the recommendations of *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009) and *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010), which have psychotherapy as a first-line treat-

⁷ The documentation for this measure is similar to that specified for MDD-T8*.

ment option. The CPG authors identify CBT and interpersonal psychotherapy (IPT) as the two evidence-based psychotherapies for MDD with the strongest, most extensive evidence base. For PTSD, the CPG authors identified TF-CBT and SIT as the two modalities of evidence-based psychotherapy. The strength of the evidence for all recommendations was given a grade of A, indicating that there is good evidence to support the claim that the intervention improved outcomes. The APA practice guidelines recommend that CBT be considered a first-line treatment option for both MDD and PTSD (APA, 2004; Work Group on Major Depressive Disorder, 2010). Other appropriate treatments for PTSD included TF-CBT variants (e.g., EMDR, imagery rehearsal) and stress inoculation. An AHRQ report on treatment for PTSD confirms these conclusions (Jonas et al., 2013).

Although there is research evidence supporting the claim that psychotherapy is effective as the primary or adjunct treatment for PTSD and MDD, this indicator does not capture the type of psychotherapy offered (i.e., evidence-based or not). Further, the threshold for success on the measure is met after a single psychotherapy session, which is unlikely to be adequate to achieve a response. For this reason, this indicator should be used descriptively only.

This measure was modified from a measure used in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011). Modifications include a change in what constitutes a break in care from five months to six months to match the time frame that is more generally used. Additionally, the requirement for a six-month break in PTSD-related medication was deleted because of the risk of eliminating patients from the denominator who were being treated for conditions other than PTSD. PTSD-related medications may be used to treat other conditions.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims

Feasibility

This measure is designed to be implemented using administrative claims data, making it very feasible to operationalize.

Feasibility Code

Green

Measure Set: Posttraumatic Stress Disorder—PTSD-T9*: Receipt of Care in the First Eight Weeks

Measure

Percentage of PTSD patients who received four psychotherapy visits or two medication-management visits within the first eight weeks⁸

Numerator

Patients in the denominator who had four psychotherapy visits *or* two medication-management visits within eight weeks of a new treatment episode for PTSD

Denominator

Patients in a new treatment episode for PTSD

Definitions

Medication-management visit. Outpatient visit coded as one with a focus on medication management and PTSD is the primary diagnosis

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

Psychotherapy visit. One or more psychotherapy encounters with PTSD as the primary or nonprimary diagnosis

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

New measure

Rationale for Measure Inclusion

This measure was developed for this project via a RAND consensus process involving five clinician researchers and quality measurement experts. It is designed to assess a minimally appropriate level of care for the mental health patient entering a new treatment episode. The VA/DoD CPGs for MDD and PTSD do not state explicitly the minimum or optimal number of visits during the initial treatment period (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). However, the measure is con-

⁸ The documentation for this measure is similar to that specified for MDD-T9*.

sistent with a key element of the MDD guideline, which states that “patients require frequent visits early in treatment to assess response to intervention, suicidal ideation, side effects, and psychosocial support systems” (Management of MDD Working Group, 2009). The number of psychotherapy visits (four) matches the shortest evidence-based intervention recommended in the PTSD CPG (brief CBT for acute stress disorder) (Management of Post-Traumatic Stress Working Group, 2010). The definition is also consistent with the technical specifications used in the VHA Mental Health Program Evaluation, in which any eight-week period with fewer than four psychotherapy visits was defined as a period in which the patient was *not* receiving psychotherapy (Horvitz-Lennon et al., 2009).

Two medication-management visits within eight weeks was selected as minimally appropriate follow-up because, in addition to the first visit to prescribe the new medication, a second visit would be required to meet VA/DoD practice guidelines. These guidelines recommend that the dose be titrated at four to six weeks if symptoms are nonresponsive and that the prescription be changed at eight to 12 weeks if the patient’s symptoms remain nonresponsive (Management of MDD Working Group, 2009). If the four- to six-week visit occurs on schedule with guidelines, the care would meet the threshold for this measure. Note that this measure provides a two-week buffer time period beyond CPG recommendations.

Finally, we draw attention to the different time frames specified for this measure and the T6 measures (PTSD and MDD). For medication management, this measure allows *eight weeks* in which to complete the second visit, while the T6 measures require the second visit to occur within *30 days*. The reason for this difference is that this measure assesses the minimally appropriate level of care for mental health patients, while T6 sets a higher threshold for ideal care.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims and pharmacy data

Feasibility

This measure is designed to be implemented using administrative claims data and pharmacy data, making it very feasible to operationalize.

Feasibility Code

Green

Measure Sets: Posttraumatic Stress Disorder—PTSD-T10: Response to Treatment at Six Months and PTSD-T11: Response to Treatment at 12 Months

Measure

Percentage of PTSD patients with response to treatment at six months or 12 months

Numerator

Patients who have a documented reduction of at least five points on the PCL within six months or 12 months (plus or minus 30 days)

Denominator

Patients with PTSD and PCL scores that are positive for PTSD (PCL score of at least 30)

Definition

PCL. The PCL is a 17-item self-report measure of symptoms of PTSD (Blanchard, Jones-Alexander, et al., 1996). Slight variants of the scale exist, all of which are scored identically. The PCL-C is the general (civilian) version, and scale items refer to a “stressful experience from the past.” The PCL-M is the military version, and items refer to a “stressful military experience.” The PCL-S (specific) limits responses to one particular stressful event by requiring the respondent to nominate a single stressful event; subsequent items refer to “the stressful experience.”

Care Continuum

Treatment

Measure Type

Outcome

Care Setting

Outpatient

Measure Sources

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NQF, “NQF #1884 Measure Under Consideration: Depression Response at 6 Months—Progress Towards Remission,” last updated March 7, 2013f.

NQF, “NQF #1885 Measure Under Consideration: Depression Response at 12 Months—Progress Towards Remission,” last updated March 7, 2013g.

Oxman, Thomas E., Allen J. Dietrich, John W. Williams Jr., Charles C. Engel, Mathew Friedman, Paula Schnurr, Stanley Rosenberg, and Sheila L. Barry, *RESPECT-Mil Primary Care Clinician’s Manual: Three Component Model for Primary Care Management of Depression and PTSD (Military Version)*, 3CM, August 2008. As of May 12, 2014:

http://www.pdhealth.mil/respect-mil/downloads/PCC_Final.pdf

Rationale for Measure Inclusion

These measures are based on an indicator developed for the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011). At the time of the VHA program evaluation, there were no existing performance indicators to measure this component of quality mental health care. However, clinical experts in PTSD and in general mental health care, both internal and external to the VHA, endorsed the newly developed indicator of symptom improvement as being high-priority, relevant, useful, and meaningful within the VA system (Farmer et al., 2010). Note that an important caveat of this measure is that it can be reliably and validly derived only if the PCL is regularly administered to patients with PTSD. This measure is also related to two MDD measures currently under consideration for NQF endorsement (NQF, undated [a]). These measures address a response to treatment (progress toward remission) within six and 12 months based on changes in PHQ-9 scores. Also, influencing the adaptation of these measures is the use of the PCL in the Re-Engineering Systems of Primary Care Treatment in the Military (RESPECT-Mil) program to monitor response to treatment (Oxman et al., 2008). In this program, the measure of response included an additional item to the PCL that addressed functional status. Functioning is not included in these measures because it is addressed in a separate measure using a validated measure of functioning.

It is unclear to what extent the PCL is currently regularly administered across MHS facilities. However, *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010) does specify that providers should assess PTSD symptoms during each visit. Although the authors of the guidelines make no requirement that a validated symptom measure be employed for this purpose, they do suggest that clinicians consider a measure, such as the PCL. Moreover, the guideline authors specify that a more comprehensive reassessment and evaluation of progress should be completed at least every 90 days and that this reassessment should include a standardized measure of PTSD symptoms, such as the PCL (Management of Post-Traumatic Stress Working Group, 2010). Thus, during the six-month follow-up period included in one of these quality measures, guideline-consistent care would include at least two comprehensive symptom assessments.

For a variety of reasons, we recommend that the PCL (Blanchard, Jones-Alexander, et al., 1996) be the standardized measurement tool for this purpose. First, the PCL has been validated with active-duty service members (Bliese et al., 2008). Importantly, the psychometric properties of the PCL are strong, with good internal consistency ($\alpha = 0.94\text{--}0.97$) and reliable scale scores across short test-retest periods ($r = 0.88\text{--}0.96$) (Blanchard, Jones-Alexander, et al., 1996; Ruggiero et al., 2003; Weathers, Litz, et al., 1993). Note that long-measurement-period reliability is neither expected nor desirable for a measure designed to be sensitive to symptom change over time. In fact, to the contrary, it is important to establish that measures employed for this purpose are sensitive to symptom change in response to treatment—a criterion passed by the PCL (Monson et al., 2008). Validity of the PCL as an indicator of DSM-IV diagnosis and symptom strength is strong, with the PCL being nearly collinear ($r = 0.93$) with CAPS, the gold-standard measure for psychiatric diagnosis (Blanchard, Jones-Alexander, et al., 1996). Moreover, there is good convergent validity with other validated measures of PTSD ($r = 0.77\text{--}0.93$) (Blanchard, Jones-Alexander, et al., 1996; Weathers, Litz, et al., 1993).

Investigators active in PCL refinement recommend that reductions in scale scores of 10 to 20 be considered clinically meaningful change and that reductions of five to ten points be considered reliable changes (Monson et al., 2008). We selected the minimum five-point threshold for this measure for two reasons. First, it is the threshold used to assess initial response to treat-

ment in the RESPECT-Mil protocol for primary care management of PTSD (Oxman et al., 2008). Although this protocol is designed to assess initial response (after six weeks of care), we maintained the threshold here as a *minimum* standard of care. As treatment facilities are able to maximize performance on this achievable aim, administrators may wish to set new goals for treatment success.

The recommended threshold for identifying a case as a probable PTSD case in a specialty mental health clinic is 45 to 50 (National Center for PTSD, 2012). The recommended cutoff identified by the scale author (50) (Weathers, Litz, et al., 1993) is associated with good sensitivity (0.78–0.82) and specificity (0.83–0.86) (Blanchard, Jones-Alexander, et al., 1996). In a small sample of motor vehicle accident victims, lowering the cutoff to 44 was associated with improved sensitivity (0.94), similar specificity (0.86), and strong diagnostic efficiency (0.90) (Blanchard, Jones-Alexander, et al., 1996). Thresholds to identify PTSD in primary care settings, in which the prevalence of PTSD is much lower, are shifted downward to improve identification (under the assumption that a thorough assessment would occur after the screening). The recommended threshold for identifying PTSD in these settings is 30 for both civilians and active-duty service members (Blanchard, Jones-Alexander, et al., 1996; Bliese et al., 2008; National Center for PTSD, 2012; Oxman et al., 2008). For this measure, we selected a lower threshold to ensure that all patients diagnosed with PTSD are included in the denominator even if their initial PCL scores were lower than the conventional cutoff for identification in specialty-care settings.

Of note is the fact that the diagnostic criteria for PTSD were updated in the 2013 revision of the DSM (APA, 2013). At the time of this report, the National Center for PTSD was working to revise the PCL to reflect these changes (National Center for PTSD, 2014). It will be important to track the outcome of its efforts in order to determine the ongoing suitability of the instrument and the need to update items or scoring protocols.

Given that these are outcome measures, it is important to consider case-mix adjustment when comparing results. PCL scores can be stratified by baseline score. Other potential risk-adjustment variables include gender, ZIP Code, race and ethnicity, country of origin, and primary language.

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

The denominators for these measures (patients with PTSD) can be partially calculated from administrative claims data. However, the determination of the PCL score that would trigger the relevant measure requires either medical record abstraction or the documentation of the PCL score that is data accessible. Examples of data accessibility would be an EHR with an extractable data field designated for the documentation of the PCL score or a patient data portal used for completion of the PCL tool. Lacking electronic access to these data elements would necessitate the use of the more labor-intensive medical record review. These data sources

would also be required to access the subsequent PCL score at six and 12 months after the triggering score.

Feasibility Code

Yellow to orange

Measure Sets: Posttraumatic Stress Disorder—PTSD-T12: Remission at Six Months and PTSD-T13: Remission at 12 Months

Measure

Percentage of PTSD patients in PTSD-symptom remission at six months or 12 months

Numerator

Patients with PCL scores indicative of PTSD remission (PCL score less than 28) within six months or 12 months (plus or minus 30 days)

Denominator

Patients with PTSD and initial PCL scores that are positive for PTSD (PCL score of at least 30)

Definition

PCL. The PCL is a 17-item self-report measure of symptoms of PTSD (Blanchard, Jones-Alexander, et al., 1996). Slight variants of the scale exist, all of which are scored identically. The PCL-C is the general (civilian) version, and scale items refer to a “stressful experience from the past.” The PCL-M is the military version, and items refer to a “stressful military experience.” The PCL-S (specific) limits responses to one particular stressful event by requiring the respondent to nominate a single stressful event; subsequent items refer to “the stressful experience.”

Care Continuum

Treatment

Measure Type

Outcome

Care Setting

Outpatient

Measure Sources

Adapted from the following:

NQF, “NQF #0710 Depression Remission at 12 Months,” last updated March 7, 2013e.

NQF, “NQF #0711 Depression Remission at 6 Months,” last updated April 24, 2013j.

Oxman, Thomas E., Allen J. Dietrich, John W. Williams Jr., Charles C. Engel, Mathew Friedman, Paula Schnurr, Stanley Rosenberg, and Sheila L. Barry, *RESPECT-Mil Primary Care Clinician’s Manual: Three Component Model for Primary Care Management of Depression and PTSD (Military Version)*, 3CM, August 2008. As of May 12, 2014:

http://www.pdhealth.mil/respect-mil/downloads/PCC_Final.pdf

Rationale for Measure Inclusion

Although it may not be possible for all cases, the ideal PTSD treatment outcome is symptom remission. This measure provides a quantitative metric to track success achieving this outcome as a percentage of patients who report symptom remission on a standardized scale. To be consistent with the remission definition in the RESPECT-Mil protocol for treatment of PTSD in

primary care (Oxman et al., 2008), we selected a PCL score of less than 28 as the metric for remission.⁹ The PCL used to monitor response in RESPECT-Mil included an additional that assessed functional status. Functional status is not included in the measures described here; it is addressed in a separate measure using a validated measure of functioning. These measures are patterned on the NQF-endorsed measures for MDD, which look for remission at six and 12 months based on change in the PHQ-9 score.

VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress (Management of Post-Traumatic Stress Working Group, 2010) specifies that providers should assess PTSD symptoms during each visit. Although the authors of the guidelines make no requirement that a validated symptom measure be employed for this purpose, they do suggest that clinicians consider an outcome measure, such as the PCL. Moreover, the guideline authors specify that a more comprehensive reassessment and evaluation of progress should be completed every 90 days at minimum and that this reassessment should include a standardized measure of PTSD symptoms, such as the PCL (Management of Post-Traumatic Stress Working Group, 2010). Thus, during the six-month follow-up period included in this quality measure, guideline-consistent care would include at least two comprehensive, quantitative symptom assessments.

For a variety of reasons, we recommend that the PCL (Blanchard, Jones-Alexander, et al., 1996) be the standardized measurement tool for this purpose. First, the PCL has been validated with active-duty service members (Bliese et al., 2008). Importantly, the psychometric properties of the PCL are strong, with good internal consistency ($\alpha = 0.94\text{--}0.97$) and reliable scale scores across short test-retest periods ($r = 0.88\text{--}0.96$) (Blanchard, Jones-Alexander, et al., 1996; Ruggiero et al., 2003; Weathers, Litz, et al., 1993). Note that long-measurement-period reliability is neither expected nor desirable for a measure designed to be sensitive to symptom change over time. In fact, to the contrary, it is important to establish that measures employed for this purpose are sensitive to symptom change in response to treatment—a criterion passed by the PCL (Monson et al., 2008). Validity of the PCL as an indicator of DSM-IV diagnosis and symptom strength is strong, with the PCL being nearly collinear ($r = 0.93$) with CAPS, the gold-standard measure for psychiatric diagnosis (Blanchard, Jones-Alexander, et al., 1996). Moreover, there is good convergent validity with other validated measures of PTSD ($r = 0.77\text{--}0.93$) (Blanchard, Jones-Alexander, et al., 1996; Weathers, Litz, et al., 1993). Of note is the fact that the diagnostic criteria for PTSD were updated in the 2013 revision of the DSM (APA, 2013). At the time of this report, the National Center for PTSD was working to revise the PCL to reflect these changes (National Center for PTSD, 2014). It will be important to track the outcome of its efforts in order to determine the ongoing suitability of the instrument and the need to update items or scoring protocols.

Given that these are outcome measures, it is important to consider case-mix adjustment when comparing results. PCL scores can be stratified by baseline score. Other potential risk-adjustment variables include gender, ZIP Code, race and ethnicity, country of origin, and primary language.

⁹ RESPECT-Mil total scale scores for the PCL (Oxman et al., 2008) are an algebraic transformation from the original Blanchard, Jones-Alexander, et al. (1996) PCL scoring. To convert RESPECT-Mil PCL thresholds to conventional scale scores, add 17 to the RESPECT-Mil PCL score.

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

The denominators for these measures (patients with PTSD) can be partially calculated from administrative claims data. However, the determination of the PCL result that would trigger the measure requires either medical record abstraction or the documentation of the PCL score that is data accessible. Documentation of the PCL result could most feasibly be accessed if the PCL score were recorded in an EHR with extractable data fields or collected from patients via a web-based portal. Lacking electronic access to these data elements would necessitate the use of the more labor-intensive medical record review.

Feasibility Code

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T14*: Improvement in Functional Status

Measure

Percentage of PTSD patients in a new treatment episode with improvement in functional status at six months¹⁰

Numerator

Patients in the denominator with improvement in functional status from the first visit for PTSD to six months after the first visit

Denominator

Patients with a new treatment episode for PTSD and who have at least two measures of functional status during the first six months of the new treatment episode

Definition

New treatment episode. A PTSD-related admission or transfer to an inpatient or residential mental health bed for whom PTSD is the primary diagnosis *or* an outpatient encounter in which PTSD is the primary diagnosis following a break in care (defined as no outpatient encounters in which PTSD is either the primary or nonprimary diagnosis for six or more months)

Care Continuum

Treatment

Measure Type

Outcome

Care Setting

Outpatient

Measure Source

Adapted from the following:

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Rationale for Measure Inclusion

General functioning or health-related quality of life (HRQOL) is widely recognized as an important outcome (Moriarty, Zack, and Kobau, 2003). In fact, it can be thought of as the complement to symptom-reduction or disease-remission measures, which is consistent with the World Health Organization's definition of health as "a state of complete physical, mental and social well-being—not merely the absence of disease or infirmity" (World Health Organization, 1948). The postdeployment measure on which this measure is based did not specify the

¹⁰ The documentation for this measure is similar to that specified for MDD-T14*.

instrument to be used to measure change in function. Clinicians and researchers who wish to track patient functioning over time and in response to treatment have a variety of functioning measures from which to choose. However, many of these measures are lengthy (e.g., SF-36; McHorney, Ware, and Raczek, 1993), and some of the most popular short measures (e.g., Sheehan Disability Scale [SDS] [Sheehan, Harnett-Sheehan, and Raj, 1996], five-dimension European Quality of Life [EQ-5D] [Rabin and Charro, 2001]) are associated with licensing fees. The Centers for Disease Control and Prevention (CDC) four-item HRQOL Healthy Days instrument (HRQOL-4) (CDC, 2000) is one good option that balances the need for a validated instrument of functioning with a preference for a brief and no-cost instrument.

The CDC HRQOL-4 is a four-item measure that includes a global assessment of self-reported health (“Would you say that your general health is: Excellent, Very Good, Good, Fair, Poor?”). Two questions assess the number of days during the past 30 days on which the respondent’s (1) physical health and (2) mental health were not good. The sum of these two items is known as the Unhealthy Days measure. The final item asks the respondent to estimate the number of days on which poor physical or mental health kept him or her from engaging in his or her typical daily activities.

The CDC HRQOL-4 has been widely used in population-based public health surveys, such as the state-based Behavioral Risk Factor Surveillance System (BRFSS) (Nelson et al., 2000), the National Health and Nutrition Examination Survey (NHANES) (CDC, 2014), and the Medicare Health Outcomes Survey (HOS) (NCQA, 2013b). Benchmarking data for comparisons with state and national samples are available on the CDC HRQOL website (CDC, 2012).

The test-retest reliabilities of measure items are moderate (intraclass correlation [ICC] = 0.57–0.75) (Andresen, Catlin, et al., 2003). Note that strong test-retest reliability is neither expected nor desired in measures that are designed to be sensitive to clinical change over time. In fact, to the contrary, it is important to establish that measures employed as indicators of treatment outcome are sensitive to change in response to treatment. This criterion is met by the CDC HRQOL-4. Moriarty, Zack, and Kobau (2003) observe that the “number of days in the past 30 days” response format of the Healthy Days measures makes them particularly well suited to respond to short-term changes. The measure is responsive to seasonal effects on populations (Moriarty, Zack, and Kobau, 2003) and shifts in medical utilization (Albert, 2000).

Concurrent validity of the measure has been established via strong correlations between the CDC HRQOL-4 and established measures of functioning, such as the SF-36 and EQ-5D (Andresen, Fouts, et al., 1999; Jia et al., 2011; Newschaffer, 1998). The measure has also been shown to distinguish between known disease groups (Currey et al., 2003).

Although the CDC HRQOL instrument has been used as a population health surveillance measure, to our knowledge, it has not been implemented as part of a quality measure. The validity of its use for this purpose will require pilot-testing. Additional work will also be necessary to determine the degree of improvement that must be observed before confirming that a patient has met the threshold to be classified as “improved” on the domain of functional status. That is, how many additional healthy days are required in order for a patient to be classified as improved? In the absence of this important information about change thresholds, investigators may wish to benchmark final scores against a population norm instead. For example, CDC reported that the average number of unhealthy days per month across the U.S. population is 6.0 (Zack et al., 2004). As expected, individuals with medical conditions report more

unhealthy days. For example, on average, patients with diabetes report 8.6 unhealthy days per month, patients with asthma report 11.1 unhealthy days per month, and patients with liver conditions report 14.5 unhealthy days per month (Zahran et al., 2005). Of course, it would be most useful to benchmark against the number of unhealthy days reported by patients with active PTSD or MDD. Research in this area is limited, but, in a sample of Los Angeles County residents, those with depression reported an average of 20.1 unhealthy days (Shih and Simon, 2008).

Because this is an outcome measure, adjustment for case mix is important to consider when evaluating outcomes in patient populations. Without case-mix adjustment, the sicker patients who generally receive more care and often have worse outcomes may distort the relationship between process and outcomes such that better care appears to worsen results.

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

The denominator for this measure (a new treatment episode for PTSD) can be partially calculated from administrative data. However, the functioning score (baseline and six months later) requires either medical record abstraction *or* the documentation of a quantifiable tool that is data accessible. Examples of data accessibility would be an EHR with an extractable data field designated for the documentation of the scale score or a patient data portal used for completion of the tool.

Feasibility Code

Yellow to orange

Measure Set: Posttraumatic Stress Disorder—PTSD-T15*: Follow-Up After Hospitalization for Mental Illness

Measure

Percentage of psychiatric inpatient hospital discharges of patients with PTSD with follow-up¹¹

Numerator

Inpatient discharges in the denominator in which the inpatient discharge was followed with an outpatient visit, intensive outpatient encounter, or partial hospitalization with a mental health practitioner

- (a) within 30 days of discharge
- (b) within seven days of discharge

Denominator

Patients with a PTSD diagnosis who were discharged alive from an acute inpatient setting (including acute care psychiatric facilities) with a primary mental health diagnosis on or between January 1 and December 1 of the measurement year make up the denominator. The denominator for this measure is based on discharges, not patients. Therefore, all discharges for patients who have more than one discharge on or between January 1 and December 1 of the measurement year should be included. If the discharge is followed by readmission or direct transfer to an acute facility for a mental health primary diagnosis (within the 30-day follow-up period), only the readmission discharge or the discharge from the facility to which the patient was transferred counts. Although rehospitalization might not be for the selected mental health disorder, it is likely to be for a related condition.

The denominator excludes both the initial discharge and the readmission or direct-transfer discharge if the readmission or direct-transfer discharge occurs after December 1 of the measurement year. It excludes discharges followed by readmission or direct transfer to a nonacute facility for a mental health primary diagnosis within the 30-day follow-up period. These discharges are excluded from the measure because readmission or transfer may prevent an outpatient follow-up visit from taking place. The denominator also excludes discharges in which the patient was transferred directly or readmitted within 30 days after discharge to an acute or nonacute facility for a non-mental health primary diagnosis. These discharges are excluded from the measure because rehospitalization or transfer may prevent an outpatient follow-up visit from taking place.

Definition

Eligible follow-up. Any outpatient visits, intensive outpatient encounters, or partial hospitalizations that occur within the relevant time frame, including those that occurred on the date of discharge

Care Continuum

Treatment

¹¹ The documentation for this measure is similar to that specified for MDD-T15*.

Measure Type

Process

Care Setting

Outpatient

Measure Sources

NCQA, "HEDIS 2013," c. 2012a. As of August 26, 2013:

<http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures/HEDIS2013.aspx>

NQF, "NQF #0576 Follow-Up After Hospitalization for Mental Illness," last updated February 28, 2013d.

Rationale for Measure Inclusion

This is an NQF-endorsed measure developed by NCQA (NQF, undated [a]) and included in HEDIS 2013 (NCQA, 2012a). In its rationale statement, NCQA says,

as treatment of mentally ill patients continues to shift from inpatient to outpatient settings, coordinating and maintaining continuity of care are important aspects of health care quality. There are several clinical reasons for ensuring adequate and timely follow-up care for patients after discharge from an institution or hospital for mental illness:

- Preventing readmission
- Keeping track of those who will eventually require readmission
- Providing transitional care from inpatient to outpatient setting.

It is important to provide regular follow-up therapy to patients after they have been hospitalized for mental illness. An outpatient visit with a mental health practitioner after discharge is recommended to ensure that the patient's transition to the home and work environments is supported and that gains made during hospitalization are not lost. It also helps health care providers to detect problems early and provide continuing care.

Missed appointments increase the likelihood of rehospitalization and increase the cost of outpatient care (Mitchell and Selmes, 2007). In terms of clinical characteristics, individuals with co-occurring serious mental illness and substance use disorders have high rates of treatment disengagement, as do individuals with higher levels of psychopathology (Kreyenbuhl, Nossel, and Dixon, 2009).

Disengagement from mental health services can be a significant problem that can lead to exacerbation of psychiatric symptoms, repeated hospitalizations, first-episode or recurrent homelessness, violence against others, and suicide (Dixon et al., 2009; Fischer et al., 2008). Communication between inpatient and outpatient clinicians is an intervention associated with improved odds of a successful linkage to postdischarge outpatient care (Boyer et al., 2000).

The care continuity targeted by this measure is not specifically included in the 2010 *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010). However, the guideline does make references to the potential use of case management to coordinate and increase continuity of care (Rosen et al., 2006). The 2009 *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009) also recommends the use of a case manager to coordinate communication between primary and mental health care specialists as one com-

ponent of case management (Bower et al., 2006; Gilbody, Bower, et al., 2006; J. W. Williams et al., 2007). This measure has face validity, and it is the standard of care to provide patients with adequate follow-up after an inpatient psychiatric stay. Furthermore, this indicator is an industry-standard measure, as indicated by its adaptation by HEDIS.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims

Feasibility

This measure can be operationalized using administrative claims data to identify the mental health inpatient discharges (denominator) and the recommended care for follow-up within seven and 30 days of discharge (numerator), making it highly feasible.

Feasibility Code

Green

Measure Set: Posttraumatic Stress Disorder—PTSD-T16*: Psychiatric Inpatient Capacity

Measure

Number of inpatient psychiatric beds available per 10,000 active-duty service members¹²

Numerator

Number of inpatient psychiatric beds available at the MTF across the measurement period

Denominator

Among MTFs with psychiatric inpatient capacity, number of active-duty service members served by the MTF, divided by 10,000

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

MTF with psychiatric inpatient capacity. Any MTF in which the average number of inpatient beds is greater than zero across the measurement period

Number of inpatient psychiatric beds. The number of psychiatric beds at an MTF may fluctuate over time. For this measure, the *average* number of available inpatient psychiatric beds across the measurement period should be used.

Care Continuum

Treatment

Measure Type

Structure

Care Setting

Inpatient

Measure Source

Adapted from the following:

World Health Organization, *Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies*, October 2010. As of May 13, 2014:

<http://www.who.int/healthinfo/systems/monitoring/en/>

Rationale for Measure Inclusion

This measure provides an indication of the availability of inpatient psychiatric services and may allow comparisons across MTFs to identify underserved areas (World Health Organization, 2010). In some services (e.g., Air Force), most inpatient care is provided by civilian hospitals, so this measure may be less relevant. Although structure measures receive little attention in the classic Donabedian model of organizational quality (Berwick, 1996; IOM, 2001), they may nonetheless provide an important window into the availability of quality care (Cleary

¹² The documentation for this measure is similar to that specified for MDD-T16*.

and O’Kane, undated; World Health Organization, 2010). Multiple VA/DoD CPGs indicate that inpatient psychiatric care is warranted and expected under certain clinical circumstances (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). The extent to which patients can access that inpatient care, of course, depends on a variety of factors, but perhaps the most elemental is whether or not psychiatric inpatient beds in their service areas are available. Thus, for MTFs with psychiatric inpatient capacity, the number of inpatient psychiatric beds available for every 10,000 active-duty service members is an indicator of the availability of psychiatric inpatient care. We note that the link between structure measures and outcomes can be complex and difficult to measure (Landon, Wilson, and Cleary, 1998) and sometimes weak (Landon, Zaslavsky, et al., 2001). The validity of this particular measure as an indicator of quality has not yet been established.

Potential Data Sources***Numerator***

Management data; management data from health care provider survey

Denominator

Management data

Feasibility

The denominator for this measure can be calculated from the Defense Enrollment Eligibility Reporting System (DEERS) enrollment file. The numerator may require a survey of facility management staff to report the number of beds available at that facility if this information is not otherwise recorded and easily accessible.

Feasibility Code

Green to yellow

Measure Sets: Posttraumatic Stress Disorder—PTSD-PE1*: Patient Report of Getting Treatment Quickly, PTSD-PE2*: Patient Report of Receiving Enough Information to Manage His or Her Condition, PTSD-PE3*: Patient Report of Receiving Information About Treatment Options, PTSD-PE4*: Patient Report of Being Helped by Counseling or Treatment Received, PTSD-PE5*: Patient Report of Improvement in Symptoms and Functioning, and PTSD-PE6*: Patient Report of Overall Rating of Counseling or Treatment Received

Measure

- PTSD-PE1*: percentage of patients with PTSD who report getting treatment quickly¹³
- PTSD-PE2*: percentage of patients with PTSD who report being given as much self-management information as they wanted¹⁴
- PTSD-PE3*: percentage of patients with PTSD who report being told about treatment options¹⁵
- PTSD-PE4*: percentage of patients with PTSD who report being helped by counseling or treatment received¹⁶
- PTSD-PE5*: percentage of patients with PTSD who report being better than they were one year ago¹⁷
- PTSD-PE6*: percentage of patients with PTSD who rated counseling and treatment received as 9 or 10 (out of 10)¹⁸

Numerator

- PTSD-PE1*: patients in the denominator who report “always” getting treatment quickly
 - getting help by phone
 - getting urgent treatment as soon as it was needed
 - getting an appointment as soon as it was wanted
- PTSD-PE2*: patients in the denominator who indicate that they received enough information to manage their own conditions
- PTSD-PE3*: patients in the denominator who indicate that they received information about treatment options
 - self-help or consumer-run programs
 - different treatments that are available for the condition
- PTSD-PE4*: patients in the denominator who report being helped a lot by the counseling or treatment received
- PTSD-PE5*: patients in the denominator who report being much better than they were a year ago regarding
 - ability to deal with daily problems
 - ability to deal with social situations

¹³ The documentation for this measure is similar to that specified for MDD-PE1*.

¹⁴ The documentation for this measure is similar to that specified for MDD-PE2*.

¹⁵ The documentation for this measure is similar to that specified for MDD-PE3*.

¹⁶ The documentation for this measure is similar to that specified for MDD-PE4*.

¹⁷ The documentation for this measure is similar to that specified for MDD-PE5*.

¹⁸ The documentation for this measure is similar to that specified for MDD-PE6*.

- ability to accomplish things
- ability to deal with symptoms or problems
- PTSD-PE6*: patients in the denominator whose overall rating of the counseling or treatment they received is 9 or 10 on a scale of 0 (worst) to 10 (best)

Denominator

Patients with PTSD who received any ambulatory or outpatient behavioral health care services during the measurement period, including outpatient visits or treatment sessions, medications, partial treatment, or day or night treatment, *and* answered the requisite questions for the selected measures (PTSD-PE1* through PTSD-PE6*) on the ECHO survey

Definitions

ECHO survey. The ECHO survey collects consumers' ratings of their behavioral health treatment and is designed to assess a variety of aspects of behavioral health care (Shaul et al., 2001).

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

Care Continuum

All phases

Measure Type

Patient experience

Care Setting

Outpatient

Measure Source

NQF, "NQF #0008 Experience of Care and Health Outcomes (ECHO) Survey," last updated September 17, 2012b.

Rationale for Measure Inclusion

NQF has endorsed the use of the ECHO as a self-report measure of patient satisfaction and self-reported treatment outcomes associated with mental health care (NQF, undated [a]). A version of the ECHO was selected for inclusion in HEDIS in 2002, and it is also a registered Consumer Assessment of Healthcare Providers and Systems (CAHPS) measure (ECHO Development Team, 2002). The initial pilot ECHO was developed via a collaborative effort between clinicians, patients, and QI groups (the Center for Mental Health Services, the Behavioral Health Measurement Advisory Panel, the Washington Circle, Human Services Research Institute, the Mental Health Statistics Improvement Program, the Consumer Assessment of Behavioral Health Services instrument development team, and the CAHPS instrument development group) (Shaul et al., 2001).

In a large sample of mental health consumers ($N = 3,449$), an exploratory factor analysis revealed nine subscales, which represented (1) communication and interaction with clinicians, (2) current mental or emotional status, (3) the health plan, (4) perceived improvement, (5) access to treatment, (6) information about treatment, (7) perceived efficacy of treatment,

(8) office staff, and (9) cultural competence (Shaul et al., 2001). Not all items loaded on a factor. Internal consistency for the factors ranged from 0.62 to 0.93. Informed by these analyses, the original scale was modified by removing items that performed poorly to create a final scale of 50 items. The authors offered evidence that individual ECHO items were positively correlated to ratings on two overall satisfaction items embedded in the scale as an indication of scale validity (Shaul et al., 2001). Benchmarks for scale scores and subscale scores are available from the National CAHPS Benchmarking Database at no cost. CAHPS recommends adjusting the data for respondent age, education, and general health status (AHRQ, 2014).

Potential Data Sources***Numerator***

Patient-reported data/survey

Denominator

Administrative claims; patient-reported data/survey

Feasibility

The data for the numerators of these measures are collected by patient survey. These data could be collected using a patient data portal or by a mail survey. The denominators can be identified using administrative claims data (to identify patients with PTSD and receipt of behavioral health care) and patient survey (to determine cases that answered the requisite composite survey items).

Feasibility Code

Yellow

Measure Set: Posttraumatic Stress Disorder—PTSD-RU1*: Psychiatric Inpatient Admissions

Measure

Number of psychiatric admissions per 100 patients with PTSD¹⁹

Numerator

Number of psychiatric admissions during the measurement period for patients in the denominator

Denominator

Number of patients with a PTSD diagnosis, divided by 100

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

Psychiatric admission. Any hospitalization in which a psychological condition is the primary diagnosis

Care Continuum

All phases

Measure Type

Resource use

Care Setting

Outpatient

Measure Source

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Rationale for Measure Inclusion

Inpatient psychiatric care is appropriate and recommended when the symptoms of a PH condition are severe or when the patient poses a threat to him- or herself or others (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). However, inpatient care also imposes the most restrictions on patients and is a substantial cost driver of total treatment expenditures (Luppa et al., 2007). For these reasons and others, it is generally recommended that each patient receive care in the least restrictive setting appropriate for the severity of his or her condition. Although it will always be the case that some patients are best served by inpatient care, high-quality outpatient care delivered in a timely fashion should avert some potential hospitalizations.

¹⁹ The documentation for this measure is similar to that specified for MDD-RU1*.

This measure provides the MHS a tracking metric to follow the rate of inpatient hospitalization across time. Although there is no clear benchmark for the appropriate rate of psychiatric hospitalization among patients with PH conditions, by tracking trends over time and in response to improvements in outpatient psychological care, the MHS will be in a position to monitor use and respond to indications of overuse.

Potential Data Sources***Numerator***

Administrative claims

Denominator

Administrative claims

Feasibility

The numerator and denominator for this measure can be calculated with administrative claims data.

Feasibility Code

Green

Measure Set: Posttraumatic Stress Disorder—PTSD-RU2*: Average Costs per Member per Month

Measure

Total cost of health care resources used by TRICARE beneficiaries (members) with PTSD, divided by the number of months beneficiaries are enrolled in TRICARE (member-months)²⁰

Numerator

The total costs required for treating beneficiaries diagnosed with PTSD for all of their member-months during the measurement period. Costs can be subset to create complementary measures, including the following:

- total inpatient costs
- total outpatient costs
- total pharmacy costs
- total PH-related pharmacy costs
- PTSD-specific costs
- other PH costs
- physical health costs
- medication costs
- psychotherapy costs.

Denominator

- Total number of member-months attributable to beneficiaries with PTSD
- Total number of member-months attributable to all beneficiaries

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually). In the more extensive discussion in Appendix D, we recommend an annual measurement period.

Member month. Any month during the measurement period for which an applicable beneficiary (either all beneficiaries or those diagnosed with PTSD) is enrolled in the TRICARE system and not separated from the military

Care Continuum

All phases

Measure Type

Resource use

Care Settings

Outpatient, inpatient, and residential

²⁰ The documentation for this measure is similar to that specified for MDD-RU2*. A more extensive description of this measure appears in Appendix D.

Measure Sources

Adapted from the following:

NQF, *National Voluntary Consensus Standards for Cost and Resource Use: Final Report*, April 2012a. As of September 13, 2013:

<http://www.qualityforum.org/Publications/2012/04/>

National_Voluntary_Consensus_Standards_for_Cost_and_Resource_Use.aspx

Sorbero, Melony E., Thomas E. Mannle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

We recommend using average cost PMPM as a measure of resource use, focusing on the costs for beneficiaries who have been diagnosed with PTSD during the measurement period. This measure can be calculated from existing MHS data, is used regularly by civilian-sector health plans to track the cost of treatment within managed care organizations (Kongstvedt, 2009), is similar to other measures that have been vetted by NQF (NQF, 2012a), and is similar to other measures regularly tracked already within the MHS. The measure calculates the cost of delivering PH treatment to a population diagnosed with PTSD, taking into account that individuals can enter or exit the population during an analysis period.

Potential Data Sources***Numerator***

Administrative claims and management data

Denominator

Administrative claims and management data

Feasibility

The denominator for this measure can be calculated from administrative claims data (beneficiaries with PTSD) and information about TRICARE enrollment found in the MHS Data Repository (MDR). The numerator is based on data readily available in the MDR.

Feasibility Code

Green

Descriptions of Measures of Quality of Care for Major Depressive Disorder

Measure Set: Major Depressive Disorder—MDD-S1*: Screening and Follow-Up for Common Psychological Health Conditions

Measure

Percentage of patients screened for PTSD, MDD, and alcohol misuse and, if positive, appropriate follow-up initiated¹

Numerator

Patients in the denominator who are

(a) screened annually for the following PH conditions using a standardized tool:

- PTSD
- MDD
- alcohol misuse

(b) and, if screened positive, appropriate follow-up was initiated within seven days for any of these:

- PTSD
- MDD
- alcohol misuse

Denominator

(a) All patients

(b) Patients who screened positive for PTSD, MDD, or alcohol misuse

Definitions

Screening

Screening for PTSD. Many standardized tools are available for screening for PTSD. The four-item PC-PTSD is the most widely used (IOM, 2012; Prins et al., 2003). Screening tools may be as short as a single item (Gore et al., 2008), or a longer tool may be used that screens for the condition and assesses additional symptoms that can facilitate the determination of the

¹ The documentation for this measure is similar to that specified for PTSD-S1*.

presence or absence of a diagnosis of PTSD (e.g., PCL). Examples of validated PTSD tools for adults include the following:

- DTS (Zlotnick et al., 1996)
- DEQ (Kubany et al., 2000)
- five-item primary care anxiety screener (Means-Christensen et al., 2006)
- IES-R (Tehrani, Cox, and Cox, 2002)
- LASC (King et al., 1995)
- M-3 checklist (Gaynes et al., 2010)
- M-PTSD and M-PTSD-DS (Keane, Caddell, and Taylor, 1988)
- MPSS-SR (Falsetti et al., 1993)
- Penn Inventory (Hammarberg, 1992)
- Posttraumatic Adjustment Scale (O'Donnell, Creamer, Parslow, et al., 2008)
- PDHA and PDHRA (incorporates the PC-PTSD) (Bailey, 1998)
- PC-PTSD (Prins et al., 2003)
- PTSD Brief Screen (Leskin and Westrup, 1999)
- PCL-C and PCL-M (Blanchard, Hickling, et al., 2003; Weathers, Huska, and Keane, 1991)
- PPTSD-R (Vrana and Lauterbach, 1991)
- R-CMS (Norris and Perilla, 1996)
- SPTSS (Carlson, 2002)
- Short PTSD Rating Interview (Connor and Davidson, 2001)
- Short Screening Scale for DSM-IV PTSD (Breslau et al., 1999)
- single-item PTSD screener (Gore et al., 2008)
- SPAN (J. Davidson, 2002)
- Trauma Screening Questionnaire (Brewin et al., 2002)
- TSC-40 (Briere, 1996)
- TSI (Briere et al., 1995)
- PCL-2 and PCL-6 (Lang and Stein, 2005).

Positive screen for PTSD. What constitutes a positive screen depends on the standardized screening tool used. The PC-PTSD, for example, 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 VA (Prins et al., 2003). 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.

Screening for MDD. There are several standardized tools used for screening for depression. Screening tools may be quite short with one or two items (e.g., PHQ-2), or a longer tool may be used that screens and assesses additional symptoms that can facilitate the determination of the presence or absence of a diagnosis of MDD (e.g., PHQ-9). Examples of validated MDD tools for adults include the following:

- BDI and BDI-II (Beck, Steer, and Brown, 1996; Beck, 1988)
- CESD (Devins et al., 1988)
- GDS (Lyness et al., 1997)
- DUKE-AD (Parkerson, Broadhead, and Tse, 1996)
- HRSD₁₇ (Hamilton, 1960)

- HRSD₂₁ (Hamilton, 1967)
- HSRD₂₄ (J. B. W. Williams, 1988)
- IDS-SR₃₀ (Rush, Gullion, et al., 1996)
- MADRS (Svanborg and Åsberg, 2001)
- MOSDQ (Burnam et al., 1988)
- PHQ-2 (Kroenke, Spitzer, and Williams, 2003)
- PHQ-9 (Spitzer, Kroenke, and Williams, 1999)
- PRIME-MD (Spitzer, Kroenke, and Williams, 1999)
- QIDS-SR₁₆ (Rush, Trivedi, Ibrahim, et al., 2003)
- SSDS (Barney et al., 2010)
- SCID MDD module (First et al., 1996)
- Zung Depression Scale (Gabrys and Peters, 1985).

Positive screen for MDD. What constitutes a positive screen depends on the standardized screening tool used. For example, the PHQ-2 score ranges from 0 to 6 with a positive screen designated as a score greater than 2 (Kroenke, Spitzer, and Williams, 2003).

Screening for alcohol misuse. There are several screening tools used for screening for misuse of alcohol. The following are recommended as brief, feasible screening tools (Bradley and Berger, 2013):

- AUDIT-C (Bradley, DeBenedetti, et al., 2007)
- AUDIT (Babor and Grant, 1989)
- SASQ (Seale et al., 2006).

Positive screen for alcohol misuse. What constitutes a positive screen depends on the standardized screening tool used. For example, the AUDIT-C score ranges from 0 to 12 with a positive screen designated as a score of at least 5 within VA medical centers. Although previous work has indicated that two or three points for women and four points for men may balance sensitivity in most settings (Bradley and Berger, 2013), VA adopted the higher cut point to decrease the burden of false positives on providers (Lapham et al., 2012).

Follow-Up for Positive Screen

Received appropriate follow-up. The elicitation of a positive screen for PTSD, MDD, or alcohol misuse requires further assessment of the patient to determine whether the patient has the condition in question. This requires one or more of the following:

- additional evaluation: This could be the administration of a standardized tool designed to assess for the presence or absence of the condition (e.g., PHQ-9 for depression, PCL for PTSD, or AUDIT for alcohol misuse or abuse) or a structured diagnostic interview that resulted in the determination of the presence or absence of the condition.
- referral to a practitioner qualified to diagnose and treat the condition in question
- pharmacologic intervention
- other interventions or follow-up appropriate to the treatment of the condition in question.

Care Continuum

Screening

Measure Type

Process

Care Setting

Outpatient

Measure Source

Adapted from the following:

NQF, "NQF #0418 Screening for Clinical Depression," last updated September 23, 2011b.

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

VA, *2012 Performance and Accountability Report: Honoring All Who Have Served*, 2012a. As of May 7, 2014: <http://www.va.gov/budget/report/>

Rationale for Measure Inclusion

Among active-duty service members, PTSD, MDD, and alcohol use disorders are among the most common psychiatric conditions, and all have been linked to combat deployments (Kessler, Sonnega, et al., 1995; Schell and Marshall, 2008). Moreover, these conditions frequently go undiagnosed, and all are associated with functional disability that results in significant personal and societal costs (Allison-Aipa et al., 2010; Gahm and Lucenko, 2008; Gerrity, Corson, and Dobscha, 2007; Thomas et al., 2010; E. C. Williams et al., 2012). Screening, defined as the examination of an otherwise-healthy population to identify people who are at higher risk for a condition (Morrison, 1992), is recommended when the identified condition is an important health problem, the screening instrument is valid, and adequate resources are available to ensure that someone who screens positive receives a thorough assessment and, if needed, follow-up care (IOM, 2012; Morrison, 1992). Given the prevalence of PTSD, MDD, and alcohol use disorders among service members, and significant disability associated with each, population screening should be considered (Post-Deployment Health Guideline Expert Panel, 2001). Note that, although screening for common postdeployment conditions is recommended, screening should always be conducted in concert with adequate clinical resources to ensure that every patient who screens positive for a condition receives a timely and thorough diagnostic assessment and that adequate staffing is in place to provide those who are diagnosed with a condition with guideline-appropriate treatment (U.S. Preventive Services Task Force, 2009).

Posttraumatic Stress Disorder

VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress (Management of Post-Traumatic Stress Working Group, 2010) rates the evidence as only fair in support of the claim that screening for PTSD improves health outcomes and that the cost-benefit ratio is favorable. Similarly, the IOM cautions that, although it is generally accepted that early screening for psychological disorders and subsequent treatment would improve outcomes, the evidence to support this claim is not yet strong (IOM, 2012). Nonetheless, *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* currently recommends that all new patients be screened for PTSD and then subsequently screened annually, and it assigns the strength of this recommendation a grade of B. This corresponds to a recommendation that clinicians

provide the service to eligible patients and a finding of at least fair evidence that the intervention improves health outcomes. Of the many screening and assessment tools available to assess PTSD symptoms, the VA/DoD guidelines for care recommend one of the following four validated scales: PC-PTSD (Prins et al., 2003), PTSD Brief Screen (Leskin and Westrup, 1999), Short Screening Scale for DSM-IV PTSD (Breslau et al., 1999), or the PCL-M (Blanchard, Jones-Alexander, et al., 1996; Weathers, Huska, and Keane, 1991). Note that the PC-PTSD is among the most commonly used of the screening instruments and has been embedded in DoD's PDHA and PDHRA (IOM, 2012).

Major Depressive Disorder

Clinical Practice Guideline: Management of Major Depressive Disorder (MDD) recommends that the PHQ-2 (Kroenke, Spitzer, and Williams, 2003) be completed annually by all patients seen in primary care settings. This recommendation is provided a grade of A, indicating a strong recommendation with good evidence that the treatment improves important health outcomes (Management of MDD Working Group, 2009). As in PTSD, it is critical that screening be paired with the resources to follow up any positive screen with a thorough assessment that includes both a diagnostic assessment and a suicide risk evaluation.

Alcohol Misuse

Clinical Practice Guideline: Management of Substance Use Disorders (Management of Substance Use Disorders Working Group, 2009) recommends routine annual screening for alcohol abuse for all general health and mental health patients. The authors give the level of evidence a grade of A, which corresponds to a strong recommendation with good evidence that the treatment improves important health outcomes. Providers are encouraged to choose one of the two following brief screens: AUDIT-C (Bradley, DeBenedetti, et al., 2007) or SASQ (Seale et al., 2006). Note that the commonly used CAGE screener is not considered an appropriate screen for past-year risky or hazardous drinking because the items focus on lifetime use and consequences (Management of Substance Use Disorders Working Group, 2009; Ewing, 1984).

This measure was based on an NQF-endorsed measure (0418), which recommends annual screening for depression (NQF, undated [a]); screening measures for substance use and comorbid psychiatric conditions included in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011); and a postdeployment health care expert panel recommending screening for common postdeployment conditions (Post-Deployment Health Guideline Expert Panel, 2001).

Potential Data Sources

Numerator

For (a) and (b): patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

For (a): administrative claims. For (b): patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

The (a) denominator for this measure (patients seen in outpatient care) can be identified most feasibly through administrative claims. The (b) denominator (patients who screened positive

for screened conditions) would require electronic access to the screening result or medical record review. The numerator also has two parts: (a) screening with a standardized tool and, if positive, (b) appropriate follow-up. Screening is most feasible when using a brief, standardized tool that is quantitative and has a clear definition of a positive versus negative screen. For this compound measure, brief screens for the three conditions can be combined for efficiency of screening. Quantitative scores from standardized tools that are entered into an EHR with extractable data fields or by patients via a web-based portal can facilitate identifying the numerator for screening from the electronic clinical data. The confirmation of appropriate follow-up for a positive screen is more complex in that multiple actions can satisfy this part of the measure. However, the administration of a brief screening tool followed by a more thorough assessment tool for cases that are positive (e.g., positive PHQ-2 followed by PHQ-9 or the use of PHQ-9 for screening and assessment) satisfies the follow-up requirement, and the data may be very feasible to access electronically. Other actions that satisfy the implementation of appropriate follow-up (e.g., referral) would likely require medical record review.

Feasibility Code, Screening

Green to yellow

Feasibility Code, Appropriate Follow-Up

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-A1: Baseline Symptom Assessment with the Patient Health Questionnaire (nine items)

Measure

Percentage of MDD patients in a new treatment episode with an assessment of symptoms with the PHQ-9

Numerator

Patients in the denominator who have a PHQ-9 assessment of MDD symptoms within the first 30 days of a new treatment episode

Denominator

Patients with MDD in a new treatment episode

Definitions

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed, *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months)

PHQ-9. The PHQ-9 is the depression module of the full PHQ scale and is in the public domain (Kroenke, Spitzer, and Williams, 2001). Each item corresponds to one of the DSM-IV criteria for MDD and is administered as a self-report scale completed by the patient. The measure can be scored continuously (from 0 to 27) or via a diagnostic algorithm that matches item responses to the diagnostic criteria of the DSM-IV (Kroenke, Spitzer, and Williams, 2001).

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Sources

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NQF, “NQF #0712 Depression Utilization of the PHQ-9 Tool,” last updated April 24, 2013k.

Rationale for Measure Inclusion

This measure was adapted from the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011) by changing what constitutes a break in care from five to six months, which is used more commonly in the literature. Additionally, the VA

program-evaluation measure required a six-month break in MDD-related medication; we did not include a medication break in our definition of break in care because of the risk of eliminating patients from the denominator who were being treated for conditions other than MDD (MDD-related medications may be used to treat other conditions). Consistently with NQF recommendations (NQF, undated [a]), we have also changed the measure to specify that the PHQ-9 be the standardized tool used rather than allowing any standardized tool.

This measure is consistent with *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009), which states that the PHQ-9 ought to be used as part of an initial assessment for any patient with a positive depression screen or for whom depression is suspected. The strength of evidence for this recommendation was given a grade of B, which indicates that the authors believe that there exists “at least fair evidence that the intervention improves health outcomes and concludes that benefits outweigh harm” (Management of MDD Working Group, 2009). Note that the guidelines recommend that the PHQ-9 be included as an adjunct assessment tool even when a full diagnostic interview is conducted (Management of MDD Working Group, 2009). This measure is also consistent with the Institute for Clinical Systems Improvement guideline for care of MDD in primary care settings, which recommends routine monitoring of symptoms with a standardized tool, such as the PHQ-9 (Trangle et al., 2012).

Measurement of MDD symptoms at the start of care using a standardized instrument allows clinicians to track treatment response quantitatively and, when necessary because of treatment nonresponse, to adjust the treatment plan. In order to make a determination of symptom improvement (or nonresponse) at a future time point, a baseline assessment of symptoms is necessary.

The PHQ-9 (Kroenke, Spitzer, and Williams, 2001) is the recommended standardized measurement tool for a variety of reasons. Although there are multiple validated tools to assess depression, the PHQ-9 is particularly efficient, simple to administer, and easy to score and interpret (Kroenke, Spitzer, and Williams, 2001). Internal reliability of the scale is strong ($\alpha = 0.86\text{--}0.89$), and 48-hour test-retest reliability is also strong ($r = 0.84$) even when the mode of administration differs (patient-completed versus interviewer-administered) (Kroenke, Spitzer, and Williams, 2001). In a 2007 meta-analysis of 14 psychometric evaluations of the PHQ-9, Gilbody, Richards, et al. (2007) reported a pooled sensitivity estimate of the measure of 0.80 and a specificity estimate of 0.92. Across the full range of the scale, the diagnostic performance of the scale is strong (area under the curve [AUC] = 0.95) (Kroenke, Spitzer, and Williams, 2001). Importantly, diagnostic performance did not differ depending on the scoring strategy (a diagnostic algorithm versus continuous scoring with a cut point of 10) or based on the prevalence of depression in the evaluated population (Gilbody, Richards, et al., 2007). In a summary of optimal cut points for identifying probable depression, the authors of this meta-analysis note that empirical optimal cut points have varied from 9 (community sample) to 12 (inpatient TBI sample) (Gilbody, Richards, et al., 2007). Finally, the scale performs as expected, with strong correlations between the PHQ-9 and SF-20 HRQOL scales ($r = 0.33\text{--}0.73$) and between self-reported disability days ($r = 0.24$) and health care utilization (physician visits, $r = 0.24$) (Kroenke, Spitzer, and Williams, 2001), all of which suggest good construct validity.

Potential Data Sources***Numerator***

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure (a new treatment episode of MDD) can be derived from administrative claims data (most easily) or from medical records. Documentation of the assessment of MDD symptoms with the PHQ-9 could most feasibly be accessed if the PHQ-9 score were recorded in an EHR with extractable data fields or collected from patients via a web-based portal. Other means of documenting use of the PHQ-9 (e.g., clinical notation of the score) would require medical record review, which makes that option the most labor-intensive.

Feasibility Code

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-A2: Assessment for Manic or Hypomanic Behaviors

Measure

Percentage of MDD patients in a new treatment episode assessed for manic or hypomanic behaviors

Numerator

Documentation of an assessment of the presence or absence of the symptoms or behaviors associated with mania or hypomania before or at the visit when MDD treatment is initiated

Denominator

MDD patients in a new treatment episode

Definitions

Assessment for symptoms or behaviors associated with mania or hypomania. Documentation of the presence or absence of current or prior symptoms of mania or hypomania or reference to presence or absence (prior or current) of specific symptoms of mania or hypomania, such as any of the following:

- a period of elevated, expansive or irritable mood
- grandiosity (unrealistic beliefs in one's ability, intelligence, and powers; may be delusional)
- decreased need for sleep
- more talkative than usual
- flight of ideas
- racing thoughts, distractibility
- high sex drive
- tendency to show poor judgment, such as impulsively deciding to quit a job
- increased reckless behaviors (such as lavish spending sprees, impulsive sexual indiscretions, abuse of alcohol or drugs, or ill-advised business decisions).

There are several standardized bipolar screening tools that may be used to assess for mania or hypomania, including the following:

- Altman Self-Rating Mania Scale (ASRM) (Altman et al., 1997)
- Bech-Rafaelsen Mania Scale (MAS) (Bech and Rafaelsen, 1980)
- Bipolar Spectrum Diagnostic Scale (BSDS) (Nassir Ghaemi et al., 2005)
- Brief Bipolar Disorder Symptom Scale (BDSS) (Dennehy et al., 2004)
- Clinical Global Impressions Scale for use in bipolar illness (CGI-BP) (Spearing et al., 1997)
- Hypomanic Personality Scale (Eckblad and Chapman, 1986)
- Mood Disorder Questionnaire (MDQ) (Hirschfeld et al., 2000)
- Self-Report Manic Inventory (SRMI) (Shugar et al., 1992)
- Young Mania Rating Scale (YMRS) (Young et al., 1978).

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed for whom MDD is the primary diagnosis *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months)

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

NQF, “NQF #0109—Bipolar Disorder and Major Depression: Assessment for Manic or Hypomanic Behaviors,” last updated February 6, 2013b.

Rationale for Measure Inclusion

Some patients experiencing a major depressive episode have bipolar disorder rather than a depressive disorder. For these patients, the appropriate treatment differs considerably from the treatment for MDD; in fact, typical pharmacological treatments for depression may precipitate a manic episode (e.g., Altshuler et al., 1995). For this reason, it is critical that a provider assessing a patient who is currently depressed rule out a history of manic or hypomanic episodes before proceeding with treatment.

Clinical Practice Guideline: Management of Major Depressive Disorder (MDD) (Management of MDD Working Group, 2009) says that the “possible existence of bipolar disorder should be assessed in patients presenting with depressive symptoms, using a clinical interview or bipolar questionnaire.” Similarly, APA’s *Practice Guideline for Treatment of Patients with Major Depressive Disorder* (Work Group on Major Depressive Disorder, 2010) notes that major depressive episodes are common in the course of bipolar disorder, so it is critical that providers consider “bipolar disorders as part of the differential diagnosis of major depressive disorder” and that

all patients who present for treatment for a major depressive episode should be screened for a past history of manic or hypomanic episodes and for past adverse reactions to antidepressants that might be consistent with a “switch” into hypomania or mania. (Work Group on Major Depressive Disorder, 2010, p. 24)

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure applies to MDD patients with a new treatment episode and can be identified with administrative claims data or with medical record review. The performance of the assessment for current and prior manic or hypomanic symptoms and behaviors could most feasibly be accessed if a standardized assessment instrument (e.g., BSDS) were recorded in an EHR with extractable data fields or collected from patients via a web-based portal. Other means of documenting screening for bipolar symptoms (e.g., clinical notation of their presence or absence) would require medical record review, which makes that option the most labor-intensive.

Feasibility Code

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-A3*: Assessment for Suicide Risk

Measure

Percentage of patients in a new treatment episode for MDD assessed for suicide risk²

Numerator

Patients in the denominator who are assessed for current suicide risk during the same visit in which a new treatment episode began or in the 14 days prior

Denominator

Patients with a new treatment episode for MDD

Definitions

Assessment of suicide risk. Suicide risk assessment must include questions about the following:

- suicidal ideation
- patient’s intent to initiate a suicide attempt

and, if either is present,

- patient’s plans for a suicide attempt
- whether the patient has means for completing suicide.

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed for whom MDD is the primary diagnosis *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months).

SI. SI includes any reference to the patient not wanting to live anymore, comments about killing oneself or doing oneself serious harm, passing thoughts of death, or similar thoughts. Absence of SI is documentation of specific denial of SI (e.g., “no suicidal thoughts,” “denies SI”). Using the PHQ-9, which includes an item assessing SI, would count for assessment of SI, but a PHQ of fewer items would not.

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

² The documentation for this measure is similar to that specified for PTSD-A3*.

Measure Sources

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NQF, “NQF #0104 Adult Major Depressive Disorder: Suicide Risk Assessment,” last updated March 15, 2013i.

Rationale for Measure Inclusion

Given the increased risk of attempted and completed suicide associated with most psychiatric conditions (Cavanagh et al., 2003; Kelly and Mann, 1996), it is important for providers to assess SI among new or returning patients and, when present, to implement a safety plan and begin quality mental health services (Ramchand, Acosta, et al., 2011). Case-control studies show that one-half to three-quarters of all suicides can be attributed to psychiatric disorders, typically mood and anxiety disorders (Cavanagh et al., 2003). MDD, the most strongly related disorder, increases the risk for death by suicide by 20 times relative to the general population (Cavanagh et al., 2003; Harris and Barraclough, 1997). The demographic profile of active-duty service members (younger and more likely to be male than the civilian population) also matches the demographic risk factors for completed suicide (Goldsmith et al., 2002; McKeown, Cuffe, and Schulz, 2006). For these reasons, it is important that any patient with a new treatment episode for a PH condition be assessed for suicide risk.

This measure is based on the NQF-endorsed measure 0104, which recommends screening for suicide risk any patient with a new treatment episode of MDD (NQF, undated [a]). The measure used in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011) looked for an annual assessment of SI, whereas this measure looks for a suicide risk assessment at the time of a new treatment episode. Assessing SI is a routine part of the mental status exam conducted in psychiatry, and APA recommends that it be used as part of standard practice (Work Group on Major Depressive Disorder, 2010). This recommendation received a grade of I, which indicates that it was “recommended with substantial clinical confidence.” This measure’s required components for a suicide risk assessment (ideation, intent, plans, and means) is consistent with recommendations in *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide* (Assessment and Management of Risk for Suicide Working Group, 2013).

The VA/DoD guideline also recommends that treatment providers consider nonmodifiable risk factors for suicide (e.g., younger age, male gender, family history of suicide, same-sex orientation) and modifiable risk factors (e.g., unstable housing, financial problems, psychiatric disorders) in order to determine whether the relative risk of a completed suicide is low, intermediate, or high (Assessment and Management of Risk for Suicide Working Group, 2013). These risk factors and ultimate risk status are not included in the current measure. Determining acute risk status for suicide (low, intermediate, or high) requires complex clinical judgment; integrating all risks into a single acute risk category would be difficult to perform reliably or consistently with the clinician responsible for the patient’s clinical care. It is our judgment that instantiating these guidelines into a quality measure will require a record of the *clinician’s judgment* of the patient’s risk category. Such a record of the clinician’s judgment is not currently a field in the EHR and would therefore require medical record review. That being said,

we suspect that even a medical record review would reveal that not all guideline-specified risk factors are documented in the record. However, we also believe that, as these recently released guidelines are promulgated, it is possible that a field will be added to the EHR requiring providers to indicate—when a patient is positive for SI—whether the acute risk of an attempt is low, intermediate, or high. Were this to occur, the quality measure in place should be updated to include this field in the criteria for passing the measure.

Potential Data Sources***Numerator***

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure can be defined with administrative claims data or by medical record review. The numerator for this measure would typically require medical record review. Although it would be possible to identify cases in the numerator that were negative for suicide risk (e.g., negative response to PHQ-9 item addressing SI) using electronic clinical data (e.g., if the PHQ-9 item responses were recorded in extractable fields of an electronic medical record or by a patient via a data portal), any case in which such a specific tool was not used or the screen was positive (and, therefore, required further assessment of intent and plan) would require medical record review.

Feasibility Code

Orange

Measure Set: Major Depressive Disorder—MDD-A4*: Assessment of Recent Substance Use

Measure

Percentage of MDD patients in a new treatment episode assessed for recent substance use³

Numerator

Patients in the denominator who have an assessment of recent substance abuse, including type, quantity, and frequency, within the first 30 days of a new treatment episode for MDD

Denominator

Patients in a new treatment episode for MDD

Definitions

Assessment of use. Documentation of no recent alcohol and no recent drug use *or* documentation of recent alcohol or drug use, including type, quantity, and frequency for all substances used. An appropriate screening tool may be used.

- type: An assessment of alcohol, marijuana, cocaine, heroin or other opiates, amphetamine or methamphetamine, or note indicating that the patient denied all other substance use
- quantity (for alcohol only): Any evidence of a quantity assessment, including number of drinks per day, number of drinks per week, any note about binge drinking (at least five drinks in one drinking episode for men, at least four drinks in one drinking episode for women)
- frequency: Note about daily, monthly, weekly, or occasional use.

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed for whom MDD is the primary diagnosis *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months)

Recent use. Use in the past three months

Care Continuum

Assessment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

³ The documentation for this measure is similar to that specified for PTSD-A4*.

Measure Sources

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NQF, “NQF #0110 Bipolar Disorder and Major Depression: Appraisal for Alcohol or Chemical Substance Use,” last updated February 6, 2013c.

Rationale for Measure Inclusion

Mental health patients who are currently using alcohol or other drugs do not respond to treatment as well as patients who are not using alcohol or drugs (Le Fauve et al., 2004). Moreover, the impairment associated with their mental health conditions appears to be more severe and chronic than for patients without concurrent substance use (Kessler, 2004). *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* and the 2009 *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* both recommend that current substance use patterns of patients with these disorders be assessed in order to identify substance abuse or dependency, including alcohol, nicotine, and prescribed and illicit drugs (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010).

This measure was modified from its use in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011) in a variety of ways. First, what constitutes a break in care was changed from five months to six months to match the time frame that is more generally used. Additionally, the requirement for a six-month break in MDD-related medication was deleted because of the risk of eliminating patients from the denominator who were being treated for conditions other than MDD because MDD-related medications may also be used to treat other conditions.

There is a similar measure endorsed by NQF (0110), which recommends assessing comorbid alcohol and substance use in patients with bipolar or unipolar depression (NQF, undated [a]).

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims

Feasibility

The denominator for this measure applies to any patient with a new treatment episode for MDD and can be identified with administrative claims data. Determining whether an assessment of recent substance use had occurred could be completed using patient self-reported survey data or medical record review. Portions of this assessment could be accessible from electronic clinical data if, for example, a web-based patient portal were developed and used to collect screening information such as this or documented in extractable fields of an EHR. More likely, manual medical record review would be used to determine whether clinical documenta-

tion of an assessment of recent substance use were present and whether it covered the required aspects of assessment.

Feasibility Code

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-T1: Periodic Symptom Assessment with the Patient Health Questionnaire (nine items)

Measure

Percentage of MDD patients with symptom assessment using the PHQ-9 during the four-month measurement period

Numerator

Patients in the denominator who have a PHQ-9 administered at least once during the four-month measurement period

Denominator

Patients diagnosed with MDD or dysthymia (primary diagnosis in specialty care, primary or nonprimary diagnosis in primary care) and who have an encounter within each four-month measurement period

Definitions

Four-month measurement period with an encounter. Time window in which MDD patients are either seen at an office visit or contacted via another method (phone, email) during a four-month time period defined by dates of service that fall into that time period (e.g., June 1, 2012, to September 30, 2012)

PHQ-9. The PHQ-9 is the depression module of the full PHQ scale and is in the public domain (Kroenke, Spitzer, and Williams, 2001). Each item corresponds to one of the DSM-IV criteria for MDD and is administered as a self-report scale completed by the patient. The measure can be scored continuously (from 0 to 27) or via a diagnostic algorithm that matches item responses to the diagnostic criteria of the DSM-IV (Kroenke, Spitzer, and Williams, 2001).

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

NQF, “NQF #0712 Depression Utilization of the PHQ-9 Tool,” last updated April 24, 2013k.

Rationale for Measure Inclusion

There is an increasing emphasis on the need to deliver treatment that is evidence-based and effective. Harding and colleagues (2011) make the case for measurement-based care as the standard for psychiatric practice to align treatment for PH disorders with physical health care. Standardized, repeated measurement of MDD symptoms allows clinicians to track individual patient response to treatment and allows administrators and organizations to monitor the treatment outcomes of larger patient groups. Systematic measurement of response to treat-

ment is considered an important component of enhanced primary care. In randomized trials, compared with treatment as usual, enhanced primary care for depression roughly doubles the likelihood of a treatment response (Bower et al., 2006; Gilbody, Bower, et al., 2006; J. W. Williams et al., 2007). Note that dysthymia is included in the denominator because that is the definition specified by NQF for this measure.

This measure is consistent with *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009), which recommends that depressive symptoms be carefully assessed at follow-up visits and that the PHQ-9 be used to monitor treatment response four to six weeks after initiation of treatment and periodically thereafter until full remission is achieved. The authors of the VA/DoD CPG gave the strength of the recommendations a grade of B, which corresponds to a judgment that “at least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harms” (Management of MDD Working Group, 2009). Guidelines issued by the Institute for Clinical Systems Improvement also recommend the PHQ-9 as the preferred tool to detect and monitor depression in the primary care setting (Trangle et al., 2012).

We note that the NQF recommendation that symptoms be reassessed every four months is more specific but consistent with the VA/DoD guideline, which recommends “periodic” reassessment after the initial reassessment at four to six weeks (Management of MDD Working Group, 2009; NQF, undated [a]). No specific rationale for the selection of four months rather than another time period is provided in the NQF measure documentation (NQF, undated [a]); however, measurement should be more frequent than just the beginning and end of care to ensure inclusion of patients who may not complete care and to provide intermittent assessments for treatment adjustments, if indicated. This window provides a sufficiently lengthy time period to allow a lenient estimate of compliance with measurement-based care standards.

The PHQ-9 (Kroenke, Spitzer, and Williams, 2001) is the recommended standardized measurement tool for a variety of reasons. Although there are multiple validated tools to assess depression, the PHQ-9 is particularly efficient, simple to administer, and easy to score and interpret (Kroenke, Spitzer, and Williams, 2001). Internal reliability of the scale is strong ($\alpha = 0.86\text{--}0.89$), and 48-hour test-retest reliability is also strong ($r = 0.84$) despite different modes of administration (patient-completed versus interviewer-administered) (Kroenke, Spitzer, and Williams, 2001). In a 2007 meta-analysis of 14 psychometric evaluations of the PHQ-9, Gilbody, Richards, et al. (2007) reported a pooled sensitivity estimate of the measure of 0.80 and a specificity estimate of 0.92. Across the full range of the scale, diagnostic performance is strong (AUC = 0.95) (Kroenke, Spitzer, and Williams, 2001). Importantly, diagnostic performance did not differ depending on the scoring strategy (a diagnostic algorithm versus continuous scoring with a cut point of 10) or based on the prevalence of depression in the evaluated population (Gilbody, Richards, et al., 2007). In a summary of optimal cut points for identifying probable depression, Gilbody and colleagues (2007) note that empirical optimal cut points have varied from 9 (community sample) to 12 (inpatient TBI sample). Finally, the scale performs as expected, with strong correlations between the PHQ-9 and SF-20 HRQOL scales ($r = 0.33\text{--}0.73$) and between self-reported disability days ($r = 0.24$) and health care utilization (physician visits, $r = 0.24$) (Kroenke, Spitzer, and Williams, 2001), all of which suggest good construct validity. Importantly, the scale is sensitive to change in clinical status (Löwe, Kroenke, et al., 2004; Löwe, Unützer, et al., 2004).

Potential Data Sources***Numerator***

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure (patients with MDD and an encounter during the measurement period) can be calculated from administrative claims data, depending on what types of encounters are accessible. Documentation of the use of the PHQ-9 to assess severity and treatment response could most feasibly be accessed if the PHQ-9 score were recorded in an EHR with extractable data fields or collected from patients via a web-based portal. Other means of documenting use of the PHQ-9 (e.g., clinical notation of the score) would require medical record review, which makes that option the most labor-intensive.

Feasibility Code

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-T2*: Availability of Specialty-Care Visits

Measure

Percentage of days when third available specialty-care appointment is within two days⁴

Numerator

Days in the measurement period on which the third available specialty-care visit is within two days (48 hours)

Denominator

Days in the measurement period

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

Third available appointment. Using a paper or electronic scheduling assistant, for a dummy patient in a given measurement point, the new or return specialty-care appointment that is third in a sequential list of available specialty-care appointments

Care Continuum

Treatment

Measure Type

Structure

Care Setting

Outpatient

Measure Source

Institute for Healthcare Improvement, “Third Next Available Appointment,” undated; referenced 2011. As of September 18, 2013:
<http://www.ihl.org/knowledge/Pages/Measures/ThirdNextAvailableAppointment.aspx>

Rationale for Measure Inclusion

Long wait times for care can interfere with treatment engagement (Festinger et al., 1995; Gallucci, Swartz, and Hackerman, 2005; MacDonald, Brown, and Ellis, 2000). If a first appointment is scheduled weeks after an initial call, a patient who was prepared to address his or her substance use or psychiatric symptoms at the time he or she contacted a facility may no longer be motivated to engage in treatment. For this reason, many treatment facilities strive to ensure that specialty-care appointments are scheduled to occur quickly after patients call to request care.

One measurement strategy to gauge the availability of care is to record the next available appointment at a given measurement point. For example, a dummy patient can be entered

⁴ The documentation for this measure is similar to that specified for PTSD-T2*.

into a scheduling system, and the first appointment available for that patient is recorded. This strategy to measure treatment availability is not ideal. Same-day cancellations and other unexpected events can create openings in a clinic's schedule that arbitrarily and uncharacteristically make it appear that specialty-care appointments are readily available. Moreover, a real patient may be unable to attend a first available appointment that occurs within hours of his or her call. Thus, we recommend that the third available appointment be used instead as a more accurate and sensitive measure of care availability (Institute for Healthcare Improvement, undated; Oldham, 2001). A third-available-appointment measure of treatment availability reduces the likelihood of chance and unexpected occurrences from the measure (Institute for Healthcare Improvement, undated).

To make the distinction between first-available and third-available measures more concrete, imagine two very different clinics. Clinic A is overstaffed, fills only 70 percent of its available appointments and can nearly always provide a new patient with an intake appointment on the day that the patient prefers. Clinic B operates with significant staff shortages, filling 100 percent of its available appointments, and new patients experience long delays. Both clinics typically have one patient who cancels his or her appointment on a given day. Now imagine that an administrator wanted to determine wait times for specialty care and decided to use the first available appointment as an indicator. Clinic A would have 30 percent of its appointments unfilled on the first day and would be coded with a wait time of one day or less. However, clinic B would also have an appointment available on the first day (a patient canceled that day) and would be coded as having a wait time of one day or less. That is, using a first-available-appointment indicator, both clinics would pass and would appear identical. However, from the patient's perspective, these two clinics are dramatically different. When a patient calls clinic A, he or she would be offered many different appointment times from which to choose; a patient who calls clinic B would be offered a same-day appointment that the patient is unlikely to be able to attend (because of travel times, work schedules, or child-care interference). When the patient declines the same-day cancellation appointment, the clinic's 100-percent schedule rate would mean that the next unscheduled appointment may be weeks or even months away. In some ways, same-day cancellations are noise in the system that makes it difficult for the administrator to observe real wait times. However, if *three* appointments were available within two days, it is likely that the clinic consistently maintains some immediate availability for new patients. Clinics with full booking and long wait times are much less likely to have three appointments available within two days and are therefore more likely to be identified using a third-available-appointment measure.

Potential Data Sources

Numerator

Management data

Denominator

Management data

Feasibility

Data collection can be completed electronically or manually. For manual data collection, the schedule is searched forward from the index day to count to the day of the third avail-

able appointment. An electronic scheduling system may allow searches to be conducted and recorded automatically, which would make the measure quite feasible to implement.

Feasibility Code

Green to yellow

Measure Set: Major Depressive Disorder—MDD-T3*: Appropriate Follow-Up for Endorsed Suicidal Ideation

Measure

Percentage of patient contacts with SI with appropriate follow-up⁵

Numerator

Documentation of appropriate follow-up for the SI, intent, or behavior

Denominator

Outpatient visits or contacts in which the MDD patient endorsed SI, intent, or behavior

Definitions

Appropriate follow-up. Appropriate follow-up is specific to the patient's presentation with regard to suicide ideation with or without intent and suicidal behavior.

For positive suicidal ideation and *negative* for intent:

- provision of resource list (to outpatient): Patient given a list of resources to call or visit if in danger
- appointment for follow-up

For positive suicidal ideation and *positive* for intent: provision of a family intervention or patient referral for hospitalization

For suicidal behavior: patient referral for hospitalization

Family intervention. Documented provider discussion with patient and family members regarding the patient's SI and strategies to keep the patient safe

SI. Any reference to the patient not wanting to live anymore, comments about killing oneself or doing oneself serious harm, and thoughts of death as a solution

Suicidal behavior. Any attempt to kill oneself. It includes attempted suicide and suicidal gestures (a suicidal action unlikely to be being fatal).

Suicidal intent. Documentation indicating imminent threat of suicide: Patient has a specific plan to kill him- or herself (e.g., location, timing) or has selected and has access to the means for suicide (e.g., pills, firearms).

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

⁵ The documentation for this measure is similar to that specified for PTSD-T3*.

Measure Source

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

Given that SI may predict suicidal behavior, it is important that providers with patients who endorse these thoughts provide immediate and appropriate follow-up care to reduce their patients' risk.

For patients who are actively suicidal, inpatient psychiatric hospitalization is a common prevention measure to ensure their safety. Although hospitalization typically prevents suicide during the stay, hospitalization alone has not been demonstrated to reduce the risk of suicide following discharge (Goldsmith et al., 2002; Work Group on Suicidal Behaviors, 2003). Rather, specific interventions that are conducted during the inpatient stay are the key (Brown et al., 2005; Linehan et al., 2006). Unfortunately, hospital stays are often too short to allow any specific intervention to be delivered (Goldsmith et al., 2002). Nonetheless, without other strategies to keep a patient safe who poses a short-term danger to him- or herself, hospitalization may be an appropriate strategy (Assessment and Management of Risk for Suicide Working Group, 2013; Work Group on Suicidal Behaviors, 2003).

An alternative is to develop a safety plan with a suicidal patient and his or her family and support network. These plans, often written up and signed as “contracts” between the patient and the provider, are widely used by mental health providers (Miller, Jacobs, and Gutheil, 1998). No evidence exists to support their effectiveness (Goldsmith et al., 2002), but detecting a treatment effect in programs targeting low-base-rate behaviors, such as suicide, is difficult (Work Group on Suicidal Behaviors, 2003). One component of safety plans, means restriction, does hold promise (Ramchand, Acosta, et al., 2011). *Means restriction* refers to any strategy that removes a suicidal patient's access to lethal means. This typically refers to removal of firearms from the patient's residence or access to firearms while on duty but also includes public health initiatives, such as packaging medications that are lethal when overdosed in blister packs or engineering shower rods to fail if an individual attempts to use one to hang him- or herself (Ramchand, Acosta, et al., 2011). Safety plans, particularly when they involve the patient's family, can and should include means-restriction plans. Given that firearms are the most common route to suicide among service members, DoD providers may wish to pay particular attention to developing plans with the patient and family to restrict firearm access (Hilton et al., 2009; Blue Ribbon Work Group on Suicide Prevention in the Veteran Population, 2008). Note that safety plans, which put specific suicide risk reduction strategies into place, are distinct from no-suicide contracts, in which the patient simply promises not to engage in suicidal behavior. No-suicide contracts are not recommended because of the lack of supportive empirical evidence and concern that providers may not closely monitor suicidal patients who sign such contracts (Assessment and Management of Risk for Suicide Working Group, 2013).

This measure is based on *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide* (Assessment and Management of Risk for Suicide Working Group, 2013). The recommendations are also consistent with VA/DoD CPGs for the management of substance-use disorder, PTSD, MDD, and psychoses, which state that, when a patient is a threat to him- or herself or others, a plan should be implemented to ensure safety

until the patient can be further evaluated and treated by a mental health professional (Bongar, 2002; Management of MDD Working Group, 2009; Management of Substance Use Disorders Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). The APA CPGs also recommend thorough assessment of suicidality during intake evaluations (Work Group on Suicidal Behaviors, 2003). The indicator was developed by RAND researchers, incorporating consultation with suicide experts and VA clinical leadership for the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011).

It is important to note that *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide* (Assessment and Management of Risk for Suicide Working Group, 2013) specifies that the recommended course of treatment be tied to a clinical judgment of whether the acute risk for suicide is low, intermediate, or high. This recommendation is not instantiated in the current measure definition. Decisions about acute risk status require the clinical provider to integrate data about suicide ideation, thoughts, planning, impulse control, previous attempts, persistence of ideation, and the strength of intent to act into a single risk status judgment. That acute risk status judgment (low, intermediate, high) is then mapped onto several possible clinical responses. When acute risk status is low, the provider can choose to consult with a behavioral health provider or address the safety issues and treat the presenting problems. When acute risk status is intermediate, the recommendations are to limit access to lethal means, conduct a complete behavioral evaluation (or refer to a behavioral health provider to do so), and determine an appropriate referral. The appropriate referral is left to the judgment of the clinician, who must select the “least restrictive level of care necessary to ensure safety.” When acute risk status is high, the guidelines recommend maintenance of direct observational control of the patient and transfer to an emergency care setting for hospitalization. As these guidelines are promulgated, it is possible that fields will be added to the electronic record to capture more-complex decisions, such as assignment to an acute risk category. Moreover, choices for the “least restrictive level of care necessary to ensure safety” may be further operationalized into an “if-then” decisionmaking tool to guide provider action. If these two steps occur, it will be important to update the quality measure to more precisely match the VA/DoD guideline (Assessment and Management of Risk for Suicide Working Group, 2013).

Potential Data Sources

Numerator

Medical record (EHR, paper record)

Denominator

Medical record (EHR, paper record)

Feasibility

The data source for this measure is the patient’s medical record. Because of the complexity of the screening and assessment for SI and intent, and the application of an appropriate follow-up, this measure requires review of medical record documentation.

Feasibility Code

Orange

Measure Set: Major Depressive Disorder—MDD-T4*: Documented Treatment Plan

Measure

Percentage of MDD patients with a documented treatment plan⁶

Numerator

Number of patients in the denominator with a treatment plan in the medical record

Denominator

Patients with MDD

Definition

Treatment plan. This is a written plan that provides a means for the systematic documentation of a patient's care and progress and includes the following:

- a list of the patient's problems
- a measurable goal or set of goals related to each of the listed problems
- a specific plan of treatment to achieve each goal.

Care Continuum

Treatment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Rationale for Measure Inclusion

Treatment plans are believed to improve care coordination, interdisciplinary communication, and evidence-based care and provide a basis from which to track appropriate provision of care and treatment response. Many of the elements of measurement-based care, such as baseline assessment, monitoring over time, changes to the treatment plan when patient does not respond, and tracking to avoid relapse during the maintenance phase, all depend on there being a treatment plan in place. By making the initial course of treatment explicit, the provider is supported in ensuring that an adequate treatment trial is completed, that patient response can be assessed early and often, and that patients who are not improving can be provided with

⁶ The documentation for this measure is similar to that specified for PTSD-T4*.

alternative treatment in a timely manner. In many ways, the documentation itself is the means to other ends, but, ultimately, it may be a vital element of quality care.

This measure is consistent with recommendations in *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010) and *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009), the APA practice guideline for MDD (Work Group on Major Depressive Disorder, 2010), and the IOM's *Treatment for Posttraumatic Stress Disorder in Military and Vetern [sic] Populations: Initial Assessment* (IOM, 2012), all of which include recommendations regarding the presence and content of patient treatment plans.

Potential Data Sources**Numerator**

Medical record (EHR, paper record)

Denominator

Administrative claims

Feasibility

This denominator for this measure (patients with MDD) can be identified from administrative claims data. However, the assessment for the presence and adequacy of a treatment plan requires medical record review.

Feasibility Code

Orange

Measure Set: Major Depressive Disorder—MDD-T5: Duration of Antidepressant Treatment

Measure

Percentage of MDD patients newly treated with antidepressant medication for 12 weeks or six months

Numerator

(a) Effective acute-phase treatment: At least 84 days (12 weeks) of continuous treatment with antidepressant medication during the 114-day period following the initial prescription. Gaps in medication treatment up to a total of 30 days during the 114-day period are allowed.

(b) Effective continuation-phase treatment: At least 180 days (six months) of continuous treatment with antidepressant medication during the 231-day period following the initial prescription. Gaps in medication treatment up to a total of 51 days during the 231-day period are allowed.

(c) Effective acute-phase treatment: At least 84 days (12 weeks) of continuous treatment with antidepressant medication during the 114-day period following the initial prescription. Gaps in medication treatment up to a total of 30 days during the 114-day period are allowed *or* the patient may have a documented reason for discontinuing antidepressant treatment in less than 84 days of the start of the antidepressant treatment.

(d) Effective continuation-phase treatment: At least 180 days (six months) of continuous treatment with antidepressant medication during the 231-day period following the initial prescription. Gaps in medication treatment up to a total of 51 days during the 231-day period are allowed *or* the patient may have a documented reason for discontinuing antidepressant treatment in less than 180 days of the start of the antidepressant treatment.

Denominator

Patients with MDD with a new prescription for antidepressant medication

Definitions

New prescription for antidepressant. Prescription given for an antidepressant for someone for whom no prescription for any antidepressant was filled in the prior 90 days

Antidepressants

- miscellaneous antidepressants: bupropion, vilazodone
- monoamine oxidase inhibitors: isocarboxazid, phenelzine, selegiline, tranylcypromine
- phenylpiperazine antidepressants: nefazodone, trazodone
- psychotherapeutic combinations: chlordiazepoxide amitriptyline, amitriptyline perphenazine, fluoxetine-olanzapine
- SNRI antidepressants: desvenlafaxine, duloxetine, venlafaxine
- SSRI antidepressants: citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline
- tetracyclic antidepressants: maprotiline, mirtazapine
- tricyclic antidepressants: amitriptyline, amoxapine, clomipramine, desipramine, doxepin, imipramine, nortriptyline, protriptyline, trimipramine

Care Continuum

Treatment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

NCQA, “HEDIS 2013,” c. 2012a. As of August 26, 2013:

<http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures/HEDIS2013.aspx>

NQF, “NQF #0105: Antidepressant Medication Management,” last updated March 9, 2013h.

Rationale for Measure Inclusion

This indicator is consistent with recommendations in *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009). The guideline strongly recommends antidepressant medications as a first-line treatment option for patients with MDD (see also Fournier et al., 2010; Moncrieff, Wessely, and Hardy, 2004). Given limited evidence to recommend one antidepressant over another (Gartlehner et al., 2007), the guideline suggests that clinicians choose between medications based on side-effect profiles, patient and family history, concurrent medical illness, and other prescribed medications. Recommended classes of antidepressants include SSRIs, SNRIs, bupropion, and mirtazapine (Management of MDD Working Group, 2009). For patients who remit, the guidelines recommend that they continue to take the same dose for six to 12 months to reduce the risk of relapse. The CPG authors give the strength of the evidence supporting each of these recommendations a grade of A, which corresponds to a “strong recommendation that clinicians provided the intervention to eligible patients” and is reserved for recommendations for which “good evidence was found that the intervention improves important health outcomes and . . . benefits substantially outweigh harm” (Management of MDD Working Group, 2009).

The VA/DoD CPG is consistent with the civilian treatment guideline issued by APA (Work Group on Major Depressive Disorder, 2010). APA also recommends antidepressants as a treatment option for depression and that, for patients who respond to antidepressants, treatment be continued for four to nine months to reduce the risk of relapse. Guideline authors give both recommendations a grade of I, which corresponds to recommendations that are supported with “substantial clinical confidence” (Work Group on Major Depressive Disorder, 2010). Similarly, the Institute for Clinical Systems Improvement guideline recommends antidepressants for patients with depression, indicating that the time to remission can take as long as three months, and that the medication be continued for six to 12 months for patients who respond to antidepressants (Trangle et al., 2012).

The empirical literature supports the claim that an antidepressant trial should be optimized before shifting to a new treatment strategy. For example, in a trial of fluoxetine, even among patients who showed no improvement at week 6, 31 to 41 percent achieved full remission by 12 weeks (Quitkin et al., 2003). Although antidepressant treatments should be continued for at least six months after remission to reduce the risk of relapse (Management of MDD Working Group, 2009), half of patients who begin treatment with an antidepressant discontinue the medication within one to six months after initiation (Melartin et al., 2005; Simon, 2002). These early discontinuations are associated with an increased risk for relapse and future depressive episodes (Melartin et al., 2005; Simon, 2002).

This measure has been modified from the NQF measure (NQF, undated [a]) to include two options for data collection: one using exclusively administrative data and the other using medical record data to supplement the administrative data.

Potential Data Sources

Numerator

- (a) Administrative claims and pharmacy data
- (b) Administrative claims and pharmacy data
- (c) Administrative claims and pharmacy data *and* medical record (EHR, paper record)
- (d) Administrative claims and pharmacy data *and* medical record (EHR, paper record)

Denominator

Administrative claims and pharmacy data

Feasibility

This measure can be implemented as an administrative data measure or as a measure that combines administrative data and medical record review. The denominator in either case can be derived solely from administrative claims and pharmacy data. Two choices of numerator are presented here. Numerators (a) and (b) are calculated from administrative claims and pharmacy data alone, making them highly feasible, but this approach lacks information about potentially valid reasons that an initiated medication trial may have been terminated early. Numerators (c) and (d) include medical record review to supplement administrative data with medication-termination information, where applicable. Using both data sources provides more-complete data but decreases feasibility because of the effort related to medical record review.

Feasibility Code, Numerators (a) and (b)

Green

Feasibility Code, Numerators (c) and (d)

Orange

Measure Set: Major Depressive Disorder—MDD-T6: Follow-Up of New Prescription for Antidepressant

Measure

Percentage of MDD patients newly prescribed an antidepressant with a follow-up visit within 30 days

Numerator

Patients in the denominator with a follow-up visit within 30 days of a new antidepressant prescription

Denominator

Patients with MDD and a new prescription for an antidepressant

Definitions

Follow-up visit. Visit with the antidepressant-prescribing provider coded as a medication-management visit

New prescription for antidepressant. Prescription given for an antidepressant for someone for whom no prescription for an antidepressant had been filled in the prior 90 days

Antidepressants

- miscellaneous antidepressants: bupropion, vilazodone
- monoamine oxidase inhibitors: isocarboxazid, phenelzine, selegiline, tranylcypromine
- phenylpiperazine antidepressants: nefazodone, trazodone
- psychotherapeutic combinations: chlordiazepoxide amitriptyline, amitriptyline perphenazine, fluoxetine-olanzapine
- SNRI antidepressants: desvenlafaxine, duloxetine, venlafaxine
- SSRI antidepressants: citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline
- tetracyclic antidepressants: maprotiline, mirtazapine
- tricyclic antidepressants: amitriptyline, amoxapine, clomipramine, desipramine, doxepin, imipramine, nortriptyline, protriptyline, trimipramine

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

New measure

Rationale for Measure Inclusion

Although there is clear evidence that antidepressant medications are associated with symptom reduction (Fournier et al., 2010), one-third of patients will discontinue treatment within a month of receiving the prescription (Simon, 2002). For this reason, it is important for providers to maintain contact with patients in order to assess side effects and barriers to treatment engagement. Providers who follow up with patients have the opportunity to work collaboratively with them to troubleshoot strategies to maintain treatment engagement.

This measure was developed to provide an assessment of attention to follow-up after an antidepressant prescription. We believe that the 30-day follow-up window represents an adequate trial to allow the provider to make a determination of initial response and evaluate side effects experienced by the patient. The follow-up visit provides an opportunity to titrate dosage, substitute a different antidepressant, or discontinue pharmacological treatment. Although the RAND team selected a 30-day window for the first follow-up, we note that this time period was selected based on clinical judgment. Research has not yet been conducted to determine the precise threshold for the time period. Validation research will be necessary in order to determine the time frame that jointly maximizes the time available for the provider and patient to schedule a visit and ensures that the time frame is no longer than the period after which treatment engagement suffers.

Finally, we draw attention to the different time frames specified for this measure and the T9 measures (PTSD and MDD). This measure requires two medication-management visits (prescribing visit and follow-up medication-management visit) within *30 days*, while the T9 measure allows *eight weeks* in which to complete the second medication-management visit. The reason for this difference is that the T9 measure assesses the *minimally* appropriate level of care for mental health patients, while this measure sets a higher threshold for ideal care.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims and pharmacy data

Feasibility

This measure can be implemented using administrative claims data and pharmacy data, making it very feasible to operationalize.

Feasibility Code

Green

Measure Set: Major Depressive Disorder—MDD-T7: Evidence-Based Psychotherapy

Measure

Percentage of MDD patients who receive evidence-based psychotherapy for MDD

Numerator

(a) Patients in the denominator during the measurement period who received any evidence-based psychotherapy visits, and

(b) The total number of evidence-based psychotherapy visits received during the measurement period

Denominator

(a) All MDD patients

(b) Patients with an MDD diagnosis who are receiving any psychotherapy

Definitions

CBT. CBT combines elements of cognitive and behavioral approaches, emphasizing both behavioral activation and changes in negatively biased patterns of cognition. It is based on the concept that thoughts, feelings, and behaviors are interrelated and influence each other. The treatment is structured and time-limited and includes instructional components and between-session homework assignments to solidify newly acquired skills. Typically, cognitive behavioral therapists have an agenda for each session, which ensures that specific techniques and concepts are taught during the short course of treatment. Typical CBT sessions for MDD would include the following:

- discussion of thoughts that may contribute to depression
- discussion of behaviors that may contribute to depression
- instruction on managing depression, such as reattribution and increasing participation in enjoyable activities
- education about the diagnosis of depression
- collaboration between the therapist and patient to determine homework assignments to allow the patient to practice skills learned in session. Examples include the following:
 - activity monitoring and scheduling (e.g., developing an activity schedule, completing an activity monitoring chart, recording pleasure ratings for each activity)
 - scheduling activities that the patient once enjoyed
 - practice skills learned in session
 - record maladaptive thoughts and alternative cognitions
 - conduct behavioral experiments.
- assistance in questioning negative thoughts and developing new, more-adaptive thoughts (e.g., practicing rational responses using reattribution or alternative reasoning)
- helping the patient to discover the beliefs or assumptions behind his or her maladaptive thoughts (e.g., core beliefs, cognitive schemas, patterns in thinking).

Evidence-based psychotherapy. For MDD, evidence-based psychotherapies include CBT and IPT.

IPT. IPT focuses on resolving relationship conflicts and improving role functioning in order to reduce depressive symptoms (Klerman et al., 1984). Depending on the patient, the therapist will choose to focus treatment on one of four domains: interpersonal loss, role conflict, role change, or interpersonal skills. The treatment is brief (16 to 20 sessions) and manualized (i.e., put into a structured manual).

Measurement period. Period of time during which MDD care is evaluated (e.g., monthly, quarterly, annually)

Psychotherapy. One or more psychotherapy encounters in which MDD is the primary diagnosis

Total number of evidence-based psychotherapy visits. Total number of visits during the measurement period with the same provider as the first evidence-based psychotherapy visit

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

Adapted from the following:

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

This measure has been expanded from the source to include IPT. This measure is consistent with the recommendations of *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009). This guideline identifies CBT and IPT as the two evidence-based psychotherapies for MDD with the strongest, most extensive evidence base. The guideline gives the strength of the evidence for both a grade of I (reserved for conclusions supported by at least one well-conducted RCT) and gives the strength of the recommendation an A, indicating that there is good evidence to support the claim that the intervention improved outcomes and that the benefits outweigh harm.

Selection of these two psychotherapy modalities as the first-line behavioral treatments in specialty mental health care is consistent with other systematic reviews. CBT for MDD outperforms waiting-list controls or placebo interventions with respect to MDD response and remission and performs as well as other evidence-based treatments (DeRubeis et al., 2005; Dimidjian et al., 2006; Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression, 2004; National Institute for Health and Care Excellence, 2004). Practice guidelines from APA acknowledge this evidence and include CBT as an appropriate first-line treatment for MDD (Work Group on Major Depressive Disorder, 2010).

Systematic reviews of IPT have included multiple well-conducted RCTs showing a symptom reduction relative to placebo (de Mello et al., 2005; National Institute for Health and Care Excellence, 2004). The effect sizes associated with IPT (small to moderate) were similar to those found for CBT, and comparative-effectiveness trials showed that IPT performed similarly to both CBT and antidepressants (de Mello et al., 2005; National Institute for Health and Care Excellence, 2004). APA practice guidelines also include IPT, along with CBT, as the psychotherapeutic approach with the strongest evidence (Work Group on Major Depressive Disorder, 2010). The guideline authors give IPT a grade of I, which corresponds to a recommendation with “substantial clinical confidence” (Work Group on Major Depressive Disorder, 2010).

Potential Data Sources***Numerator***

Medical record (EHR, paper record)

Denominator

Administrative claims; medical record (EHR, paper record)

Feasibility

The denominator for this measure (patients with MDD and those patients receiving any psychotherapy) can be identified with administrative claims data or medical record review. The numerator requires medical record review to determine the therapy approach used to treat the patient’s MDD and assess whether therapy was evidence-based.

Feasibility Code

Orange

Measure Set: Major Depressive Disorder—MDD-T8*: Psychotherapy for New Treatment Episode

Measure

Percentage of MDD patients in a new treatment episode who received any psychotherapy⁷

Numerator

Patients in the denominator receiving any psychotherapy within four months after starting a new treatment episode

Denominator

Patients in a new treatment episode for MDD

Definitions

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed for whom MDD is the primary diagnosis *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months)

Psychotherapy. One or more psychotherapy encounters with MDD as the primary or nonprimary diagnosis. If the initial visit triggering the new treatment episode is a psychotherapy-related encounter, there must be at least one additional psychotherapy encounter to pass.

Care Continuum

Treatment

Measure Type

Process

Care Settings

Outpatient, inpatient, and residential

Measure Source

Adapted from the following:

Sorbero, Melony E., Thomas E. Mandle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

This measure is consistent with the recommendations of *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009) and *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010), which have psychotherapy as a first-line treat-

⁷ The documentation for this measure is similar to that specified for PTSD-T8*.

ment option. The CPG authors identify CBT and IPT as the two evidence-based psychotherapies for MDD with the strongest, most extensive evidence bases. For PTSD, the CPG authors identified TF-CBT and SIT as the two modalities of evidence-based psychotherapy. The strength of the evidence for all recommendations was given a grade of A, indicating that there is good evidence to support the claim that the intervention improved outcomes. The APA practice guidelines recommend that CBT be considered a first-line treatment option for both MDD and PTSD (APA, 2004; Work Group on Major Depressive Disorder, 2010). Other appropriate treatments for PTSD included TF-CBT variants (e.g., EMDR, imagery rehearsal) and stress inoculation. An AHRQ report on treatment for PTSD confirms these conclusions (Jonas et al., 2013).

Although there is research evidence supporting the claim that psychotherapy is effective as the primary or adjunct treatment for PTSD and MDD, this indicator does not capture the type of psychotherapy offered (i.e., evidence-based or not). Further, the threshold for success on the measure is met after a single psychotherapy session, which is unlikely to be adequate to achieve a response. For this reason, this indicator should be used descriptively only.

This measure was modified from a measure used in the VHA Mental Health Program Evaluation (Farmer et al., 2010; Watkins, Pincus, Paddock, et al., 2011). Modifications include a change in what constitutes a break in care from five months to six months to match the time frame that is more generally used. Additionally, the requirement for a six-month break in MDD-related medication was deleted because of the risk of eliminating patients from the denominator who were being treated for conditions other than MDD. MDD-related medications may be used to treat other conditions.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims

Feasibility

This measure is designed to be implemented using administrative claims data, making it very feasible to operationalize.

Feasibility Code

Green

Measure Set: Major Depressive Disorder—MDD-T9*: Receipt of Care in the First Eight Weeks

Measure

Percentage of MDD patients who received four psychotherapy visits or two medication-management visits within the first eight weeks⁸

Numerator

Patients in the denominator who had four psychotherapy visits *or* two medication-management visits within eight weeks of a new treatment episode for MDD

Denominator

Patients in a new treatment episode for MDD

Definitions

Medication-management visit. Outpatient visit coded as one with a focus on medication management and MDD is the primary diagnosis

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed for whom MDD is the primary diagnosis *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months)

Psychotherapy visit. One or more psychotherapy encounters with MDD as the primary or nonprimary diagnosis

Care Continuum

Treatment

Measure Type

Process

Care Setting

Outpatient

Measure Source

New measure

Rationale for Measure Inclusion

This measure was developed for this project via a RAND consensus process involving five clinician researchers and quality measurement experts. It is designed to assess a minimally appropriate level of care for a mental health patient entering a new treatment episode. The VA/DoD CPGs for MDD and PTSD do not state explicitly the minimum or optimal number of visits during the initial treatment period (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). However, the measure is con-

⁸ The documentation for this measure is similar to that specified for PTSD-T9*.

sistent with a key element of the MDD guideline, which states that “patients require frequent visits early in treatment to assess response to intervention, suicidal ideation, side effects, and psychosocial support systems” (Management of MDD Working Group, 2009). The number of psychotherapy visits (four) matches the shortest evidence-based intervention recommended in the PTSD CPGs (brief CBT for acute stress disorder) (Management of Post-Traumatic Stress Working Group, 2010). The definition is also consistent with the technical specifications used in the VHA Mental Health Program Evaluation, in which any eight-week period with fewer than four psychotherapy visits was defined as a period in which the patient was *not* receiving psychotherapy (Horvitz-Lennon et al., 2009).

Two medication-management visits within eight weeks was selected as minimally appropriate follow-up because, in addition to the first visit to prescribe the new medication, a second visit would be required to meet VA/DoD practice guidelines. These guidelines recommend that the dose be titrated at four to six weeks if symptoms are nonresponsive and that the prescription be changed at eight to 12 weeks if the patient’s symptoms remain nonresponsive (Management of MDD Working Group, 2009).

Finally, we draw attention to the different time frames specified for this measure and the T6 measures (PTSD and MDD). For medication management, this measure allows *eight weeks* in which to complete the second visit, while the T6 measures require the second visit to occur within *30 days*. The reason for this difference is that this measure assesses the minimally appropriate level of care for mental health patients, while T6 sets a higher threshold for ideal care.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims and pharmacy data

Feasibility

This measure is designed to be implemented using administrative claims data and pharmacy data, making it very feasible to operationalize.

Feasibility Code

Green

Measure Sets: Major Depressive Disorder—MDD-T10: Response to Treatment at Six Months and MDD-T11: Response to Treatment at 12 Months

Measure

Percentage of MDD patients with response to treatment at six months or 12 months

Numerator

Patients who have a six-month or 12-month (plus or minus 30 days) PHQ-9 score that is reduced by at least 50 percent from the initial PHQ-9 score

Denominator

Patients with MDD or dysthymia and an initial PHQ-9 score positive for depression (PHQ-9 score greater than 9)

Definition

PHQ-9. The PHQ-9 is the depression module of the full PHQ scale and is in the public domain (Kroenke, Spitzer, and Williams, 2001). Each item corresponds to one of the DSM-IV criteria for MDD and is administered as a self-report scale completed by the patient. The measure can be scored continuously (from 0 to 27) or via a diagnostic algorithm that matches item responses to the diagnostic criteria of the DSM-IV (Kroenke, Spitzer, and Williams, 2001).

Care Continuum

Treatment

Measure Type

Outcome

Care Setting

Outpatient

Measure Sources

NQF, “NQF #1884 Measure Under Consideration: Depression Response at 6 Months—Progress Towards Remission,” last updated March 7, 2013f.

NQF, “NQF #1885 Measure Under Consideration: Depression Response at 12 Months—Progress Towards Remission,” last updated March 7, 2013g.

Rationale for Measure Inclusion

There is an increasing emphasis on the need to deliver treatment that is evidence-based and effective. Harding and colleagues (Harding et al., 2011) make the case for measurement-based care as the standard for psychiatric practice to align treatment for PH disorders with physical health care. Standardized, repeated measurement of MDD symptoms allows clinicians to track individual patient response to treatment and allows administrators and organizations to monitor the treatment outcomes of larger patient groups. Systematic measurement of response to treatment is considered an important component of collaborative care. In randomized trials, compared with treatment as usual, collaborative care for depression roughly doubles the likeli-

hood of a treatment response (Bower et al., 2006; Gilbody, Bower, et al., 2006; J. W. Williams et al., 2007).

The measure is consistent with *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009), which recommends that the PHQ-9 be used to monitor treatment response following the initiation of treatment and after each change in treatment. The guideline authors give the strength of this recommendation a B, which corresponds to the judgment that “at least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harms” (Management of MDD Working Group, 2009). Guidelines issued by the Institute for Clinical Systems Improvement also recommend the PHQ-9 as the preferred tool to monitor depression in the primary care setting (Trangle et al., 2012).

These two measures are measures currently under consideration by NQF for endorsement. (NQF, undated [a]). Note that dysthymia is included in the denominators because that is the definition specified by NQF for these measures. Documentation of the proposed NQF measure includes the statement that the “measure itself is determined to have face validity based on expert panel and workgroups. . . . Experts agreed on the use of common tool (PHQ-9) and that response is defined as greater than 50% improvement from the initial PHQ-9 score” (NQF, undated [a]). The Institute for Clinical Systems Improvement also suggests a 50-percent reduction on a standardized rating scale as a measure of treatment response (Trangle et al., 2012) and cites as support for this threshold two Sequenced Treatment Alternatives to Relieve Depression (STAR*D) reports that describe the use of a 50-percent reduction in the QIDS-SR₁₆ as the measure of treatment response (Rush, Trivedi, Wisniewski, et al., 2006; Trivedi et al., 2006). *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009) suggests that a five-point reduction or total score less than 10 be used as the measure of significant improvement. This alternative recommendation is consistent with an empirical evaluation of the minimal clinically important difference in PHQ-9 scores. Löwe, Unützer, et al. (2004) reported analyses revealing that a five-point (or greater) change in PHQ-9 scores reflects clinically significant change. Thus, for patients with very low inclusion scores (e.g., PHQ-9 = 10), this magnitude of change is consistent with a 50-percent reduction in scores. However, for patients with severe depression (e.g., PHQ-9 = 20), a 50-percent reduction is a considerably more stringent criterion than that suggested by the Löwe analyses.

For a variety of reasons, the PHQ-9 (Kroenke, Spitzer, and Williams, 2001) is the recommended standardized measurement tool. Although there are many validated tools to assess depression, the PHQ-9 is particularly efficient, simple to administer, and easy to score and interpret (Kroenke, Spitzer, and Williams, 2001). Internal reliability of the scale is strong ($\alpha = 0.86\text{--}0.89$), and 48-hour test-retest reliability is also strong ($r = 0.84$) despite different modes of administration (patient-completed versus interviewer-administered) (Kroenke, Spitzer, and Williams, 2001). In a 2007 meta-analysis of 14 psychometric evaluations of the PHQ-9, Gilbody, Richards, et al. (2007) reported a pooled sensitivity estimate of the measure of 0.80 and a specificity estimate of 0.92. Across the full range of the scale, diagnostic performance is strong (AUC = 0.95) (Kroenke, Spitzer, and Williams, 2001). Importantly, diagnostic performance did not differ depending on the scoring strategy (a diagnostic algorithm versus continuous scoring with a cut point of 10) or based on the prevalence of depression in the evaluated population (Gilbody, Richards, et al., 2007). In a summary of optimal cut points for identifying probable depression, Gilbody and colleagues (2007) noted that empirical optimal cut points have varied from 9 (community sample) to 12 (inpatient TBI sample). Finally, the

scale performs as expected, with strong correlations between the PHQ-9 and SF-20 HRQOL scales ($r = 0.33$ – 0.73) and between self-reported disability days ($r = 0.24$) and health care utilization (physician visits, $r = 0.24$) (Kroenke, Spitzer, and Williams, 2001), suggesting good construct validity. Importantly, the scale is sensitive to change in clinical status (Löwe, Kroenke, et al., 2004; Löwe, Unützer, et al., 2004).

Given that these are outcome measures, it is important to consider case-mix adjustment. At a minimum, the PHQ-9 scores can be stratified by baseline score. Other potential risk-adjustment variables include gender, ZIP Code, race and ethnicity, country of origin, and primary language.

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

One component of the denominators for these measures (a new treatment episode of MDD or dysthymia [now known as persistent depressive disorder]) can be calculated from administrative claims data. Documentation of the result of the PHQ-9 that would trigger the measure could most feasibly be accessed if the PHQ-9 score were recorded in an EHR with extractable data fields or collected from patients via a web-based portal. These data sources would also be required to access the subsequent PHQ-9 score at six and 12 months after the triggering score. Other means of documenting PHQ-9 results (e.g., clinical notation of the score) would require medical record review, which makes that option the most labor-intensive.

Feasibility Code

Yellow to orange

Measure Sets: Major Depressive Disorder—MDD-T12: Remission at Six Months and MDD-T13: Remission at 12 Months

Measure

Percentage of MDD patients with MDD-symptom remission at six months or 12 months

Numerator

Patients who achieve remission at six months or 12 months (plus or minus 30 days) as demonstrated by a PHQ-9 score of less than 5

Denominator

Patients with MDD or dysthymia and an initial PHQ-9 score positive for depression (PHQ-9 score greater than 9)

Definitions

PHQ-9. The PHQ-9 is the depression module of the full PHQ scale and is in the public domain (Kroenke, Spitzer, and Williams, 2001). Each item corresponds to one of the DSM-IV criteria for MDD and is administered as a self-report scale completed by the patient. The measure can be scored continuously (from 0 to 27) or via a diagnostic algorithm that matches item responses to the diagnostic criteria of the DSM-IV (Kroenke, Spitzer, and Williams, 2001).

Care Continuum

Treatment

Measure Type

Outcome

Care Setting

Outpatient

Measure Sources

NQF, “NQF #0710 Depression Remission at 12 Months,” last updated March 7, 2013e.

NQF, “NQF #0711 Depression Remission at 6 Months,” last updated April 24, 2013j.

Rationale for Measure Inclusion

There is an increasing emphasis on the need to deliver treatment that is evidence-based and effective. Harding and colleagues (2011) make the case for measurement-based care as the standard for psychiatric practice to align treatment for PH disorders with physical health care. Standardized, repeated measurement of MDD symptoms allows clinicians to track individual patient response to treatment and allows administrators and organizations to monitor the treatment outcomes of larger patient groups. Systematic measurement of treatment outcomes is considered an important component of enhanced primary care. In randomized trials, compared with treatment as usual, enhanced primary care for depression roughly doubles the likelihood of a treatment response (Bower et al., 2006; Gilbody, Bower, et al., 2006; J. W. Williams et al., 2007).

The measures presented here are NQF-endorsed (NQF, undated [a]). Note that dysthymia is included in the denominators because that is the definition specified by NQF for these measures. These measures are consistent with recommendations in *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009), which recommends that the PHQ-9 be used to monitor patients until full remission is achieved. The guideline authors give the strength of this recommendation a grade of B, which corresponds to the judgment that “at least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harms” (Management of MDD Working Group, 2009, p. 3). It is also consistent with Institute for Clinical Systems Improvement guidelines, which specify that, when selecting a treatment option, “the primary goal should be to achieve remission” (Trangle et al., 2012, p. 32). The guideline authors include the PHQ-9 as one possible standardized measure of depression and suggest that a score less than 5 be used as the threshold for remission (Trangle et al., 2012). According to research conducted by the PHQ-9 developers, scores less than 5 almost always signify the absence of a depressive disorder (Kroenke, Spitzer, and Williams, 2001).

For a variety of reasons, the PHQ-9 (Kroenke, Spitzer, and Williams, 2001) is the recommended standardized measurement tool. Although there are many validated tools to assess depression, the PHQ-9 is particularly efficient, simple to administer, and easy to score and interpret (Kroenke, Spitzer, and Williams, 2001). Internal reliability of the scale is strong ($\alpha = 0.86\text{--}0.89$), and 48-hour test-retest reliability is also strong ($r = 0.84$) despite different modes of administration (patient-completed versus interviewer-administered) (Kroenke, Spitzer, and Williams, 2001). In a 2007 meta-analysis of 14 psychometric evaluations of the PHQ-9, Gilbody, Richards, et al. (2007) reported a pooled sensitivity estimate of the measure of 0.80 and a specificity estimate of 0.92. Across the full range of the scale, diagnostic performance is strong (AUC = 0.95) (Kroenke, Spitzer, and Williams, 2001). Importantly, diagnostic performance did not differ depending on the scoring strategy (a diagnostic algorithm versus continuous scoring with a cut point of 10) or based on the prevalence of depression in the evaluated population (Gilbody, Richards, et al., 2007). In a summary of optimal cut points for identifying probable depression, Gilbody and colleagues (2007) noted that empirical optimal cut points have varied from 9 (community sample) to 12 (inpatient TBI sample). Finally, the scale performs as expected, with strong correlations between the PHQ-9 and SF-20 HRQOL scales ($r = 0.33\text{--}0.73$) and between self-reported disability days ($r = 0.24$) and health care utilization (physician visits, $r = 0.24$) (Kroenke, Spitzer, and Williams, 2001), all of which suggest good construct validity. Importantly, the scale is sensitive to change in clinical status (Löwe, Kroenke, et al., 2004; Löwe, Unützer, et al., 2004).

Given that these are outcome measures, it is important to consider case-mix adjustment. At a minimum, the PHQ-9 scores can be stratified by baseline score. Other potential risk-adjustment variables include gender, ZIP Code, race and ethnicity, country of origin, and primary language.

Potential Data Sources

Numerator

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

Although the diagnosis of MDD or dysthymia can be determined from administrative claims data, the numerators and denominators for these measures require access to patient PHQ-9 scores. Documentation of the result of the PHQ-9 could most feasibly be accessed if the PHQ-9 score were recorded in an EHR with extractable data fields or collected from patients via a web-based portal. If PHQ-9 scores are not readily accessible, data collection for this measure would require the more labor-intensive effort of medical record review.

Feasibility Code

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-T14*: Improvement in Functional Status

Measure

Percentage of MDD patients in a new treatment episode with improvement in functional status at six months⁹

Numerator

Patients in the denominator with improvement in functional status from the first visit for MDD to six months after the first visit

Denominator

Patients with a new treatment episode for MDD and who have at least two measures of functional status during the first six months of the new treatment episode

Definition

New treatment episode. An MDD-related admission or transfer to an inpatient or residential mental health bed for whom MDD is the primary diagnosis *or* an outpatient encounter in which MDD is the primary diagnosis following a break in care (defined as no outpatient encounters in which MDD is either the primary or nonprimary diagnosis for six or more months)

Care Continuum

Treatment

Measure Type

Outcome

Care Setting

Outpatient

Measure Source

Adapted from the following:

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Rationale for Measure Inclusion

General functioning or HRQOL is widely recognized as an important outcome (Moriarty, Zack, and Kobau, 2003). In fact, it can be thought of as the complement to symptom-reduction or disease-remission measures, which is consistent with the World Health Organization's definition of health as "a state of complete physical, mental and social well-being—not merely the absence of disease or infirmity" (World Health Organization, 1948). The postdeployment measure on which this measure is based did not specify the instrument to be used to measure

⁹ The documentation for this measure is similar to that specified for PTSD-T14*.

change in function. Clinicians and researchers who wish to track patient functioning over time and in response to treatment have a variety of functioning measures from which to choose. However, many of these measures are lengthy (e.g., SF-36) (McHorney, Ware, and Raczek, 1993), and some of the most popular short measures (e.g., SDS [Sheehan, Harnett-Sheehan, and Raj, 1996], EQ-5D [Rabin and Charro, 2001]) are associated with licensing fees. The CDC HRQOL-4 Healthy Days instrument (CDC, 2000) is one good option that balances the need for a validated instrument of functioning with a preference for a brief and no-cost instrument.

The CDC HRQOL-4 is a four-item measure that includes a global assessment of self-reported health (“Would you say that your general health is: Excellent, Very Good, Good, Fair, Poor?”). Two questions assess the number of days during the past 30 days on which the respondent’s (1) physical health and (2) mental health were not good. The sum of these two items is known as the Unhealthy Days measure. The final item asks the respondent to estimate the number of days on which poor physical or mental health kept him or her from engaging in his or her typical daily activities.

The CDC HRQOL-4 has been widely used in population-based public health surveys, such as the state-based BRFSS (Nelson et al., 2000), the NHANES (CDC, 2014), and the Medicare HOS (NCQA, 2013b). Benchmarking data for comparisons with state and national samples are available on the CDC HRQOL website (CDC, 2012).

The test-retest reliabilities of measure items are moderate ($ICC = 0.57-0.75$) (Andresen, Catlin, et al., 2003). Note that strong test-retest reliability is neither expected nor desired in measures that are designed to be sensitive to clinical change over time. In fact, to the contrary, it is important to establish that measures employed as indicators of treatment outcome are sensitive to change in response to treatment. This criterion is met by the CDC HRQOL-4. Moriarty, Zack, and Kobau (2003) observe that the “number of days in the past 30 days” response format of the Healthy Days measures makes them particularly well suited to respond to short-term changes. The measure is responsive to seasonal effects on populations (Moriarty, Zack, and Kobau, 2003) and shifts in medical utilization (Albert, 2000).

Concurrent validity of the measure has been established via strong correlations between the CDC HRQOL-4 and established measures of functioning, such as the SF-36 and EQ-5D (Andresen, Fouts, et al., 1999; Jia et al., 2011; Newschaffer, 1998). The measure has also been shown to distinguish between known disease groups (Currey et al., 2003).

Although the CDC HRQOL instrument has been used as a population health surveillance measure, to our knowledge, it has not been implemented as part of a quality measure. The validity of its use for this purpose will require pilot-testing. Additional work will also be necessary to determine the degree of improvement that must be observed before confirming that a patient has met the threshold to be classified as “improved” on the domain of functional status. That is, how many additional healthy days are required in order for a patient to be classified as improved? In the absence of this important information about change thresholds, investigators may wish to benchmark final scores against a population norm instead. For example, CDC reported that the average number of unhealthy days per month across the U.S. population is 6.0 (Zack et al., 2004). As expected, individuals with medical conditions report more unhealthy days. For example, on average, patients with diabetes report 8.6 unhealthy days per month, patients with asthma report 11.1 unhealthy days per month, and patients with liver conditions report 14.5 unhealthy days per month (Zahran et al., 2005). Of course, it would be most useful to benchmark against the number of unhealthy days reported by patients with

active PTSD or MDD. Research in this area is limited, but, in a sample of Los Angeles County residents, those with depression reported an average of 20.1 unhealthy days (Shih and Simon, 2008).

Because this is an outcome measure, adjustment for case mix is important to consider when evaluating outcomes in patient populations. Without case-mix adjustment, the sicker patients who generally receive more care and often have worse outcomes may distort the relationship between process and outcomes such that better care appears to worsen results.

Potential Data Sources***Numerator***

Patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Denominator

Administrative claims; patient-reported data/survey; electronic clinical data; medical record (EHR, paper record)

Feasibility

The denominator for this measure (a new treatment episode for MDD) can be partially calculated from administrative data. However, the functioning score (baseline and six months later) requires either medical record abstraction *or* the documentation of a quantifiable tool that is data accessible. Examples of data accessibility would be an EHR with an extractable data field designated for the documentation of the scale score or a patient data portal used for completion of the tool.

Feasibility Code

Yellow to orange

Measure Set: Major Depressive Disorder—MDD-T15*: Follow-Up After Hospitalization for Mental Illness

Measure

Percentage of psychiatric inpatient hospital discharges of patients with MDD with follow-up¹⁰

Numerator

Inpatient discharges in the denominator in which the inpatient discharge was followed with an outpatient visit, intensive outpatient encounter, or partial hospitalization with a mental health practitioner

- (a) within 30 days of discharge
- (b) within seven days of discharge

Denominator

Patients with an MDD diagnosis who were discharged alive from an acute inpatient setting (including acute care psychiatric facilities) with a primary mental health diagnosis on or between January 1 and December 1 of the measurement year make up the denominator. The denominator for this measure is based on discharges, not patients. Therefore, all discharges for patients who have more than one discharge on or between January 1 and December 1 of the measurement year should be included. If the discharge is followed by readmission or direct transfer to an acute facility for a mental health primary diagnosis (within the 30-day follow-up period), only the readmission discharge or the discharge from the facility to which the patient was transferred counts. Although rehospitalization might not be for the selected mental health disorder, it is likely to be for a related condition.

The denominator excludes both the initial discharge and the readmission or direct-transfer discharge if the readmission or direct-transfer discharge occurs after December 1 of the measurement year. It excludes discharges followed by readmission or direct transfer to a nonacute facility for a mental health primary diagnosis within the 30-day follow-up period. These discharges are excluded from the measure because readmission or transfer may prevent an outpatient follow-up visit from taking place. It excludes discharges in which the patient was transferred directly or readmitted within 30 days after discharge to an acute or nonacute facility for a non-mental health primary diagnosis. These discharges are excluded from the measure because rehospitalization or transfer may prevent an outpatient follow-up visit from taking place.

Definition

Eligible follow-up. Any outpatient visits, intensive outpatient encounters, or partial hospitalizations that occur within the relevant time frame, including those that occurred on the date of discharge

Care Continuum

Treatment

¹⁰ The documentation for this measure is similar to that specified for PTSD-T15*.

Measure Type

Process

Care Setting

Outpatient

Measure Sources

NCQA, "HEDIS 2013," c. 2012a. As of August 26, 2013:

<http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures/HEDIS2013.aspx>

NQF, "NQF #0576 Follow-Up After Hospitalization for Mental Illness," last updated February 28, 2013d.

Rationale for Measure Inclusion

This is an NQF-endorsed measure developed by the NCQA (2012b) and included in HEDIS 2013 (NCQA, 2012a). In its rationale statement, NCQA says,

as treatment of mentally ill patients continues to shift from inpatient to outpatient settings, coordinating and maintaining continuity of care are important aspects of health care quality. There are several clinical reasons for ensuring adequate and timely follow-up care for patients after discharge from an institution or hospital for mental illness:

- Preventing readmission
- Keeping track of those who will eventually require readmission
- Providing transitional care from inpatient to outpatient setting.

It is important to provide regular follow-up therapy to patients after they have been hospitalized for mental illness. An outpatient visit with a mental health practitioner after discharge is recommended to ensure that the patient's transition to the home and work environments is supported and that gains made during hospitalization are not lost. It also helps health care providers to detect problems early and provide continuing care.

Missed appointments increase the likelihood of rehospitalization and increase the cost of outpatient care (Mitchell and Selmes, 2007). In terms of clinical characteristics, individuals with co-occurring serious mental illness and substance use disorders have high rates of treatment disengagement, as do individuals with higher levels of psychopathology (Kreyenbuhl, Nossel, and Dixon, 2009).

Disengagement from mental health services can be a significant problem that can lead to exacerbation of psychiatric symptoms, repeated hospitalizations, first-episode or recurrent homelessness, violence against others, and suicide (Dixon et al., 2009; Fischer et al., 2008). Communication between inpatient and outpatient clinicians is an intervention associated with improved odds of a successful linkage to postdischarge outpatient care (Boyer et al., 2000).

The care continuity targeted by this measure is not specifically included in the 2010 *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress* (Management of Post-Traumatic Stress Working Group, 2010). However, the guideline does make references to the potential use of case management to coordinate and increase continuity of care (Rosen et al., 2006). The 2009 *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)* (Management of MDD Working Group, 2009) also recommends the use of a case manager to coordinate communication between primary and mental health care specialists as one com-

ponent of case management (Bower et al., 2006; Gilbody, Bower, et al., 2006; J. W. Williams et al., 2007). This measure has face validity, and it is the standard of care to provide patients with adequate follow-up after an inpatient psychiatric stay. Furthermore, this indicator is an industry-standard measure, as indicated by its adaptation by HEDIS.

Potential Data Sources***Numerator***

Administrative claims

Denominator

Administrative claims

Feasibility

This measure can be operationalized using administrative claims data to identify the mental health inpatient discharges (denominator) and the recommended care for follow-up within seven and 30 days of discharge (numerator), making it highly feasible.

Feasibility Code

Green

Measure Set: Major Depressive Disorder—MDD-T16*: Psychiatric Inpatient Capacity

Measure

Number of inpatient psychiatric beds available per 10,000 active-duty service members¹¹

Numerator

Number of inpatient psychiatric beds available at the MTF across the measurement period

Denominator

Among MTFs with psychiatric inpatient capacity, number of active-duty service members served by the MTF, divided by 10,000

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

MTF with psychiatric inpatient capacity. Any MTF in which the average number of inpatient beds is greater than zero across the measurement period

Number of inpatient psychiatric beds. The number of psychiatric beds at an MTF may fluctuate over time. For this measure, the *average* number of available inpatient psychiatric beds across the measurement period should be used.

Care Continuum

Treatment

Measure Type

Structure

Care Setting

Inpatient

Measure Source

Adapted from the following:

World Health Organization, *Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies*, October 2010. As of May 13, 2014:

<http://www.who.int/healthinfo/systems/monitoring/en/>

Rationale for Measure Inclusion

This measure provides an indication of the availability of inpatient psychiatric services and may allow comparisons across MTFs to identify underserved areas (World Health Organization, 2010). In some services (e.g., Air Force), most inpatient care is provided by civilian hospitals, so this measure may be less relevant. Although structure measures receive little attention in the classic Donabedian model of organizational quality (Berwick, 1996; IOM, 2001), they may nonetheless provide an important window into the availability of quality care (Cleary

¹¹ The documentation for this measure is similar to that specified for PTSD-T16*.

and O’Kane, undated; World Health Organization, 2010). Multiple VA/DoD CPGs indicate that inpatient psychiatric care is warranted and expected under certain clinical circumstances (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). The extent to which patients can access that inpatient care, of course, depends on a variety of factors, but perhaps the most elemental is whether or not psychiatric inpatient beds in their service areas are available. Thus, for MTFs with psychiatric inpatient capacity, the number of inpatient psychiatric beds available for every 10,000 active-duty service members is an indicator of the availability of psychiatric inpatient care. We note that the link between structure measures and outcomes can be complex and difficult to measure (Landon, Wilson, and Cleary, 1998) and sometimes weak (Landon, Zaslavsky, et al., 2001). The validity of this particular measure as an indicator of quality has not yet been established.

Potential Data Sources

Numerator

Management data; management data from health care provider survey

Denominator

Management data

Feasibility

The denominator for this measure can be calculated from the DEERS enrollment file. If this information is not otherwise recorded and easily accessible, the numerator may require a survey of facility management staff to report the number of beds available at that facility.

Feasibility Code

Green to yellow

Measure Sets: Major Depressive Disorder—MDD-PE1*: Patient Report of Getting Treatment Quickly, MDD-PE2*: Patient Report of Receiving Enough Information to Manage His or Her Condition, MDD-PE3*: Patient Report of Receiving Information About Treatment Options, MDD-PE4*: Patient Report of Being Helped by Counseling or Treatment Received, MDD-PE5*: Patient Report of Improvement in Symptoms and Functioning, MDD-PE6*: Patient Report of Overall Rating of Counseling or Treatment Received

Measures

- MDD-PE1*: percentage of patients with MDD who report getting treatment quickly¹²
- MDD-PE2*: percentage of patients with MDD who report being given as much self-management information as they wanted¹³
- MDD-PE3*: percentage of patients with MDD who report being told about treatment options¹⁴
- MDD-PE4*: percentage of patients with MDD who report being helped by counseling or treatment received¹⁵
- MDD-PE5*: percentage of patients with MDD who report being better than they were one year ago¹⁶
- MDD-PE6*: percentage of patients with MDD who rated counseling and treatment received as 9 or 10 (out of 10)¹⁷

Numerators

MDD-PE1*: patients in the denominator who report “always” getting treatment quickly

- (a) getting help by phone
- (b) getting urgent treatment as soon as it was needed
- (c) getting an appointment as soon as it was wanted

MDD-PE2*: patients in the denominator who indicate that they received enough information to manage their own conditions

MDD-PE3*: patients in the denominator who indicate that they received information about treatment options

- (a) self-help or consumer-run programs
- (b) different treatments that are available for the condition

MDD-PE4*: patients in the denominator who report being helped a lot by the counseling or treatment received

MDD-PE5*: patients in the denominator who report being much better than they were a year ago regarding

- (a) ability to deal with daily problems
- (b) ability to deal with social situations

¹² The documentation for this measure is similar to that specified for PTSD-PE1*.

¹³ The documentation for this measure is similar to that specified for PTSD-PE2*.

¹⁴ The documentation for this measure is similar to that specified for PTSD-PE3*.

¹⁵ The documentation for this measure is similar to that specified for PTSD-PE4*.

¹⁶ The documentation for this measure is similar to that specified for PTSD-PE5*.

¹⁷ The documentation for this measure is similar to that specified for PTSD-PE6*.

- (c) ability to accomplish things
- (d) ability to deal with symptoms or problems

MDD-PE6*: patients in the denominator whose overall rating of the counseling or treatment they received is 9 or 10 on a scale of 0 (worst) to 10 (best)

Denominator

Patients with MDD who received any ambulatory or outpatient behavioral health care services during the measurement period, including outpatient visits or treatment sessions, medications, partial treatment, or day or night treatment, *and* answered the requisite questions for the selected measures (PE1–PE6) on the ECHO survey

Definitions

ECHO survey. The ECHO survey collects consumers' ratings of their behavioral health treatment and is designed to assess a variety of aspects of behavioral health care (Shaul et al., 2001).

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

Care Continuum

All phases

Measure Type

Patient experience

Care Setting

Outpatient

Measure Source

NQF, "NQF #0008 Experience of Care and Health Outcomes (ECHO) Survey," last updated September 17, 2012b.

Rationale for Measure Inclusion

NQF has endorsed the use of the ECHO as a self-report measure of patient satisfaction and self-reported treatment outcomes associated with mental health care (NQF, undated [a]). A version of the ECHO was selected for inclusion in HEDIS in 2002, and it is also a registered CAHPS measure (ECHO Development Team, 2002). The initial pilot ECHO was developed via a collaborative effort between clinicians, patients, and QI groups (the Center for Mental Health Services, the Behavioral Health Measurement Advisory Panel, the Washington Circle, Human Services Research Institute, the Mental Health Statistics Improvement Program, the Consumer Assessment of Behavioral Health Services instrument development team, and the CAHPS instrument development group) (Shaul et al., 2001).

In a large sample of mental health consumers ($N = 3,449$), an exploratory factor analysis revealed nine subscales, which represented (1) communication and interaction with clinicians, (2) current mental or emotional status, (3) the health plan, (4) perceived improvement, (5) access to treatment, (6) information about treatment, (7) perceived efficacy of treatment, (8) office staff, and (9) cultural competence (Shaul et al., 2001). Not all items aligned with

one of these subscales in the factor analysis (based on factor loadings). Internal consistency for the factors ranged from 0.62 to 0.93. Informed by these analyses, the original scale was modified by removing items that performed poorly to create a final scale of 50 items. The authors offered evidence that individual ECHO items were positively correlated to ratings on two overall satisfaction items embedded in the scale as an indication of scale validity (Shaul et al., 2001). Benchmarks for scale scores and subscale scores are available from the National CAHPS Benchmarking Database at no cost. CAHPS recommends adjusting the data for respondent age, education, and general health status (AHRQ, 2014).

Potential Data Sources***Numerator***

Patient-reported data/survey

Denominator

Administrative claims; patient-reported data/survey

Feasibility

The data for the numerators of these measures are collected by patient survey. These data could be collected using a patient data portal or by a mail survey. The denominators can be identified using administrative claims data (to identify patients with MDD and receipt of behavioral health care) and patient survey (to determine cases that answered the requisite composite survey items).

Feasibility Code

Yellow

Measure Set: Major Depressive Disorder—MDD-RU1*: Psychiatric Inpatient Admissions

Measure

Number of psychiatric admissions per 100 patients with MDD¹⁸

Numerator

Number of psychiatric admissions during the measurement period for patients in the denominator

Denominator

Number of patients with an MDD diagnosis, divided by 100

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually)

Psychiatric admission. Any hospitalization in which a psychological condition is the primary diagnosis

Care Continuum

All phases

Measure Type

Resource use

Care Setting

Outpatient

Measure Source

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Rationale for Measure Inclusion

Inpatient psychiatric care is appropriate and recommended when the symptoms of a PH condition are severe or when the patient poses a threat to him- or herself or others (Management of MDD Working Group, 2009; Management of Post-Traumatic Stress Working Group, 2010). However, inpatient care also imposes the most restrictions on patients and is a substantial cost driver of total treatment expenditures (Luppa et al., 2007). For these reasons and others, it is generally recommended that each patient receive care in the least restrictive setting appropriate for the severity of his or her condition. Although it will always be the case that some patients are best served by inpatient care, high-quality outpatient care delivered in a timely fashion should avert some potential hospitalizations.

¹⁸ The documentation for this measure is similar to that specified for PTSD-RU1*.

This measure provides the MHS a tracking metric to follow the rate of inpatient hospitalization across time. Although there is no clear benchmark for the appropriate rate of psychiatric hospitalization among patients with PH conditions, by tracking trends over time and in response to improvements in outpatient psychological care, the MHS will be in a position to monitor use and respond to indications of overuse.

Potential Data Sources

Numerator

Administrative claims

Denominator

Administrative claims

Feasibility

The numerator and denominator for this measure can be calculated with administrative claims data.

Feasibility Code

Green

Measure Set: Major Depressive Disorder—MDD-RU2*: Average Costs per Member per Month

Measure

Total cost of health care resources used by TRICARE beneficiaries (members) with MDD, divided by the number of months beneficiaries are enrolled in TRICARE (member-months)¹⁹

Numerator

The total costs required for treating beneficiaries diagnosed with MDD for all of their member-months during the measurement period. Costs can be subset to create complementary measures, including the following:

- total inpatient costs
- total outpatient costs
- total pharmacy costs
- total PH-related pharmacy costs
- MDD-specific costs
- other PH costs
- physical health costs
- medication costs
- psychotherapy costs.

Denominator

- (a) Total number of member-months attributable to beneficiaries with MDD
- (b) Total number of member-months attributable to all beneficiaries

Definitions

Measurement period. Period of time during which care is evaluated (e.g., monthly, quarterly, annually). In the more extensive discussion in Appendix D, we recommend an annual measurement period.

Member-month. Any month during the measurement period for which an applicable beneficiary (either all beneficiaries or those diagnosed with MDD) is enrolled in the TRICARE system and not separated from the military

Care Continuum

All phases

Measure Type

Resource use

Care Settings

Outpatient, inpatient, and residential

¹⁹ The documentation for this measure is similar to that specified for PTSD-RU2*. A more extensive description of this measure appears in Appendix D.

Measure Sources

Adapted from the following:

NQF, *National Voluntary Consensus Standards for Cost and Resource Use: Final Report*, April 2012a. As of September 13, 2013:

<http://www.qualityforum.org/Publications/2012/04/>

National_Voluntary_Consensus_Standards_for_Cost_and_Resource_Use.aspx

Sorbero, Melony E., Thomas E. Mannle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Rationale for Measure Inclusion

We recommend using average cost PMPM as a measure of resource use, focusing on the costs for beneficiaries who have been diagnosed with MDD during the measurement period. This measure can be calculated from existing MHS data, is used regularly by civilian-sector health plans to track the cost of treatment within managed care organizations (Kongstvedt, 2009), is similar to other measures that have been vetted by NQF (NQF, 2012a), and is similar to other measures regularly tracked already within the MHS. The measure calculates the cost of delivering PH treatment to a population diagnosed with MDD, taking into account that individuals can enter or exit the population during an analysis period.

Potential Data Sources

Numerator

Administrative claims and management data

Denominator

Administrative claims and management data

Feasibility

The denominator for this measure can be calculated from administrative claims data (beneficiaries with MDD) and information about TRICARE enrollment found in the MDR. The numerator is based on data readily available in the MDR.

Feasibility Code

Green

Additional Rationale for Cost Measure

This appendix provides an in-depth discussion of the creation of a cost measure for PH conditions in DoD. Much of this basic information is provided in the measure description (Appendixes B and C), but this discussion offers more-extensive detail. First, we include an introduction that presents key definitions related to the measurement of costs, specifies the focus of this appendix, outlines two important considerations when creating cost measures (i.e., objectives for measuring costs and key methodological issues), and recommends a specific cost measure for use by DCoE. Second, we include a detailed measure description section that outlines specifications for the recommended measure. Finally, the appendix concludes with a brief description of the benefits, limitations, and illustrative applications of the recommended cost measure.

Introduction

Definitions of Cost Measure

Varied terms have been used across the health care literature to describe concepts related to costs, including *cost* and *value*. This section defines these two terms, differentiates between them, and argues for focusing DCoE's related measure construction on costs rather than value.

Cost

Cost describes the level of expenditures related to resources used to produce a given unit of health care. This unit of health care can include anything from a year of TRICARE coverage to a single hospital stay. For example, TRICARE reported that, in fiscal year 2011, inpatient, outpatient, and pharmaceutical services cost the plan \$4,319 per active-duty member, on average. Also, the average cost of an outpatient encounter for PTSD was \$265 in that fiscal year (TRICARE Management Activity and Office of the Chief Financial Officer, 2012). For some purposes, a marginal cost figure is required instead of an average or median cost figure. Marginal cost measures the dollar value of the extra resources needed to add one unit of health care (e.g., one more hospital stay). Marginal cost is lower than average cost except when no further capacity exists.

Value

Value differs from cost by incorporating health outcomes into its definition. It has been defined in a variety of ways and usually means something different to various stakeholders. We chose a single unified definition that clearly and succinctly defines value as “outcomes relative to costs” (Porter, 2010). These outcomes can be defined in a variety of ways, such as the utility (Chung, Kaleba, and Wozniak, 2008) (i.e., outcomes adjusted for quality of life) or symptom-free days

(Lave et al., 1998). For example, a measure of value might report that pharmaceutical treatment for depression delivered in a primary care office costs between \$12.66 and \$16.87 per additional depression-free day (Lave et al., 1998).

Report Focus

Importantly, this report focuses on the development of cost measures rather than on value measures. Because value measures implicitly incorporate outcomes, these measures require a level of theoretical development and computational burden above and beyond the mere calculation of costs. Although researchers have noted the importance of developing measures of value, those measures (particularly those related to mental health) remain in their nascent stages. One particular strain of literature has focused on estimating the value of depression care particularly for the cost of producing outcomes related to depression (e.g., depression-free days [Domino et al., 2008; Katon et al., 2006; Lave et al., 1998; Simon et al., 2001] or quality-adjusted life-years, or QALYs). However, we could find no evidence that measures used in these studies (e.g., cost per change in QALYs or cost per change in depression-free days) have been adopted or endorsed by any organization for routine monitoring purposes. To develop measures of value, the outcome of care (e.g., change in symptom-free days) must be clearly linked to the cost incurred for that care. Also, the outcome measures used in the literature to develop value measures require longitudinal primary data collection from patients over the course of their treatment, and all of the costs in a research study are often known to the researcher. In a research setting, this is possible but extremely difficult in a regular quality monitoring system.

Because there are no easily applied value measures for routinely monitoring PH care, we would need to develop novel measures to apply these concepts to the MHS. The cost measure that we recommend in this report could conceivably be tied to one or more of the outcome measures described in this report to form a value measure. However, such a project would require significant time, resources, and research that are beyond the scope of this effort. Therefore, at the present time, we suggest separate measures of costs and outcomes.

Objectives for Measuring Cost

The selection of the most appropriate cost measurement approach depends largely on the objectives for which the resulting cost measures will be used. Some of these possible objectives within the MHS include the following: (1) comparing the cost of care across the military services or MTFs or TRICARE regions to identify unnecessary variation and opportunities to reduce costs, (2) comparing costs between direct care (e.g., care provided at MTFs) and purchased care (e.g., care provided in the civilian sector), and (3) monitoring the costs of care across all providers (e.g., tracking the costs over time). Later in this appendix, we discuss how our proposed cost measure might be used specifically related to PTSD and MDD.

Overview of Key Methodological Issues When Measuring Cost

Multiple methodological issues guide selection of cost measures. Two key methodological issues are considered here. First, we discuss measure type. These measure types are differentiated according to the data used to construct them and the measure's relative ability to specify or estimate the actual cost of care. Second, we also consider the measure's level of aggregation. In this setting, level of aggregation is considered to be the unit at which costs are reported. For example, costs might be reported as total cost estimate across all members, across members diagnosed with a particular PH condition, or across a treatment episode.

Measure Type

It is important to note that cost measures are a particular type of resource use measure, which relate to the resources expended during the care of a condition (Romano, Hussey, and Ritley, 2010). Resource use measures can be organized into three general classifications: (1) utilization measures, (2) standardized resource use measures, and (3) cost-of-care measures. They differ in the extent to which they are able to capture four factors that drive the expenditures used to deliver a specific type of health care:

- the number of beneficiaries who receive a certain type of care and the amount of those services that they receive
- the diagnostic and treatment services provided to these beneficiaries
- the resources (e.g., provider time, clinical support staff time, equipment and supplies) used in providing the diagnostic and treatment services
- the cost of the resources used plus any overhead costs (e.g., administration).

Utilization measures provide general information about the number of beneficiaries who receive certain types of services and the amount of those services that they receive. This type of measure is commonly used to track service use in the civilian and military health care systems. Examples include general measures, such as total and average inpatient days or numbers of physician visits, and more-specific measures, such as average number of visits for a particular diagnosis or procedure. One advantage of utilization measures is that they are relatively easy to generate because the data are typically available in administrative databases. However, these measures do not capture the labor, supplies, and other resources used to produce the services or the cost of the resources. In this way, utilization measures might be considered, at best, coarse proxies for cost.

Standardized resource use measures adjust utilization for case mix by incorporating resources that are judged to be appropriate for producing health care. This type of measure is often expressed in relative-value units (RVUs), which are independent of actual prices. RVUs standardize services and count them based on the relative value of resources used to produce each service. RVUs were developed for Medicare payment; for each medical procedure, RVUs capture expert judgment of the physician skill and time required to deliver those services and data on other provider expenses. Thus, they represent the expected relative resources used to perform each procedure across the United States. RVUs are widely used to translate utilization-based measures into a measure of resources (Glass and Anderson, 2002). One advantage of using RVUs is that they can be executed in settings in which comprehensive, accurate cost data are not available. For example, the direct-care system is an in-house provider, so no payment transactions are recorded for the care provided. However, direct-care-system RVUs can be estimated based on the number and intensity of the services provided. Also, RVUs can be used to compare resource use across MTFs that have different support service structures. However, it is important to note that RVUs do not measure actual costs expended by a specific provider. Specifically, because RVUs are based on information about the resources used within the civilian health care sector, they may not reflect the relative resources used to provide health care within the MHS, especially within the direct-care system. So, they do not reflect differences in efficiency or cost structure between the MHS and other providers or for specific MTFs within the MHS, and they cannot be used for budgeting purposes. Also, standardized resource use measures are not necessarily intuitive to users and may be difficult to interpret. Furthermore, some

have argued that RVUs specifically do not capture the resource use necessary to deliver mental health services compared with physical health services and may undervalue mental health.

The *costs of care* to the payer can also be calculated. These costs can be relatively easily tracked in a third-party reimbursement system (e.g., the purchased-care sector in TRICARE), in which payments are based on a detailed list of services provided. Tracking these costs is more difficult in the direct-care system of the MHS, in which the actual costs of producing an admission or office visit (both direct and indirect costs) are less easily determined. The MDR attaches an estimate of the cost to each clinical encounter record based on aggregate data on labor and other operating costs collected at the facility and clinic levels. The MDR cost measure includes only operations and maintenance costs incurred at the MTFs. Costs incurred elsewhere in DoD (e.g., for personnel or contracting support, laundry, food, grounds keeping, or centralized contracts) can differ systematically across MTFs. For example, some MTFs obtain support services from their installations, whereas other MTFs must provide the same services internally. Therefore, the differences in costs for a particular PH condition across MTFs or military services may reflect these support service structures rather than any differences in efficiencies or care patterns.

Actual costs also capture some differences that reflect local conditions that may be unrelated to underlying efficiency in delivering care. Well-known examples are local labor market wage condition, relevant for civilian and contract employees in the MTFs, added costs attributable to teaching in facilities with residency and other training programs, and added costs attributable to military readiness activities. Depending on the purpose or purposes for which the cost measure will be used, adjustments for these differences may be indicated.

The issues in accurately and appropriately measuring costs become very important when comparing costs across purchased versus direct care or even across military services. The claim records for purchased-care services by definition represent the full cost to DoD of purchased care, whereas the direct-care cost measure is incomplete. Comparing costs across these settings would require extensive analysis to identify all costs incurred for MTFs. If the purpose of such a comparison is to determine whether care should be shifted between the two sectors, an appropriate measure of the marginal cost difference associated with increasing or decreasing MTF service volume would need to be calculated.

Level of Aggregation

An important distinguishing characteristic of different cost measures is the level at which the measures is calculated and reported. The two levels at which cost measures can be calculated and reported are (1) population level and (2) episode level. The population-level measures can be used to estimate per capita costs during a specified period of time for the covered population (e.g., TRICARE beneficiaries) or for a subset of that population (e.g., beneficiaries with PTSD). These per capita measures are generally expressed as the average costs for a member, enrollee, or beneficiary (often referred to as PMPM) or a (treated) patient per month. The measure may refer to a different time period, such as per member per year (PMPY). *Member* typically refers to all individuals enrolled in a health plan, but it may also refer to a subgroup of members with specific personal characteristics or health conditions, such as those members with MDD or PTSD.

The episode-level approach can be used to estimate the average cost required to care for a single case. For example, we might identify a new case of PTSD and measure the costs from

onset to resolution or the costs from onset over a period of time after onset, predetermined based on the usual course of the condition and its treatment.

Population- and episode-level approaches answer different questions. The population-level approach can be used to answer the following question: “How much did it cost per beneficiary per month to treat PTSD last year?” This approach can capture differences in the prevalence of PTSD and care-seeking behavior in addition to resources used in the delivery of care for those who seek care. Conversely, an episode-level approach could answer the following question: “How much did it cost to treat a single case of PTSD last year?” Episode measures focus only on the delivery of care for those who seek care. There are clear advantages to each approach, and whether to adopt one approach or the other, or both approaches, is based primarily on the potential use of the cost measures. One potential disadvantage of using a population-level measure (particularly if average costs are calculated across all TRICARE beneficiaries as opposed to only those diagnosed with a certain PH condition) is that there may be a disincentive to screen given that costs would rise as the number of diagnoses increases. One distinct disadvantage of using an episode-level approach is that this approach may require the development of sophisticated algorithms to determine the beginning and end of any one episode.

Recommended Measure: Average Costs per Member per Month

In order to recommend a measure, we first reviewed the academic literature, gray literature, and health care quality measure databases (e.g., NQMC). We looked for any cost measure that was specifically used to measure costs of care for PH conditions. From that search, we generated a list of potential cost measures. Second, we conferred with DCoE staff to better understand the priorities of the organization related to cost measures, particularly those about the potential use of the cost measure. We developed the following criteria for choosing the cost measure based on the literature, measures commonly used outside the MHS, and the data available to the MHS:

- data available through MHS data sets
- measure (or similar measure) vetted by a national organization, such as NQF or AHRQ
- measure used commonly among other health plans to measure the cost of care
- similarity to a measure already used within the MHS
- measure intuitively understandable by a wide range of stakeholders
- measure allowing for attribution of costs as accurately as possible at the MTF level.

Informed by these criteria, we recommend a PMPM cost measure, focusing on beneficiaries who have been diagnosed with a given PH disorder. This measure can be used to derive the costs of treating any particular PH disorder. However, in order to simplify the language, we use PTSD and MDD as examples. This measure can be calculated from existing MHS data, is used regularly by civilian-sector health (Kongstvedt, 2009) to track the cost of treatment, is similar to other measures that have been vetted by (NQF, 2012a), and is similar to other measures regularly tracked already within the MHS. The measure calculates the cost of delivering PH treatment to a population diagnosed with PTSD or MDD, taking into account that individuals can enter or exit the population during an analysis period. If the time period used is appropriately selected, the measure can be defined at the MTF level for that MTF’s enrolled patients or for the patients treated by the MTF. The following section describes how this measure could be calculated in the DoD context.

Measure Description

Measure Description

The recommended measure is calculated as the total costs of all diagnostic or treatment health care resources used by members with PTSD or MDD, divided by the number of months for eligible beneficiaries.¹

Numerator

The summation of all costs of care delivered to beneficiaries with PTSD or MDD during eligible member-months. A beneficiary would be considered eligible for the measure if he or she had at least one inpatient admission or two outpatient visits for which the condition is listed as a diagnosis at any point in the 12 months prior to the reporting period. For example, if the PMPM measure is to be reported as of January 1, 2013, for a beneficiary who had one PTSD-related inpatient admission or two PTSD-related outpatient visits between January 1, 2012, and December 31, 2012, his or her costs would be included in the numerator. If a service member had both PTSD and MDD, he or she would be included in both denominators. This inclusion criterion has been used in other previous studies of PTSD and MDD utilization (Watkins, Pincus, Smith, et al., 2011).

In this case, the PMPM measure should be considered a “total cost of care” measure. However, observing total cost of care PMPM provides limited information on its own. One benefit of using the PMPM measure is that the data can be cut in a variety of ways to help provide key information on why any differences are observed either over time, across MTFs, across settings, or across different care locations. Therefore, we propose that eight other PMPM measures are also developed, including the following:

- total inpatient costs
- total outpatient costs
- total pharmacy costs
- total PH-related pharmacy costs
- PTSD-related costs or MDD-related costs
- other PH costs
- physical health costs
- medication costs
- psychotherapy costs.

Importantly, the construction of the cost estimates in the encounter data is developed differently depending on where the encounter occurred: direct care or purchased care.

Direct Care

Direct care includes any encounter that occurs within a DoD MTF. Because the vast majority of costs related to MTFs are fixed (because most military personnel are required for readiness), costs in this setting are calculated by apportioning those fixed costs to each individual service. These costs are apportioned based on standard work unit measures, essentially measuring the

¹ It is important to note that *members* in *PMPM* does not refer strictly to service members. In this context, *members* refers to all TRICARE-enrolled beneficiaries more broadly and not strictly service members. So *member* and *beneficiary* are used interchangeably in this appendix.

amount of work that it takes to deliver a given service. In the inpatient setting, each service is assigned a relative weighted product (RWP), while, in the ambulatory setting, the services are assigned an Ambulatory Patient Group (APG). RWPs and APGs are weights expressed as ratios, wherein higher numbers represent services that require more-extensive resources to deliver. For each service, the RWP and APG are then multiplied by standard dollar value (different for inpatient and outpatient) to derive the cost of the service. However, it is important to note that only a fraction of all costs of care delivery are captured by the standard dollar values. As we indicated above, some support service costs are incurred by the installation or elsewhere within DoD; these are not captured in the standard dollar value of MTF care. For example, cost associated with headquarters management, Health Affairs, TRICARE Management Activity, service surgeons general, service information management, service intermediate commands, and central information systems are not currently included in the cost calculation. Therefore, MTF care costs are, to an unknown extent, underestimated.

Finally, pharmacy costs are calculated based on the DoD full transaction cost and include the associated member copayments, dispensing costs, and ingredient costs.

Purchased Care

Purchased-care costs represent payments for services provided through TRICARE. In this setting, the costs represent transacted costs between the three TRICARE contractors and civilian providers, most of which have contracted to be in the TRICARE network. These transacted costs are negotiated with the TRICARE contractors and, for the most part, are based on the Medicare fee-for-service fee schedule.

Denominator

For this measure, the denominator is equal to the total number of applicable member-months attributable to eligible beneficiaries. Importantly, eligible beneficiaries can either be all beneficiaries or be only beneficiaries with PTSD or MDD (i.e., diagnosed with a given condition). Using these different denominators, these two measures largely balance each other and provide different and unique information. Most substantially, the two measures allow for an examination of the extent to which cost growth is driven by more-expensive cases (disease-specific denominator) as opposed to more diagnoses (all-beneficiaries denominator).

Determining the number of eligible member-months for each person in the denominator is much more challenging. First, member-months should be counted only when a service member is covered by TRICARE on the 15th of any given month. Using the 15th of the month is similar to the approach used in commercial health plans to determine eligibility. If a member was on active duty on January 1, 2012, but discharged on July 17, 2012, the number of member-months for this service member would be counted as seven. Certain members and their dependents retain TRICARE coverage for up to 180 days after separation; these include reserve component members who have been activated for more than 30 days and active component members who are involuntarily separated. Because these individuals may have other coverage during this period from VA or their civilian employers, they should not be included in the cost measure calculation.

Although determining member-months based on TRICARE coverage is relatively straightforward, determining which months should be considered eligible based on clinical criteria is less straightforward. We could count eligible months using two methods: (1) count all months in the year as eligible (all months) or (2) count two months before the initial diagnosis

and all months after initial diagnosis (fraction of months).² The all-months approach is a more traditional approach to calculating PMPM costs, whereas the fraction-of-months approach is less traditional and would be considered more similar to an episode-level measure.

Each of these approaches has potential limitations. The all-months approach would include months before a beneficiary developed the health condition or first sought care. This would overstate the eligible member-months. On the other hand, for beneficiaries who develop PTSD but do not quickly begin a course of treatment, the fraction-of-months approach could undercount the eligible months. For example, imagine that a beneficiary first experiences significant symptoms of depression in January 2012 but is not diagnosed with MDD until July 2012. In this case, January through July should be considered eligible months, but, in the fraction-of-months approach, only two would be included. Finally, there is no precedent in the literature for a fraction-of-months approach, and calculating such a measure would likely require extensive coding. Therefore, we would largely be inventing a whole new approach. Conversely, there are NQF-endorsed measures that use similar all-months inclusion criteria for different conditions, such as diabetes and asthma (NQF, 2012a).

The choice of denominator approach should be guided by a variety of considerations. For example, the characteristics of the disease may drive which months we might consider eligible. For example, to the extent that PTSD and MDD are considered chronic conditions, we would likely err on the side of considering more months as eligible. Meaning that, even when the individual has no records of PTSD or MDD diagnoses, he or she still has the condition, and the months between claims should still be counted. In contrast, to the extent that the conditions are considered episodic, the months between PTSD or MDD claims should not necessarily be counted. However, estimating the beginning and end of any particular episode of care can require complicated algorithms and significant data.

Furthermore, the purpose of the measure should also guide the choice of the approach. To the extent that the measure will be used to generally observe trends in costs over time, the exact approach is less important as long as it is consistent over time. To the extent that the measure will be used for budgeting and an exact estimate of the costs is important, a more precise approach should be developed.

Recognizing the limitations of each approach, we suggest calculating the denominator two ways. First, we suggest creating a PMPM measure that includes in the denominator only beneficiaries with PTSD or MDD and using the all-months approach, wherein each month that a beneficiary is covered by TRICARE will be included in the denominator. We have decided to suggest the all-months approach because we believe that PTSD and MDD are chronic disease-like and that beneficiaries likely utilize significant health care services in the months before an official diagnosis. In other words, we would rather err on the side of including months in which a beneficiary may not actually have symptoms than to exclude months in which he or she did have symptoms but was yet to be diagnosed with PTSD. This all-months approach is also consistent with the approach used in a VA study of PH costs (Watkins, Pincus, Smith, et al., 2011). Second, we also suggest calculating another PMPM measure also using the all-months approach but using all beneficiary months as the denominator rather than those

² This is only an illustrative example. There is no standard for determining which months should be counted in a fraction-of-months measure. A measure could be constructed that incorporates three months or no months before the initial diagnosis.

months of the beneficiaries with PTSD or MDD alone. As we described above, these two measures provide different and equally useful information about costs.

Exclusions

Finally, we suggest including in the denominator only those beneficiaries who have TRICARE Prime as their primary coverage. Only a portion of costs for beneficiaries with other coverage sources can be observed in the claims data. As shown in Table D.1, the source of care varies considerably across the TRICARE beneficiary population depending on the TRICARE plan in which a beneficiary is enrolled, his or her primary care assignment (TRICARE Prime), and other health coverage. For some groups, TRICARE is not their only coverage, and the costs captured in the MDR are incomplete. The recommended cost measure can be defined for different subpopulations but, within the MHS, is most often defined for TRICARE Prime enrollees because, with rare exceptions, they get all their care through TRICARE. The recommended measures for TRICARE Standard/Extra or Medicare enrollees only capture the share of health care costs paid by DoD and are heavily influenced by the extent of reliance on other insurance coverage, including Medicare and employer insurance. Because TRICARE covers most health care services, the costs measured in the MDR for beneficiaries without other insurance are likely to be fairly complete. However, other insurance coverage is recorded only for those beneficiaries who report having it or use TRICARE as a second payer. Some beneficiaries for whom no insurance coverage is recorded in the MDR may, in fact, have other coverage, e.g., in a civilian health maintenance organization. Therefore, separating retirees and dependents under age 65 into those with and without other coverage involves some degree of error.

Type of Measure

Resource use

Table D.1
TRICARE and Other Insurance Coverage and the System of Care

Coverage	System of Care		
	MTF	TRICARE Purchased Care	Other Civilian Care
TRICARE Prime enrolled with an MTF	Primary source	Referral by the MTF	—
TRICARE Prime enrolled with a civilian primary care provider	Referral by the civilian provider	Primary source	—
TRICARE Standard/Extra with no other health coverage	If space is available	Primary source	—
TRICARE Standard/Extra with other health coverage	If space is available		Primary source; TRICARE is second payer
TRICARE for Life (Medicare eligible)	If space is available		Primary source; TRICARE is second payer

Data Source

The data necessary to calculate the recommended measure can be found in the MDR data sets. The MDR is a comprehensive source of data available to MHS researchers. The MDR includes information on all health care events that have been paid for or provided through TRICARE. The data set includes information on each beneficiary, including coverage information and demographics, as well as a personal identifier that allows the beneficiary information to be linked to clinical encounter data. Similar to other civilian administrative data sets, the clinical encounter data include provider information, dates of care, diagnostic and procedure codes, and the cost of the encounter.

Measure Sources

Adapted from the following:

NQF, *National Voluntary Consensus Standards for Cost and Resource Use: Final Report*, April 2012a. As of September 13, 2013:

<http://www.qualityforum.org/Publications/2012/04/>

National_Voluntary_Consensus_Standards_for_Cost_and_Resource_Use.aspx

Sorbero, Melony E., Thomas E. Mannle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Feasibility

The denominator for this measure applies to all patients with a diagnosis for a PH condition and can be identified with administrative data. The numerator is based on data readily available in the MDR.

Feasibility Code

Green

Benefits and Limitations of per Member per Month Measure

The proposed PMPM treatment cost measure has a variety of important benefits and limitations.

Benefits

The PMPM measure was chosen because it offers a cluster of unique benefits. First, this measure is particularly useful for tracking the costs of delivering PTSD care over time in order to generally monitor the costs of care. This would be especially useful for inclusion in a dashboard that also includes information about the quality and outcomes related to PTSD and MDD care. Also, because the measure would be calculated using different numerators and denominators, this measure provides a full view of the various factors driving costs. Second, the data from this measure can be taken directly from extant data sources. The relevant variables are included in the MDR, which is accessible to and used regularly by DoD researchers and analysts.

Third, the PMPM measure considers the total cost of care for the population of interest in addition to a more focused cost of care targeted at specific conditions. This characteristic is especially useful for conditions, such as PTSD and MDD, that are associated with co-occurring mental and physical health conditions. For example, individuals who suffer from PTSD are much more likely than others to suffer from depression, substance abuse disorders, and lower

levels of physical health status (Kessler, Sonnega, et al., 1995; Schnurr and Jankowski, 1999). If the cost of PTSD or MDD were measured strictly by the cost of direct treatment of those conditions, the true cost of the conditions would be drastically understated. Conversely, although the PMPM measure (by construction) assesses the total cost of care, the measure can also be restricted to costs related to a smaller set of conditions. For example, one could estimate the costs of PTSD-specific care, other behavioral health care, and physical health care among beneficiaries with PTSD, as well as the costs for outpatient versus inpatient care or pharmacy versus psychotherapy. This makes the PMPM measure a diverse measure. PMPM measures are also well suited to account for the fact that beneficiaries regularly leave military service and TRICARE. Therefore, this measure can measure costs even based on partial-year data.

Fourth, PMPM is also used frequently by civilian managed care organizations, as well as in the MHS. PMPM is arguably the most common measure used to track costs in managed care systems (Kongstvedt, 2009). Consequently, PMPM costs are the basis for calculating the amount that Medicare pays to managed care organizations for providing care to their beneficiaries (CMS, 2007). Because PMPM is used commonly, a variety of NQF-endorsed PMPM measures can be used to help guide measure development. Although these NQF-endorsed measures pertain to different conditions (e.g., diabetes and heart disease), they can still be used to help guide measure development in this setting. The PMPM measure is also used extensively within the MHS (Watkins, Pincus, Smith, et al., 2011; Opsut, 2010). For example, the Office of the Assistant Secretary of Defense for Health Affairs has been monitoring the cost of care for TRICARE beneficiaries using a PMPM measure that incorporates costs from inpatient, outpatient, pharmacy, and ancillary care encounter data. Therefore, the information gathered by DCoE can be compared with that gathered by others in DoD, as well as compared externally to data collected by other, non-DOD care providers or health plans.

Limitations

One important limitation of this measure is that, for many beneficiaries, it incorporates costs for care provided in both the direct- and purchased-care systems. As mentioned earlier, the costs for direct care are not as comprehensive as purchased-care costs are. Further, the same mix of services is not provided to a beneficiary in the two sectors, and the beneficiaries treated in the two sectors are likely to differ in their demographic and clinical characteristics. Therefore, the readily computed cost measure does not provide accurate comparisons of costs between direct and purchased care and across MTFs with different mixes of purchased and direct care without adjusting the direct-care figures to account for omitted costs and patient mix. This is not, however, a characteristic that is unique to PMPM measures per se. Instead, this issue would apply to any measure that combines cost data from the MDR for the direct- and purchased-care systems.

The PMPM measure is limited to data generated during interactions with the health care system. The likelihood that a significant amount of care for PTSD occurs outside of the health care system is high. For example, many service members and dependents visit clergy for counseling and guidance or participate in peer-to-peer counseling. This is likely especially true for prevention and screening activities that occur outside of the traditional outpatient encounter or are otherwise not submitted and recorded in claims. Some beneficiaries, especially retirees and dependents, may rely on nonmilitary sources of care. However, this is again an issue that pertains to any measure that uses the MDR data system to measure costs and is not specific to the PMPM measure.

Also, PMPM cost estimates are driven significantly by patients' health status and other clinical factors. Because each MTF and military service cares for a different patient population with variation in condition severity and comorbidities, the cost measure would have to be risk-adjusted to account for these differences across MTFs if their costs were to be compared. Furthermore, risk adjustment may also be necessary when observing changes in costs over time to the extent that patient risk changes over time, as is likely the case before, during, and after a war. Presently, DoD uses a risk-adjustment system to adjust the total PMPM for care delivered at each MTF by calculating the number of equivalent lives that each member represents. An equivalent life is calculated as a ratio of the member's utilization rate and disease burden to that of the average member (Opsut, 2011). However, this equivalent-life approach may be insufficient for risk-adjusting costs related to PH-specific PMPM measures. In this case, a risk-adjustment approach that specifically accounts for differences in spending for PH conditions would be needed. However, such a measure does not exist. Therefore, we recommend caution when using these measures to compare costs across MTFs or services; if this comparison is to be made, the risk-adjustment procedure is essential for a meaningful comparison.

Also, many of the PH quality measures that we recommend are used to address underuse of services, meaning that patients are not receiving all of the services that would be recommended for them. Therefore, in this case, to achieve better quality, more services would be delivered, which translates into higher costs. However, these services may lead to a reduction in the costs for related medical conditions or improvements in functioning. In this case, the higher costs should be assessed along with the measured changes in other health care costs and functioning. This is an argument for the future development of value measures that might demonstrate the outcomes that are achieved for any increase in costs.

Although above we treat PTSD and MDD as chronic-like diseases, in many ways, they might also be considered episodic. Therefore, beneficiaries may technically have PTSD or MDD at any time during the measurement period. For example, a patient may have a flare-up of symptoms, which subside within three months. In this case, the individual should be considered to have PTSD or MDD for only three months. However, the symptom subsidence is not observable in the data, and we cannot differentiate this individual from a patient who remains symptomatic but does not receive care. Therefore, given the specific approach used to construct the inclusion criteria, the patient may be included in the data set for more than three months. In this case, the total costs may be over- or underestimated, depending on the patient's care patterns in those ineligible months. Alternative approaches would need to be tested on MDR data, perhaps checking a sample of cases with the more-detailed medical record data, to identify the optimal criteria for identifying which months should be included in the analysis.

Applying Average Costs per Member per Month to Posttraumatic Stress Disorder and Major Depressive Disorder

In this section, we discuss ways in which the average cost PMPM measure can be applied specifically to PTSD and MDD. We discuss these potential applications with particular consideration of the benefits and limitations explored above.

First, the PMPM measure might be used to compare the cost of care across MTFs to identify variation and associated opportunities to reduce costs. To do this, DCoE could calculate the PMPM cost measure for each MTF and identify those that are expending a particularly high or low level of dollars PMPM on PTSD or MDD care. In this case, the overall cost measure would simply be used as a signal to identify potential unnecessary variation, and DCoE

staff would then be able to more specifically investigate the sources of variation. Such variation might be driven by actual cost differences (e.g., differential likelihood to be treated inpatient versus outpatient) or by relative differences across the MTFs in terms of cost structures or patient risk. However, because much of the cost variation across MTFs may be driven by differences in relative patient risk and the specific cost structures of the MTFs, the cost measure should be used with caution for this purpose. Appropriate methods must still be developed to ensure accurate comparisons across MTFs. In the future, if these significant shortcomings could be addressed, cost measures might even be used to identify MTFs that are delivering care most efficiently. Again, this highlights the potential usefulness of developing a value measure, in which MTFs might be compared based on the outcomes (e.g., symptom-free days) that are produced per dollar spent. Such a measure could potentially be constructed using outcome measures presented in this report. However, as noted previously, much more work would be required to develop such measures.

The proposed cost measure might also be used to compare costs between direct and purchased care. For example, DCoE staff can also calculate the relative cost PMPM for beneficiaries being cared for primarily in the direct-care system versus the purchased-care system. These comparisons might be used to identify inefficiencies in either delivery system and to identify opportunities for improving efficiencies within MTFs or (conversely) identifying opportunities to shift PTSD or MDD care out of purchased-care systems and into the MTFs. However, the costs reported in the MDR for MTF encounters are not estimated using methods designed to support comparison with civilian network costs, and the revised methods would be required for this purpose. Also, beneficiaries cared for in the purchased-care system likely do not have the same severity as those cared for at MTFs, necessitating risk adjustment. Therefore, at this time, DCoE should hesitate to use the cost measure for MTF–civilian comparisons.

The proposed cost measure is likely most useful for tracking expenditures across all providers and over time in order to provide a general estimate of the expenditure outlays at any given time and to understand cost trends. These estimates might be used to inform budgeting and resource allocations by understanding where and how money is being spent and what DCoE can expect about expenditures in the future. When this cost measure is used to examine changes in costs over time, users should be cognizant of changes in patient risk over time, which are likely as the military enters and exits varying levels of mobilization. DCoE staff can further investigate other potential sources of cost changes over time and determine whether those changes are potentially beneficial (e.g., increasing costs due to improved screening) or not (e.g., increasing costs due to growing inefficiencies). Again, this also highlights the potential usefulness of a value measure that can be used to identify reductions in the cost per outcome. However, as mentioned previously, significant work is required before a robust value measure could be developed and deployed.

Candidate Set of Measures of Quality of Care for Posttraumatic Stress Disorder

Table E.1
Candidate Set of Measures of Quality for Posttraumatic Stress Disorder

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-S1 ^b	Percentage of patients screened for PTSD, MDD, and alcohol misuse and, if positive, appropriate follow-up initiated	(a) Number of patients screened annually for PTSD, MDD, or alcohol misuse using a standardized tool (b) If positive screen, follow-up initiated within seven days	(a) All patients (b) Patients who screened positive for PTSD, MDD, or alcohol misuse	Screening	Process	Outpatient
PTSD-A1	Percentage of PTSD patients in a new treatment episode with assessment of symptoms with the PCL	Patients who have an assessment of PTSD symptoms within the first 30 days of a new treatment episode using the PCL	PTSD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential
PTSD-A2	Percentage of PTSD patients in a new treatment episode assessed for depression	Patients who are assessed for comorbid depression within 30 days of the new treatment episode	PTSD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential
PTSD-A3 ^b	Percentage of patients in a new treatment episode assessed for suicide risk	Patients assessed for suicide risk during the same visit in which a new treatment episode was identified or in the 14 days prior	PTSD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential
PTSD-A4 ^b	Percentage of PTSD patients in a new treatment episode assessed for recent substance use	Patients who have an assessment of recent substance abuse, including type, quantity, and frequency, within the first 30 days of the new treatment episode	PTSD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential

Table E.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-T1	Percentage of PTSD patients with symptom assessment with the PCL during the four-month measurement period	Patients who had the PCL administered at least once during the four-month measurement period	PTSD patients with an encounter within each four-month period	Treatment	Process	Outpatient
PTSD-T2 ^b	Percentage of days when third available specialty-care appointment is within two days	Days in the measurement period on which third available specialty-care visit is within two days	Days in the measurement period	Treatment	Structure	Outpatient
PTSD-T3 ^b	Percentage of PTSD patient contacts with SI with appropriate follow-up	Documentation of appropriate follow-up for SI, intent, or behavior	Outpatient visits or contacts in which PTSD patients endorsed SI, intent, or behavior	Treatment	Process	Outpatient
PTSD-T4 ^b	Percentage of PTSD patients with a documented treatment plan	Patients with a treatment plan in the medical record	Patients with PTSD	Treatment	Process	Outpatient, inpatient, and residential
PTSD-T5	Percentage of PTSD patients with an adequate trial of SSRIs and SNRIs	(a) Patients who receive a trial of SSRIs for at least 60 days (b) Patients who receive a trial of SSRIs for at least 60 days or have a documented reason for discontinuing SSRI treatment in less than 60 days	Patients with PTSD with a new prescription for an SSRI or SNRI	Treatment	Process	Outpatient
PTSD-T6	Percentage of PTSD patients newly prescribed an SSRI or SNRI with a follow-up visit within 30 days	Patients who have a follow-up visit within 30 days of a new prescription for an SSRI or SNRI	Patients with PTSD with a new prescription for an SSRI or SNRI	Treatment	Process	Outpatient

Table E.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-T7	Percentage of PTSD patients who receive evidence-based psychotherapy	(a) Patients who received any evidence-based psychotherapy visits (b) Number of evidence-based psychotherapy visits received	(a) All PTSD patients (b) Patients with a PTSD diagnosis who are receiving psychotherapy	Treatment	Process	Outpatient
PTSD-T8 ^b	Percentage of PTSD patients in a new treatment episode who received any psychotherapy	Patients receiving any psychotherapy within four months after starting a new treatment episode	PTSD patients in a new treatment episode	Treatment	Process	Outpatient, inpatient, and residential
PTSD-T9 ^b	Percentage of PTSD patients who had four psychotherapy visits or two medication-management visits within the first eight weeks	Patients who had four psychotherapy visits or two medication-management visits within eight weeks of the new treatment episode	PTSD patients in a new treatment episode	Treatment	Process	Outpatient
PTSD-T10	Percentage of PTSD patients with response to treatment at six months	Patients who have a documented reduction of at least five points on the PCL within six months (plus or minus 30 days)	PTSD patients with a PCL score that is positive for PTSD (PCL score of at least 30)	Treatment	Outcome	Outpatient
PTSD-T11	Percentage of PTSD patients with response to treatment at 12 months	Patients who have a documented reduction of at least five points on the PCL within 12 months (plus or minus 30 days)	PTSD patients with a PCL score that is positive for PTSD (PCL score of at least 30)	Treatment	Outcome	Outpatient

Table E.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-T12	Percentage of PTSD patients in PTSD-symptom remission at six months	Patients with a PCL score indicative of PTSD remission (PCL score less than 28) within six months (plus or minus 30 days)	PTSD patients with a PCL score that is positive for PTSD (PCL score of at least 30)	Treatment	Outcome	Outpatient
PTSD-T13	Percentage of PTSD patients in PTSD-symptom remission at 12 months	Patients with a PCL score indicative of PTSD remission (PCL score less than 28) within six months (plus or minus 30 days)	PTSD patients with a PCL score that is positive for PTSD (PCL score of at least 30)	Treatment	Outcome	Outpatient
PTSD-T14 ^b	Percentage of PTSD patients in a new treatment episode with improvement in functional status at six months	Patients with improvement in functional status from first visit for PTSD to six months after the first visit	PTSD patients in a new treatment episode who have at least two measures of functional status during the first six months	Treatment	Outcome	Outpatient
PTSD-T15 ^b	Percentage of psychiatric inpatient hospital discharges of patients with PTSD with follow-up	Inpatient psychiatric discharges followed with a visit within (a) 30 days (b) seven days	PTSD patients discharged from an acute inpatient setting with a primary mental health diagnosis	Treatment	Process	Outpatient
PTSD-T16 ^b	Number of inpatient psychiatric beds available per 10,000 active-duty service members	Number of inpatient psychiatric beds available at the MTFs across the measurement period	Among MTFs with psychiatric inpatient capacity, number of active-duty service members served by the MTF, divided by 10,000	Treatment	Structure	Inpatient

Table E.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-PE1 ^b	Percentage of PTSD patients who report getting treatment quickly	Number of patients who report “always” getting treatment quickly or getting treatment as soon as it was needed (a) getting help by phone (b) getting urgent treatment as soon as it was wanted (c) getting an appointment as soon as it was wanted	PTSD patients who received behavioral health care services and who answered “getting treatment quickly” questions on the ECHO survey	All phases	Patient experience	Outpatient
PTSD-PE2 ^b	Percentage of PTSD patients who report being given as much self-management information as they wanted	Number of patients who report that they received enough information to manage their own conditions	PTSD patients who received behavioral health care services and who answered “getting enough information to manage the conditions” question on the ECHO survey	All phases	Patient experience	Outpatient
PTSD-PE3 ^b	Percentage of PTSD patients who report being told about treatment options	Number of patients who indicate that they received information about treatment options (a) self-help or consumer-run programs (b) different treatments available for the conditions	PTSD patients who received behavioral health care services and answered the “information about treatment options” questions on the ECHO survey	All phases	Patient experience	Outpatient

Table E.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-PE4 ^b	Percentage of PTSD patients who report being helped by counseling or treatment received	Number of patients who report being helped a lot by the counseling or treatment received	PTSD patients who received behavioral health care services and answered the “amount the treatment helped” question on the ECHO survey	All phases	Patient experience	Outpatient
PTSD-PE5 ^b	Percentage of PTSD patients who report being better than they were one year ago	Number of patients who report being much better than they were a year ago with (a) ability to deal with daily problems (b) ability to deal with social situations (c) ability to accomplish things (d) ability to deal with symptoms or problems	PTSD patients who received behavioral health care services and answered the “perceived improvement” questions on the ECHO survey	All phases	Patient experience	Outpatient
PTSD-PE6 ^b	Percentage of patients who rated counseling and treatment received as 9 or 10 (out of 10)	Number of patients whose overall rating of the counseling or treatment they received as a 9 or 10 (on a 0-to-10 scale)	PTSD patients who received behavioral health care services and answered the “overall rating of counseling and treatment” question on the ECHO survey	All phases	Patient experience	Outpatient
PTSD-RU1 ^b	Number of psychiatric inpatient admissions per 100 patients with PTSD	Number of psychiatric admissions during the measurement period	Number of patients with a PTSD diagnosis, divided by 100	All phases	Resource use	Outpatient

Table E.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
PTSD-RU2 ^b	Total costs of health care resources used by TRICARE beneficiaries (members) with PTSD, divided by number of months enrolled (member-months)	Total costs required for treating beneficiaries with PTSD for all their member-months during the measurement period	(a) Total number of member-months attributable to beneficiaries with PTSD (b) Total number of member-months attributable to all beneficiaries	All phases	Resource use	Outpatient, inpatient, and residential

^a Definitions of terms used to characterize quality measures are provided in Appendix A.

^b This measure could be applied to patients with PH conditions other than PTSD.

Candidate Set of Measures of Quality of Care for Major Depressive Disorder

Table F.1
Candidate Set of Measures of Quality of Care for Major Depressive Disorder

Number	Measurea	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-S1 ^b	Percentage of patients screened for PTSD, MDD, and alcohol misuse and, if positive, appropriate follow-up initiated	(a) Number of patients screened annually for PTSD, MDD, or alcohol misuse using a standardized tool (b) If positive screen, follow-up initiated within seven days	(a) All patients (b) Patients who screened positive for PTSD, MDD, or alcohol misuse	Screening	Process	Outpatient
MDD-A1	Percentage of MDD patients in a new treatment episode with assessment of symptoms with the PHQ-9	Patients who have an assessment of MDD symptoms within the first 30 days of a new treatment episode using the PHQ-9	MDD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential
MDD-A2	Percentage of MDD patients in a new treatment episode assessed for manic or hypomanic behaviors	Patients who are assessed for mania or hypomania prior to or concurrent with the visit initiating MDD treatment	MDD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential
MDD-A3 ^b	Percentage of MDD patients in a new treatment episode assessed for suicide risk	Patients assessed for suicide risk during the same visit in which a new treatment episode was identified or in the 14 days prior	MDD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential
MDD-A4 ^b	Percentage of MDD patients in a new treatment episode assessed for recent substance use	Patients who have an assessment of recent substance abuse, including type, quantity, and frequency, within the first 30 days of the new treatment episode	MDD patients in a new treatment episode	Assessment	Process	Outpatient, inpatient, and residential

Table F.1—Continued

Number	Measurea	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-T1 ^b	Percentage of MDD patients with symptom assessment with the PHQ-9 during the four-month measurement period	Patients who had the PHQ-9 administered at least once during the four-month measurement period	MDD patients with an encounter within each four-month period	Treatment	Process	Outpatient
MDD-T2 ^b	Percentage of days when third available specialty-care appointment is within two days	Days in the measurement period on which third available specialty-care visit is within two days	Days in the measurement period	Treatment	Structure	Outpatient
MDD-T3 ^b	Percentage of MDD patient contacts with SI with appropriate follow-up	Documentation of appropriate follow-up for SI, intent, or behavior	Outpatient visits or contacts in which the MDD patient endorsed SI, intent, or behavior	Treatment	Process	Outpatient
MDD-T4 ^b	Percentage of MDD patients with a documented treatment plan	Patients with a treatment plan in the medical record	Patients with MDD	Treatment	Process	Outpatient, inpatient, and residential
MDD-T5	Percentage of MDD patients newly treated with an antidepressant for 12 weeks or six months	Patients who receive antidepressant treatment for (a) 12 weeks (b) six months (c) 12 weeks or have a documented reason for discontinuing antidepressant treatment in less than 12 weeks (d) six months or have a documented reason for discontinuing antidepressant treatment in less than six months	Patients with MDD with a new prescription for an antidepressant	Treatment	Process	Outpatient

Table F.1—Continued

Number	Measurea	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-T6	Percentage of MDD patients newly prescribed an antidepressant with a follow-up visit within 30 days	Patients who have a follow-up visit within 30 days of the new prescription for an antidepressant	Patients with MDD with a new prescription for an antidepressant	Treatment	Process	Outpatient
MDD-T7	Percentage of MDD patients who receive evidence-based psychotherapy	(a) Patients who received any evidence-based psychotherapy visits (b) Number of evidence-based psychotherapy visits received	(a) All MDD patients (b) Patients with MDD diagnosis who are receiving psychotherapy	Treatment	Process	Outpatient
MDD-T8 ^b	Percentage of MDD patients in a new treatment episode who received any psychotherapy	Patients receiving any psychotherapy within four months after starting a new treatment episode	MDD patients in a new treatment episode	Treatment	Process	Outpatient, inpatient, and residential
MDD-T9 ^b	Percentage of MDD patients who receive four psychotherapy visits or two medication-management visits within the first eight weeks	Patients who had four psychotherapy visits or two medication-management visits within eight weeks of the new treatment episode	MDD patients in a new treatment episode	Treatment	Process	Outpatient
MDD-T10	Percentage of MDD patients with response to treatment at six months	Patients who have six-month (plus or minus 30 days) PHQ-9 scores that are reduced by at least 50 percent from their initial PHQ-9 scores	MDD or dysthymia patients with PHQ-9 scores that are positive for depression (PHQ-9 score greater than 9)	Treatment	Outcome	Outpatient

Table F.1—Continued

Number	Measurea	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-T11	Percentage of MDD patients with response to treatment at 12 months	Patients who have 12-month (plus or minus 30 days) PHQ-9 scores that are reduced by at least 50 percent from their initial PHQ-9 scores	MDD or dysthymia patients with PHQ-9 scores that are positive for depression (PHQ-9 score greater than 9)	Treatment	Outcome	Outpatient
MDD-T12	Percentage of MDD patients in MDD-symptom remission at six months	Patients who achieve remission at six months (plus or minus 30 days) as demonstrated by PHQ-9 scores of less than 5	MDD or dysthymia patients with PHQ-9 scores that are positive for depression (PHQ-9 score greater than 9)	Treatment	Outcome	Outpatient
MDD-T13	Percentage of MDD patients in MDD-symptom remission at 12 months	Patients who achieve remission at six months (plus or minus 30 days) as demonstrated by a PHQ-9 score of less than 5	MDD or dysthymia patients with a PHQ-9 score that is positive for depression (PHQ-9 score greater than 9)	Treatment	Outcome	Outpatient
MDD-T14 ^b	Percentage of MDD patients in a new treatment episode with improvement in functional status at six months	Patients with improvement in functional status from the first visit for MDD to six months after the first visit	MDD patients in a new treatment episode who have at least two measures of functional status within the first six months	Treatment	Outcome	Outpatient
MDD-T15 ^b	Percentage of psychiatric inpatient hospital discharges of patients with MDD with follow-up	Inpatient psychiatric discharges followed with a visit within (a) 30 days (b) seven days	MDD patients discharged from an acute inpatient setting with a primary mental health diagnosis	Treatment	Process	Outpatient

Table F.1—Continued

Number	Measurea	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-T16 ^b	Number of inpatient psychiatric beds available per active-duty service members	Number of inpatient psychiatric beds available at the MTFs across the measurement period	Among MTFs with psychiatric inpatient capacity, number of active-duty service members served by the MTF, divided by 10,000	Treatment	Structure	Inpatient
MDD-PE1 ^b	Percentage of MDD patients who report getting treatment quickly	Number of patients who report always getting treatment quickly getting treatment as soon as it was needed (a) getting help by phone (b) getting urgent treatment as soon as it was wanted (c) getting an appointment as soon as it was wanted	MDD patients who received behavioral health care services and who answered "getting treatment quickly" questions on the ECHO survey	All phases	Patient experience	Outpatient
MDD-PE2 ^b	Percentage of MDD patients who report being given as much self-management information as they wanted	Number of patients who report that they received enough information to manage their own conditions	MDD patients who received behavioral health care services and who answered "getting enough information to manage the condition" question on the ECHO survey	All phases	Patient experience	Outpatient

Table F.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-PE3 ^b	Percentage of MDD patients who report being told about treatment options	Number of patients who indicate that they received information about treatment options (a) self-help or consumer-run programs (b) different treatments available for the condition	MDD patients who received behavioral health care services and answered the “information about treatment options” questions on the ECHO survey	All phases	Patient experience	Outpatient
MDD-PE4 ^b	Percentage of MDD patients who report being helped by counseling or treatment received	Number of patients who report being helped a lot by the counseling or treatment received	MDD patients who received behavioral health care services and answered the “amount the treatment helped” question on the ECHO survey	All phases	Patient experience	Outpatient
MDD-PE5 ^b	Percentage of MDD patients who report being better than they were one year ago	Number of patients who report being much better than they were a year ago with (a) ability to deal with daily problems (b) ability to deal with social situations (c) ability to accomplish things (d) ability to deal with symptoms or problems	MDD patients who received behavioral health care services and answered the “perceived improvement” questions on the ECHO survey	All phases	Patient experience	Outpatient

Table F.1—Continued

Number	Measure ^a	Numerator	Denominator	Care Continuum	Measure Type	Care Setting
MDD-PE6 ^b	Percentage of MDD patients who rated counseling and treatment received as 9 or 10 (out of 10)	Number of patients whose overall rating of the counseling or treatment they received as a 9 or 10 (on a 0-to-10 scale)	MDD patients who received behavioral health care services and answered the “overall rating of counseling and treatment” question on the ECHO survey	All phases	Patient experience	Outpatient
MDD-RU1 ^b	Number of psychiatric inpatient admissions per 100 patients with MDD	Number of psychiatric admissions during the measurement period	Number of patients with an MDD diagnosis, divided by 100	All phases	Resource use	Outpatient
MDD-RU2 ^b	Total costs of health care resources used by TRICARE beneficiaries (members) with MDD, divided by number of months enrolled (member-months)	Total costs required for treating beneficiaries with MDD for all of their member-months during the measurement period	(a) Total number of member-months attributable to beneficiaries with MDD (b) Total number of member-months attributable to all beneficiaries	All phases	Resource use	Outpatient, inpatient, and residential

^a Definitions of terms used to characterize quality measures are provided in Appendix A.

^b This measure could be applied to patients with PH conditions other than MDD.

Sources of Measures

This appendix lists the sources we used for information about the measures evaluated for this research.

Agency for Healthcare Research and Quality, “National Quality Measures Clearinghouse,” undated; referenced 2013. As of September 10, 2013:
<http://www.qualitymeasures.ahrq.gov/>

Alberta Mental Health Board, “Performance Monitoring Framework for Alberta’s Mental Health System 2007,” January 17, 2008. As of May 19, 2014:
<http://www.albertahealthservices.ca/MentalHealthWellness/hi-mhw-performance-monitoring-framework-01-08.pdf>

American Medical Association, “PCPI® and PCPI Approved Quality Measures,” undated. As of May 19, 2014:
<http://www.ama-assn.org/apps/listserv/x-check/qmeasure.cgi?submit=PCPI>

Australian Council on Healthcare Standards, *Australasian Clinical Indicator Report 2004–2011*, 13th ed., September 2012. As of May 19, 2014:
http://www.achs.org.au/media/40455/achs_clinical_indicators_report_web.pdf

Chinman, Matthew, Alexander S. Young, Melissa Rowe, Sandy Forquer, Edward Knight, and Anita Miller, “An Instrument to Assess Competencies of Providers Treating Severe Mental Illness,” *Mental Health Services Research*, Vol. 5, No. 2, June 2003, pp. 97–108.

Danish National Indicator Project, “Depression (Adults): Indicators and Standards in the Danish National Indicator Project,” February 24, 2011. As of May 19, 2014:
http://www.nip.dk/files/Subsites/NIP/Om%20NIP/About%20NIP/DNIP_depression_indicatorform_20110302.pdf

Department of Defense Task Force on the Care, Management, and Transition of Recovering Wounded, Ill, and Injured Members of the Armed Forces, *Department of Defense Recovering Warrior Task Force 2011–2012 Annual Report*, August 31, 2012. As of May 7, 2014:
<http://rwtf.defense.gov/Portals/22/Documents/Reports/fy2012report.pdf>

Dumont, Jeanne M., Priscilla A. Ridgway, Steven J. Onken, Douglas H. Dornan, and Ruth O. Ralph, *Mental Health Recovery: What Helps and What Hinders? A National Research Project for the Development of Recovery Facilitating System Performance Indicators—Phase II Technical Report: Development of the Recovery Oriented System Indicators (ROSI) Measures to Advance Mental Health System Transformation*, March 2006. As of May 19, 2014:
http://www.nasmhpd.org/docs/publications/docs/2006/Phase_II_Mental_Health_Recovery.pdf

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND—University of Pittsburgh Health Institute, 2010.

Health Services Advisory Group, *National Impact Assessment of Medicare Quality Measures*, Centers for Medicare and Medicaid Services, March 2012. As of May 19, 2014:
<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/Downloads/NationalImpactAssessmentofQualityMeasuresFINAL.PDF>

Hermann, Richard, Soeren Mattke, and members of the Organisation for Economic Co-Operation and Development Mental Health Care Panel, *Selecting Indicators for the Quality of Mental Health Care at the Health Systems Level in OECD Countries*, DELSA/ELSA/WD/HTP(2004)17, October 28, 2004. As of May 19, 2014:

<http://www.oecd.org/els/healthpoliciesanddata/33865630.pdf>

Hospital Report Research Collaborative, *Hospital Report 2007: Acute Care*, 2007. As of May 19, 2014:
https://secure.cihi.ca/free_products/OHA_Acute07_EN_final_secure.pdf

———, *Hospital Report 2007: Complex Continuing Care*, 2007. As of May 19, 2014:
https://secure.cihi.ca/free_products/OHA_CCC_2007_ENG_final_secure.pdf

———, *Hospital Report 2007: Emergency Department Care*, 2007. As of May 19, 2014:
https://secure.cihi.ca/free_products/OHA_ED_07_EN_final_secure.pdf

———, *Hospital Report 2007: Rehabilitation*, 2007. As of May 19, 2014:
https://secure.cihi.ca/free_products/OHA_Rehab_07_EN_final_secure.pdf

Institute of Medicine, *Treatment for Posttraumatic Stress Disorder in Military and Vetern [sic] Populations: Initial Assessment*, Washington, D.C.: National Academies Press, 2012. As of August 23, 2013:
http://www.nap.edu/openbook.php?record_id=13364

Jeffery, Bonnie, Sylvia Abonyi, Colleen Hamilton, S. Bird, M. Denechezhe, T. Lidguerre, F. Michayluk, L. Thomas, E. Throassie, and Z. Whitecap, *Community Health Indicators Toolkit*, University of Regina and University of Saskatchewan: Saskatchewan Population Health and Evaluation Research Unit, 2006. As of May 19, 2014:
<http://www.uregina.ca/fnh/Combined%20Domains%20-%20Jun-07.pdf>

Joint Commission, *Specifications Manual for Joint Commission National Quality Measures*, version 2013a1, 2013. As of May 19, 2014:
https://manual.jointcommission.org/releases/TJC2013A/rsrc/Manual/TableOfContentsTJC/HBIPS_v2013A1.pdf

Lutterman, Ted, Vijay Ganju, Lucille Schacht, Robert Shaw, Kathleen Monihan, and Marie Huddle, *Sixteen State Study on Mental Health Performance Measures*, Rockville, Md.: Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, 2003.

Management of MDD Working Group, *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)*, Version 2.0, May 2009. As of September 25, 2013:
http://www.healthquality.va.gov/mdd/MDD_FULL_3c1.pdf

Management of Post-Traumatic Stress Working Group, *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress*, Version 2.0, 2010. As of May 7, 2014:
<http://www.healthquality.va.gov/PTSD-Full-2010c.pdf>

Management of Substance Use Disorders Working Group, *Clinical Practice Guideline: Management of Substance Use Disorders*, Version 2.0, August 2009. As of May 7, 2014:
http://www.dcoe.mil/content/navigation/documents/sud_full_601f.pdf

Mitchell, J., M. Trangle, B. Degnan, T. Gabert, B. Haight, D. Kessler, N. Mack, E. Mallen, H. Novak, D. Rossmiller, L. Setterlund, K. Somers, N. Valentino, and S. Vincent, *Adult Depression in Primary Care*, updated September 2013. As of May 19, 2014:
https://www.icsi.org/_asset/fnhdm3/Depr-Interactive0512b.pdf

National Quality Forum, “Quality Positioning System,” undated. As of May 19, 2014:
<http://www.qualityforum.org/QPS/QPSTool.aspx>

———, *MAP Pre-Rulemaking Report: 2013 Recommendations on Measures Under Consideration by HHS, Final Report*, February 2013a. As of May 19, 2014:
http://www.qualityforum.org/Publications/2013/02/MAP_Pre-Rulemaking_Report_-_February_2013.aspx

New Zealand Ministry of Health, *Key Performance Indicator Framework for New Zealand Mental Health and Addiction Services, Phase III: Implementation of the Framework—Benchmarking Participation Manual Part 1: General Information for Participating Organisations*, July 2010. As of May 19, 2014:
<http://www.ndsa.co.nz/LinkClick.aspx?fileticket=u9p3cyK1I1U%3D&tabid=95>

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Province of New Brunswick, Department of Health and Wellness, "Report to New Brunswickers on Comparable Health and Health System Indicators, 2nd Edition," 2004.

Scottish Government, *Mental Health Project Final Report: National Benchmarking Project Report 2*, November 2007. As of May 19, 2014: <http://www.scotland.gov.uk/Publications/2008/01/22113703/0>

Sorbero, Melony E., Thomas E. Mannle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND—University of Pittsburgh Health Institute, 2010.

Spaeth-Rublee, Brigitta, Harold Alan Pincus, and Phuong Trang Huynh, "Measuring Quality of Mental Health Care: A Review of Initiatives and Programs in Selected Countries," *Canadian Journal of Psychiatry*, Vol. 55, No. 9, September 2010, pp. 539–548.

Thorp, Steven R., Janel Fidler, Lucy Moreno, Elizabeth Floto, and Zia Agha, "Lessons Learned from Studies of Psychotherapy for Posttraumatic Stress Disorder via Video Teleconferencing," *Psychological Services*, Vol. 9, No. 2, May 2012, pp. 197–199.

UK Health and Social Care Information Centre, "Indicator Portal," undated. As of May 19, 2014: <http://www.hscic.gov.uk/indicatorportal>

U.S. Department of Defense and U.S. Department of Veterans Affairs, "Integrated Mental Health Strategy, Environmental Scan of DoD, VA, and Civilian Metrics, Strategic Actions #10 and #12, Steps 10.2 and 12.2," draft report, 2011.

U.S. Department of Veterans Affairs, *Department of Veterans Affairs (VA) 2011 Performance and Accountability Report*, November 15, 2011. As of May 19, 2014: http://www.amvets.org/pdfs/legislative_pdfs/2012/VA-Performance-and-Accountability-Report-11-15-11.pdf

———, "Volume II Medical Programs & Information Technology Programs, Congressional Submission, FY 2013 Funding and FY 2014 Advance Appropriations Request," 2012.

———, *2012 Performance and Accountability Report: Honoring All Who Have Served*, 2012. As of May 7, 2014: <http://www.va.gov/budget/report/>

———, Office of Inspector General, Office of Audits and Evaluations and Healthcare Inspections, *Veterans Health Administration: Review of Veterans' Access to Mental Health Care*, 12-00900-168, April 23, 2012. As of March 1, 2013: <http://www.va.gov/oig/pubs/VAOIG-12-00900-168.pdf>

Watkins, Katherine E., Harold Alan Pincus, Brad Smith, Susan M. Paddock, Thomas E. Mannle Jr., Abigail Woodroffe, Jake Solomon, Melony E. Sorbero, Carrie M. Farmer, Kimberly A. Hepner, David M. Adamson, Lanna Forrest, and Catherine Call, *Veterans Health Administration Mental Health Program Evaluation: Capstone Report*, Santa Monica, Calif.: RAND Corporation, TR-956-VHA, 2011. As of May 13, 2014: http://www.rand.org/pubs/technical_reports/TR956.html

Bibliography

Agency for Healthcare Research and Quality, “National Quality Measures Clearinghouse,” undated (a); referenced 2013. As of September 10, 2013:
<http://www.qualitymeasures.ahrq.gov/>

———, “Uses of Quality Measures,” undated (b); referenced 2013. As of August 23, 2013:
<http://www.qualitymeasures.ahrq.gov/selecting-and-using/using.aspx>

———, “CAHPS: Surveys and Tools to Advance Patient-Centered Care,” last updated May 5, 2014; referenced 2012. As of September 30, 2013:
<https://cahps.ahrq.gov/>

AHRQ—See Agency for Healthcare Research and Quality.

Albert, S. M., “Change in BRFSS Health Status/HRQOL Measures Predicts Volume of Medical Use: Results from the Harlem Men’s Health Cohort,” *Program of 17th Annual BRFSS Conference*, Memphis, Tenn., March 14–17, 2000.

Allison-Aipa, Timothy S., Corinne Ritter, Page Sikes, and Stephanie Ball, “The Impact of Deployment on the Psychological Health Status, Level of Alcohol Consumption, and Use of Psychological Health Resources of Postdeployed U.S. Army Reserve Soldiers,” *Military Medicine*, Vol. 175, No. 9, September 2010, pp. 630–637.

Altman, Edward G., Donald Hedeker, James L. Peterson, and John M. Davis, “The Altman Self-Rating Mania Scale,” *Biological Psychiatry*, Vol. 42, No. 10, November 15, 1997, pp. 948–955.

Altshuler, Lori L., Robert M. Post, Gabriele S. Leverich, Kirstin Mikalaukas, Ann Rosoff, and Laura Ackerman, “Antidepressant-Induced Mania and Cycle Acceleration: A Controversy Revisited,” *American Journal of Psychiatry*, Vol. 152, No. 8, August 1995, pp. 1130–1138.

American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, and American Osteopathic Association, “Joint Principles of the Patient-Centered Medical Home,” March 7, 2007. As of May 7, 2014:
http://www.acponline.org/running_practice/delivery_and_payment_models/pcmh/demonstrations/jointprinc_05_17.pdf

American Medical Association, “Physician Consortium for Performance Improvement,” undated. As of August 23, 2013:
<http://www.ama-assn.org/ama/pub/physician-resources/physician-consortium-performance-improvement.page>

American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., Washington, D.C., 1994.

———, *Practice Guideline for the Treatment of Patients with Acute Stress Disorder and Posttraumatic Stress Disorder*, Washington, D.C., November 2004.

———, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed., Arlington, Va., 2013.

American Psychological Association Presidential Task Force on Evidence-Based Practice, “Evidence-Based Practice in Psychology,” *American Psychologist*, Vol. 61, No. 4, May–June 2006, pp. 271–285. As of May 7, 2014:
<http://www.apa.org/practice/resources/evidence/evidence-based-statement.pdf>

- Andresen, Elena M., T. K. Catlin, K. W. Wyrwich, and J. Jackson-Thompson, "Retest Reliability of Surveillance Questions on Health Related Quality of Life," *Journal of Epidemiology and Community Health*, Vol. 57, No. 5, May 2003, pp. 339–343.
- Andresen, Elena M., Bradley S. Fouts, James C. Romeis, and Carol A. Brownson, "Performance of Health-Related Quality-of-Life Instruments in a Spinal Cord Injured Population," *Archives of Physical Medicine and Rehabilitation*, Vol. 80, No. 8, August 1999, pp. 877–884.
- Assessment and Management of Risk for Suicide Working Group, *VA/DoD Clinical Practice Guideline for Assessment and Management of Patients at Risk for Suicide*, Version 1.0, June 2013. As of May 7, 2014: http://www.healthquality.va.gov/guidelines/MH/srb/VADODCP_SuicideRisk_Full.pdf
- Asukai, Nozomu, Azusa Saito, Nobuko Tsuruta, Junji Kishimoto, and Toru Nishikawa, "Efficacy of Exposure Therapy for Japanese Patients with Posttraumatic Stress Disorder Due to Mixed Traumatic Events: A Randomized Controlled Study," *Journal of Traumatic Stress*, Vol. 23, No. 6, December 2010, pp. 744–750.
- Babor, Thomas F., and M. Grant, "From Clinical Research to Secondary Prevention: International Collaboration in the Development of the Alcohol Use Disorders Identification Test (Audit)," *Alcohol Health and Research World*, Vol. 13, No. 3, 1989.
- Bailey, Sue, Assistant Secretary of Defense for Health Affairs, "Policy for Pre- and Post-Deployment Health Assessments and Blood Samples," memorandum for Surgeon General of the Army, Surgeon General of the Navy, Surgeon General of the Air Force, and Deputy Director for Medical Readiness, J-4, the Joint Staff, Washington, D.C., October 6, 1998. As of May 7, 2014: <http://www.health.mil/-/media/MHS/Policy%20Files/Import/99-002.ashx>
- Barney, Lisa J., Kathleen M. Griffiths, Helen Christensen, and Anthony F. Jorm, "The Self-Stigma of Depression Scale (SSDS): Development and Psychometric Evaluation of a New Instrument," *International Journal of Methods in Psychiatric Research*, Vol. 19, No. 4, December 2010, pp. 243–254.
- Bech, P., and O. J. Rafaelsen, "The Use of Rating Scales Exemplified by a Comparison of the Hamilton and the Bech-Rafaelsen Melancholia Scale," *Acta Psychiatrica Scandinavica*, Vol. 62, No. S285, September 1980, pp. 128–132.
- Beck, Aaron T., "Psychometric Properties of the Beck Depression Inventory: Twenty-Five Years of Evaluation," *Clinical Psychology Review*, Vol. 8, No. 1, 1988, pp. 77–100.
- Beck, Aaron T., Gary Emery, and Ruth L. Greenberg, *Anxiety Disorders and Phobias: A Cognitive Perspective*, 15th anniversary ed., Cambridge, Mass.: Basic Books, 2005.
- Beck, Aaron T., Robert A. Steer, and Gregory K. Brown, *BDI-II, Beck Depression Inventory: Manual*, San Antonio, Texas: Psychological Corporation, 1996.
- Beck, Aaron T., Calvin H. Ward, M. Mendelson, Jeremiah Mock, and J. Erbaugh, "An Inventory for Measuring Depression," *Archives of General Psychiatry*, Vol. 4, No. 6, 1961, pp. 561–571.
- Benedek, David M., Matthew J. Friedman, Douglas Zatzick, and Robert J. Ursano, *Guideline Watch (March 2009): Practice Guideline for the Treatment of Patients with Acute Stress Disorder and Post Traumatic Stress Disorder*, Washington, D.C.: American Psychiatric Association, March 2009.
- Berwick, Donald M., "A Primer on Leading the Improvement of Systems," *BMJ*, Vol. 312, No. 7031, 1996, p. 619.
- Bisson, J., and M. Andrew, "Psychological Treatment of Post-Traumatic Stress Disorder (PTSD)," *Cochrane Database of Systematic Reviews*, No. 3, July 18, 2007, p. CD003388.
- Blanchard, Edward B., Edward J. Hickling, Trishul Devineni, Connie H. Veazey, Tara E. Galovski, Elizabeth Mundy, Loretta S. Malta, and Todd C. Buckley, "A Controlled Evaluation of Cognitive Behavioural Therapy for Posttraumatic Stress in Motor Vehicle Accident Survivors," *Behaviour Research and Therapy*, Vol. 41, No. 1, January 2003, pp. 79–96.
- Blanchard, Edward B., Jacqueline Jones-Alexander, Todd C. Buckley, and Catherine A. Forneris, "Psychometric Properties of the PTSD Checklist (PCL)," *Behaviour Research and Therapy*, Vol. 34, No. 8, August 1996, pp. 669–673.

- Bliese, Paul D., Kathleen M. Wright, Amy B. Adler, Oscar Cabrera, Carl A. Castro, and Charles W. Hoge, "Validating the Primary Care Posttraumatic Stress Disorder Screen and the Posttraumatic Stress Disorder Checklist with Soldiers Returning from Combat," *Journal of Consulting and Clinical Psychology*, Vol. 76, No. 2, April 2008, pp. 272–281.
- Blue Ribbon Work Group on Suicide Prevention in the Veteran Population, *Report of the Blue Ribbon Work Group on Suicide Prevention in the Veteran Population: Executive Summary*, c. 2008. As of September 18, 2013: http://www.mentalhealth.va.gov/suicide_prevention/Blue_Ribbon_Report-FINAL_June-30-08.pdf
- Bongar, Bruce Michael, *The Suicidal Patient: Clinical and Legal Standards of Care*, 2nd ed., Washington, D.C.: American Psychological Association, 2002.
- Boudewyns, Patrick A., and Lee Hyer, "Physiological Response to Combat Memories and Preliminary Treatment Outcome in Vietnam Veteran PTSD Patients Treated with Direct Therapeutic Exposure," *Behavior Therapy*, Vol. 21, No. 1, Winter 1990, pp. 63–87.
- Bower, Peter, Simon Gilbody, David Richards, Janine Fletcher, and Alex Sutton, "Collaborative Care for Depression in Primary Care: Making Sense of a Complex Intervention—Systematic Review and Meta-Regression," *British Journal of Psychiatry*, Vol. 189, 2006, pp. 484–493.
- Boyer, Carol A., Donna D. McAlpine, Kathleen J. Pottick, and Mark Olfson, "Identifying Risk Factors and Key Strategies in Linkage to Outpatient Psychiatric Care," *American Journal of Psychiatry*, Vol. 157, No. 10, October 1, 2000, pp. 1592–1598.
- Bradley, Katharine, and Douglas Berger, "Screening for Unhealthy Alcohol Use," in Richard Saitz, ed., *Addressing Unhealthy Alcohol Use in Primary Care*, New York: Springer New York, 2013, pp. 7–27.
- Bradley, Katharine A., Anna F. DeBenedetti, Robert J. Volk, Emily C. Williams, Danielle Frank, and Daniel R. Kivlahan, "AUDIT-C as a Brief Screen for Alcohol Misuse in Primary Care," *Alcoholism: Clinical and Experimental Research*, Vol. 31, No. 7, July 2007, pp. 1208–1217.
- Brady, Kathleen, Teri Pearlstein, Gregory M. Asnis, Dewleen Baker, Barbara Rothbaum, Carolyn R. Sikes, and Gail M. Farfel, "Efficacy and Safety of Sertraline Treatment of Posttraumatic Stress Disorder: A Randomized Controlled Trial," *JAMA*, Vol. 283, No. 14, 2000, pp. 1837–1844.
- Breslau, Naomi, Edward L. Peterson, Ronald C. Kessler, and Lonni R. Schultz, "Short Screening Scale for DSM-IV Posttraumatic Stress Disorder," *American Journal of Psychiatry*, Vol. 156, No. 6, June 1999, pp. 908–911.
- Brewin, Chris R., Suzanna Rose, Bernice Andrews, John Green, Philip Tata, Chris McEvedy, Stuart Turner, and Edna B. Foa, "Brief Screening Instrument for Post-Traumatic Stress Disorder," *British Journal of Psychiatry*, Vol. 181, No. 2, 2002, pp. 158–162.
- Briere, John, "Psychometric Review of the Trauma Symptom Checklist–40," in B. Hudnall Stamm, ed., *Measurement of Stress, Trauma, and Adaptation*, Lutherville, Md.: Sidran Press, 1996, pp. 373–376.
- Briere, John, Diana M. Elliott, Kathryn Harris, and Ann Cotman, "Trauma Symptom Inventory: Psychometrics and Association with Childhood and Adult Victimization in Clinical Samples," *Journal of Interpersonal Violence*, Vol. 10, No. 4, December 1995, pp. 387–401.
- Brown, Gregory K., Thomas Ten Have, Gregg R. Henriques, Sharon X. Xie, Judd E. Hollander, and Aaron T. Beck, "Cognitive Therapy for the Prevention of Suicide Attempts: A Randomized Controlled Trial," *JAMA*, Vol. 294, No. 5, 2005, pp. 563–570.
- Bryant, Richard A., Michelle L. Moulds, and Reginald V. D. Nixon, "Cognitive Behaviour Therapy of Acute Stress Disorder: A Four-Year Follow-Up," *Behaviour Research and Therapy*, Vol. 41, No. 4, April 2003, pp. 489–494.
- Burnam, M. Audrey, Kenneth B. Wells, Barbara Leake, and John Landsverk, "Development of a Brief Screening Instrument for Detecting Depressive Disorders," *Medical Care*, Vol. 26, No. 8, September 1988, pp. 775–789.
- Cahill, Shawn P., Barbara Olasov Rothbaum, Patrick A. Resick, and Victoria M. Follette, "Cognitive-Behavioral Therapy for Adults," in Edna B. Foa, Terence M. Keane, Matthew J. Friedman, and Judith A. Cohen, eds., *Effective Treatments for PTSD: Practice Guidelines from the International Society for Traumatic Stress Studies*, 2nd ed., New York: Guilford Press, 2009.

Carlson, Eve B., “Psychometric Study of a Brief Screen for PTSD: Assessing the Impact of Multiple Traumatic Events,” *Assessment*, Vol. 8, No. 4, January 2002, pp. 431–441.

Cavanagh, Jonathan T. O., Alan J. Carson, Michael Sharpe, and S. M. Lawrie, “Psychological Autopsy Studies of Suicide: A Systematic Review,” *Psychological Medicine*, Vol. 33, No. 3, 2003, pp. 395–405. As of May 7, 2014:

http://journals.cambridge.org/download.php?file=%2FPSM%2FPSM33_03%2FS0033291702006943a.pdf&code=fbbe6d258d94a91381bc6fe8a66d382c

CDC—See Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Adult and Community Health, *Measuring Healthy Days: Population Assessment of Health-Related Quality of Life*, Atlanta, Ga., November 2000. As of September 19, 2013:

<http://www.cdc.gov/hrqol/pdfs/mhd.pdf>

———, “Surveillance and Data,” last updated February 22, 2012; referenced 2013. As of September 19, 2013: <http://www.cdc.gov/hrqol/surveillance.htm>

———, “National Health and Nutrition Examination Survey,” last updated April 29, 2014; referenced 2013. As of October 2, 2013:

<http://www.cdc.gov/nchs/nhanes.htm>

Centers for Medicare and Medicaid Services, “Payments to Medicare Advantage Organizations (Rev. 89, 11-02-07),” *Medicare Managed Care Manual*, November 2, 2007.

———, “Hospital Compare,” last modified December 23, 2013; referenced August 26, 2013. As of May 7, 2014:

<http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare.html>

———, *A Blueprint for the CMS Measures Management System*, version 11, July 2014. As of December 20, 2014:

<http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/MeasuresManagementSystemBlueprint.html>

Centers for Medicare and Medicaid Services and Joint Commission, “Glossary of Terms,” *Specifications Manual, Version 4.2b: Discharges 01/01/2013 to 12/31/2013*, December 4, 2012. As of May 7, 2014:

http://www.jointcommission.org/specifications_manual_for_national_hospital_inpatient_quality_measures.aspx

Chung, Jeanette, Erin Kaleba, and Gregory Wozniak, “A Framework for Measuring Healthcare Efficiency and Value (Clinical Performance Evaluation),” August 2008.

Cleary, Paul D., and Margaret E. O’Kane, “Structure Measures,” *e-Source Book*, Bethesda, Md.: National Institutes of Health, Office of Behavioral and Social Sciences Research, undated. As of October 2, 2013:

<http://www.esourceresearch.org/eSourceBook/EvaluatingtheQualityofHealthCare/7StructureMeasures/tabid/812/Default.aspx>

CMS—See Centers for Medicare and Medicaid Services.

CMS and Joint Commission—See Centers for Medicare and Medicaid Services and Joint Commission.

Committee to Design a Strategy for Quality Review and Assurance in Medicare, Division of Health Care Services, Institute of Medicine, *Medicare: A Strategy for Quality Assurance*, Vol. I, Kathleen N. Lohr, ed., Washington, D.C., 1990.

Connor, K. M., and J. R. T. Davidson, “SPRINT: A Brief Global Assessment of Post-Traumatic Stress Disorder,” *International Clinical Psychopharmacology*, Vol. 16, No. 5, September 2001, pp. 279–284.

Cottraux, Jean, Ivan Note, Sai Nan Yao, Chantal de Mey-Guillard, Françoise Bonasse, Diane Djamoussian, Evelyne Mollard, Brigitte Note, and Yaohua Chen, “Randomized Controlled Comparison of Cognitive Behavior Therapy with Rogerian Supportive Therapy in Chronic Post-Traumatic Stress Disorder: A 2-Year Follow-Up,” *Psychotherapy and Psychosomatics*, Vol. 77, No. 2, 2008, pp. 101–110.

- Currey, Shannon S., Jaya K. Rao, John B. Winfield, and Leigh F. Callahan, "Performance of a Generic Health-Related Quality of Life Measure in a Clinic Population with Rheumatic Disease," *Arthritis and Rheumatism: Arthritis Care and Research*, Vol. 49, No. 5, October 15, 2003, pp. 658–664.
- Damschroder, Laura J., David C. Aron, Rosalind E. Keith, Susan R. Kirsh, Jeffery A. Alexander, and Julie C. Lowery, "Fostering Implementation of Health Services Research Findings into Practice: A Consolidated Framework for Advancing Implementation Science," *Implementation Science*, Vol. 4, No. 1, August 7, 2009. As of January 30, 2014:
<http://www.implementationscience.com/content/pdf/1748-5908-4-50.pdf>
- Davidson, Jonathan, *SPAN Addendum to DTS Manual*, New York: Multi-Health System Inc., 2002.
- Davidson, Jonathan, David Baldwin, Dan J. Stein, Enrique Kuper, Isma Benattia, Saeed Ahmed, Ron Pedersen, and Jeff Musgnung, "Treatment of Posttraumatic Stress Disorder with Venlafaxine Extended Release: A 6-Month Randomized Controlled Trial," *Archives of General Psychiatry*, Vol. 63, No. 10, 2006, pp. 1158–1165.
- Davidson, Jonathan, Barbara O. Rothbaum, Phebe Tucker, Gregory Asnis, Isma Benattia, and Jeff J. Musgnung, "Venlafaxine Extended Release in Posttraumatic Stress Disorder: A Sertraline- and Placebo-Controlled Study," *Journal of Clinical Psychopharmacology*, Vol. 26, No. 3, June 2006, pp. 259–267.
- Davidson, Jonathan R. T., Barbara O. Rothbaum, Bessel A. van der Kolk, Carolyn R. Sikes, and Gail M. Farfel, "Multicenter, Double-Blind Comparison of Sertraline and Placebo in the Treatment of Posttraumatic Stress Disorder," *Archives of General Psychiatry*, Vol. 58, No. 5, 2001, pp. 485–492.
- Davidson, P. R., and K. C. H. Parker, "Eye Movement Desensitization and Reprocessing (EMDR): A Meta-Analysis," *Journal of Consulting and Clinical Psychology*, Vol. 69, No. 2, April 2001, pp. 305–316.
- DCoE—See Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury.
- De Mello, Marcelo Feijo, Jair de Jesus Mari, Josue Bacaltchuk, Helen Verdeli, and Richard Neugebauer, "A Systematic Review of Research Findings on the Efficacy of Interpersonal Therapy for Depressive Disorders," *European Archives of Psychiatry and Clinical Neuroscience*, Vol. 255, No. 2, 2005, pp. 75–82.
- Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury, *2011 Annual Report*, 2011. As of September 10, 2013:
http://www.dcoe.health.mil/Content/Navigation/Documents/DCoE_2011_Annual_Report.pdf
- Dennehy, Ellen B., Trisha Suppes, M. Lynn Crismon, Marcia Toprac, Thomas J. Carmody, and A. John Rush, "Development of the Brief Bipolar Disorder Symptom Scale for Patients with Bipolar Disorder," *Psychiatry Research*, Vol. 127, No. 1–2, June 30, 2004, pp. 137–145.
- Department of Defense Task Force on the Care, Management, and Transition of Recovering Wounded, Ill, and Injured Members of the Armed Forces, *Department of Defense Recovering Warrior Task Force 2011–2012 Annual Report*, August 31, 2012. As of May 7, 2014:
<http://rwtf.defense.gov/Portals/22/Documents/Reports/fy2012report.pdf>
- DeRubeis, Robert J., Steven D. Hollon, Jay D. Amsterdam, Richard C. Shelton, Paula R. Young, Ronald M. Salomon, John P. O'Reardon, Margaret L. Lovett, Madeline M. Gladis, Laurel L. Brown, and Robert Gallop, "Cognitive Therapy vs Medications in the Treatment of Moderate to Severe Depression," *Archives of General Psychiatry*, Vol. 62, No. 4, April 2005, pp. 409–416.
- Devins, Gerald M., Carolee M. Orme, Charles G. Costello, Yitzchak M. Binik, Beverly Frizzell, Henderikus J. Stam, and Wendy M. Pullin, "Measuring Depressive Symptoms in Illness Populations: Psychometric Properties of the Center for Epidemiologic Studies Depression (CES-D) Scale," *Psychology and Health*, Vol. 2, No. 2, 1988, pp. 139–156.
- Dimidjian, Sona, Steven D. Hollon, Keith S. Dobson, Karen B. Schmaling, Robert J. Kohlenberg, Michael E. Addis, Robert Gallop, Joseph B. McGlinchey, David K. Markley, Jackie K. Gollan, David C. Atkins, David L. Dunner, and Neil S. Jacobson, "Randomized Trial of Behavioral Activation, Cognitive Therapy, and Antidepressant Medication in the Acute Treatment of Adults with Major Depression," *Journal of Consulting and Clinical Psychology*, Vol. 74, No. 4, August 2006, pp. 658–670.

Dixon, Lisa, Richard Goldberg, Virginia Iannone, Alicia Lucksted, Clayton Brown, Julie Kreyenbuhl, Lijuan Fang, and Wendy Potts, "Use of a Critical Time Intervention to Promote Continuity of Care After Psychiatric Inpatient Hospitalization," *Psychiatric Services*, Vol. 60, No. 4, April 1, 2009, pp. 451–458.

DoD Task Force on the Care, Management, and Transition of Recovering Wounded, Ill, and Injured Members of the Armed Forces—See Department of Defense Task Force on the Care, Management, and Transition of Recovering Wounded, Ill, and Injured Members of the Armed Forces.

Domino, Marisa Elena, Barbara J. Burns, Susan G. Silva, Christopher J. Kratochvil, Benedetto Vitiello, Mark A. Reinecke, Jeremy Mario, and John S. March, "Cost-Effectiveness of Treatments for Adolescent Depression: Results from TADS," *American Journal of Psychiatry*, Vol. 165, No. 5, May 1, 2008, pp. 588–596.

Donabedian, Avedis, "Evaluating the Quality of Medical Care," *Milbank Memorial Fund Quarterly*, Vol. 44, No. 3, Pt. 2, 1966, pp. 166–203.

———, *An Introduction to Quality Assurance in Health Care*, New York: Oxford University Press, 2003.

ECHO Development Team, "Experience of Care and Health Outcomes Survey (ECHO), Version 3.0," 2002.

Eckblad, Mark, and Loren J. Chapman, "Development and Validation of a Scale for Hypomanic Personality," *Journal of Abnormal Psychology*, Vol. 95, No. 3, August 1986, pp. 214–222.

Erickson, Darin J., Jessica Wolfe, Daniel W. King, Lynda A. King, and Erica J. Sharkansky, "Posttraumatic Stress Disorder and Depression Symptomatology in a Sample of Gulf War Veterans: A Prospective Analysis," *Journal of Consulting and Clinical Psychology*, Vol. 69, No. 1, February 2001, pp. 41–49.

Ewing, John A., "Detecting Alcoholism: The CAGE Questionnaire," *JAMA*, Vol. 252, No. 14, October 12, 1984, pp. 1905–1907.

Falsetti, Sherry A., Heidi S. Resnick, Patricia A. Resick, and D. G. Kilpatrick, "The Modified PTSD Symptom Scale: A Brief Self-Report Measure of Posttraumatic Stress Disorder," *Behaviour Therapist*, Vol. 16, 1993, pp. 161–161.

Farmer, Carrie M., Katherine E. Watkins, Brad Smith, Susan M. Paddock, Abigail Woodroffe, Jacob Solomon, Melony E. Sorbero, Kimberly A. Hepner, Lanna Forrest, Lisa R. Shugarman, Cathy Call, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Medical Record Review Report*, Alexandria, Va.: Altarum Institute and RAND—University of Pittsburgh Health Institute, 2010.

Festinger, David S., R. J. Lamb, Maria R. Kountz, Kimberly C. Kirby, and Douglas Marlowe, "Pretreatment Dropout as a Function of Treatment Delay and Client Variables," *Addictive Behaviors*, Vol. 20, No. 1, January–February 1995, pp. 111–115.

First, M. B., R. L. Spitzer, M. Gibbon, and J. B. W. Williams, *Structured Clinical Interview for Axis I DSM-IV Disorders—Patient Edition (SCID-I/P) (W/Psychotic Screen) Version 2.0*, New York: Biometrics Research Department, New York State Psychiatric Institute, 1996.

Fischer, Ellen P., John F. McCarthy, Rosalinda V. Ignacio, Frederic C. Blow, Kristen L. Barry, Teresa J. Hudson, Richard R. Owen Jr., and Marcia Valenstein, "Longitudinal Patterns of Health System Retention Among Veterans with Schizophrenia or Bipolar Disorder," *Community Mental Health Journal*, Vol. 44, No. 5, October 2008, pp. 321–330.

Foa, Edna B., Constance V. Dancu, Elizabeth A. Hembree, Lisa H. Jaycox, Elizabeth A. Meadows, and Gordon P. Street, "A Comparison of Exposure Therapy, Stress Inoculation Training, and Their Combination for Reducing Posttraumatic Stress Disorder in Female Assault Victims," *Journal of Consulting and Clinical Psychology*, Vol. 67, No. 2, April 1999, pp. 194–200.

Foa, Edna B., Jonathan R. T. Davidson, and Allen Frances, eds., "The Expert Consensus Guideline Series: Treatment of Posttraumatic Stress Disorder," *Journal of Clinical Psychiatry*, Vol. 60, Supp. 16, 1999, pp. 6–76.

Foa, Edna B., Elizabeth A. Hembree, Shawn P. Cahill, Sheila A. M. Rauch, David S. Riggs, Norah C. Feeny, and Elna Yadin, "Randomized Trial of Prolonged Exposure for Posttraumatic Stress Disorder with and Without Cognitive Restructuring: Outcome at Academic and Community Clinics," *Journal of Consulting and Clinical Psychology*, Vol. 73, No. 5, October 2005, pp. 953–964.

Foa, Edna B., Elizabeth Ann Hembree, and Barbara Olasov Rothbaum, *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences—Therapist Guide*, New York: Oxford University Press, 2007.

- Foa, Edna B., Terence Martin Keane, and Matthew J. Friedman, *Effective Treatments for PTSD: Practice Guidelines from the International Society for Traumatic Stress Studies*, New York: Guilford, 2000.
- Foa, Edna B., Barbara O. Rothbaum, David S. Riggs, and Tamera B. Murdock, "Treatment of Posttraumatic Stress Disorder in Rape Victims: A Comparison Between Cognitive-Behavioral Procedures and Counseling," *Journal of Consulting and Clinical Psychology*, Vol. 59, No. 5, October 1991, pp. 715–723.
- Fontana, Alan, and Robert Rosenheck, "A Short Form of the Mississippi Scale for Measuring Change in Combat-Related PTSD," *Journal of Traumatic Stress*, Vol. 7, No. 3, 1994, pp. 407–414.
- Fournier, Jay C., Robert J. DeRubeis, Steven D. Hollon, Sona Dimidjian, Jay D. Amsterdam, Richard C. Shelton, and Jan Fawcett, "Antidepressant Drug Effects and Depression Severity: A Patient-Level Meta-Analysis," *JAMA*, Vol. 303, No. 1, January 6, 2010, pp. 47–53.
- Gabrys, John B., and Kay Peters, "Reliability, Discriminant and Predictive Validity of the Zung Self-Rating Depression Scale," *Psychological Reports*, Vol. 57, No. 3, Part 2, December 1985, pp. 1091–1096.
- Gahm, Gregory A., and Barbara A. Lucenko, "Screening Soldiers in Outpatient Care for Mental Health Concerns," *Military Medicine*, Vol. 173, No. 1, January 2008, pp. 17–24.
- Gallucci, Gerard, Wayne Swartz, and Florence Hackerman, "Brief Reports: Impact of the Wait for an Initial Appointment on the Rate of Kept Appointments at a Mental Health Center," *Psychiatric Services*, Vol. 56, No. 3, March 1, 2005, pp. 344–346.
- Gartlehner, Gerald, Richard A. Hansen, Patricia Thieda, Angela M. DeVeugh-Geiss, Bradley N. Gaynes, Erin E. Krebs, Linda J. Lux, Laura C. Morgan, Janelle A. Shumate, Loraine G. Monroe, and Kathleen N. Lohr, *Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression*, Rockville, Md.: Agency for Healthcare Research and Quality, January 2007. As of May 9, 2014: <http://www.ncbi.nlm.nih.gov/books/NBK43023/>
- Gaynes, Bradley N., Joanne DeVeugh-Geiss, Sam Weir, Hongbin Gu, Cora MacPherson, Herbert C. Schulberg, Larry Culpepper, and David R. Rubinow, "Feasibility and Diagnostic Validity of the M-3 Checklist: A Brief, Self-Rated Screen for Depressive, Bipolar, Anxiety, and Post-Traumatic Stress Disorders in Primary Care," *Annals of Family Medicine*, Vol. 8, No. 2, March 1, 2010, pp. 160–169.
- Gerrity, Martha S., Kathryn Corson, and Steven K. Dobscha, "Screening for Posttraumatic Stress Disorder in VA Primary Care Patients with Depression Symptoms," *Journal of General Internal Medicine*, Vol. 22, No. 9, September 2007, pp. 1321–1324.
- Gilbody, Simon, Peter Bower, Janine Fletcher, David Richards, and Alex J. Sutton, "Collaborative Care for Depression: A Cumulative Meta-Analysis and Review of Longer-Term Outcomes," *Archives of Internal Medicine*, Vol. 166, No. 21, November 27, 2006, pp. 2314–2321.
- Gilbody, Simon, David Richards, Stephen Brealey, and Catherine Hewitt, "Screening for Depression in Medical Settings with the Patient Health Questionnaire (PHQ): A Diagnostic Meta-Analysis," *Journal of General Internal Medicine*, Vol. 22, No. 11, November 2007, pp. 1596–1602.
- Glass, Kathryn P., and Jeffery R. Anderson, "Relative Value Units: From A to Z (Part I of IV)," *Journal of Medical Practice Management*, Vol. 17, No. 5, March–April 2002, pp. 225–228.
- Goldsmith, Sara K., Terry C. Pellmar, Arthur M. Kleinman, and William E. Bunney, eds., *Reducing Suicide: A National Imperative*, Washington, D.C.: National Academies Press, 2002. As of May 9, 2014: <http://www.nap.edu/catalog/10398.html>
- Gore, Kristie L., Charles C. Engel, Michael C. Freed, Xian Liu, and David W. Armstrong III, "Test of a Single-Item Posttraumatic Stress Disorder Screener in a Military Primary Care Setting," *General Hospital Psychiatry*, Vol. 30, No. 5, September–October 2008, pp. 391–397.
- Greenberg, Greg A., Robert A. Rosenheck, and Alan Fontana, "Continuity of Care and Clinical Effectiveness: Treatment of Posttraumatic Stress Disorder in the Department of Veterans Affairs," *Journal of Behavioral Health Services and Research*, Vol. 30, No. 2, April–June 2003, pp. 202–214.
- Hamilton, Max, "A Rating Scale for Depression," *Journal of Neurology, Neurosurgery, and Psychiatry*, Vol. 23, No. 1, February 1960, pp. 56–62.

———, “Development of a Rating Scale for Primary Depressive Illness,” *British Journal of Clinical Psychology*, Vol. 6, No. 4, December 1967, pp. 278–296.

Hammarberg, Melvyn, “Penn Inventory for Posttraumatic Stress Disorder: Psychometric Properties,” *Psychological Assessment*, Vol. 4, No. 1, March 1992, pp. 67–76.

Harding, Kelli Jane, A. John Rush, Melissa Arbuckle, Madhukar H. Trivedi, and Harold Alan Pincus, “Measurement-Based Care in Psychiatric Practice: A Policy Framework for Implementation,” *Journal of Clinical Psychiatry*, Vol. 72, No. 8, August 2011, pp. 1136–1143.

Harris, E. Clare, and Brian Barraclough, “Suicide as an Outcome for Mental Disorders: A Meta-Analysis,” *British Journal of Psychiatry*, Vol. 170, No. 3, March 1997, pp. 205–228.

Hermann, R., and R. Palmer, “Common Ground: A Framework for Selecting Core Quality Measures,” *Psychiatric Services*, Vol. 53, No. 3, 2002, pp. 281–287.

Hilton, Susan M., David B. Service, Valerie A. Stander, Aaron D. Werbel, and Bonnie R. Chavez, *Department of the Navy Suicide Incident Report (DONSIR): Summary of Findings, 1999–2007*, San Diego, Calif.: Naval Health Research Center, March 24, 2009. As of May 9, 2014:
<http://oai.dtic.mil/oai/oai?&verb=getRecord&metadataPrefix=html&identifier=ADA515110>

Hirschfeld, Robert M. A., Janet B. W. Williams, Robert L. Spitzer, Joseph R. Calabrese, Laurie Flynn, Paul E. Keck Jr., Lydia Lewis, Susan L. McElroy, Robert M. Post, Daniel J. Rapport, James M. Russell, Gary S. Sachs, and John Zajecka, “Development and Validation of a Screening Instrument for Bipolar Spectrum Disorder: The Mood Disorder Questionnaire,” *American Journal of Psychiatry*, Vol. 157, No. 11, November 1, 2000, pp. 1873–1875.

Horvitz-Lennon, Marcela, Katherine E. Watkins, Harold Alan Pincus, Lisa R. Shugarman, Brad Smith, Teryn Mattox, and Thomas E. Mannle Jr., *Veterans Health Administration Mental Health Program Evaluation Technical Manual*, Santa Monica, Calif.: RAND Corporation, WR-682-VHA, 2009. As of May 9, 2014:
http://www.rand.org/pubs/working_papers/WR682.html

Institute for Healthcare Improvement, “Third Next Available Appointment,” undated; referenced 2011. As of September 18, 2013:
<http://www.ihl.org/knowledge/Pages/Measures/ThirdNextAvailableAppointment.aspx>

Institute of Medicine, *Crossing the Quality Chasm: A New Health System for the 21st Century*, Washington, D.C.: National Academy Press, 2001. As of May 9, 2014:
<http://www.nap.edu/books/0309072808/html/>

———, *Treatment of Posttraumatic Stress Disorder: An Assessment of the Evidence*, Washington, D.C.: National Academies Press, 2008. As of May 9, 2014:
http://www.nap.edu/catalog.php?record_id=11955

———, *Treatment for Posttraumatic Stress Disorder in Military and Vetern [sic] Populations: Initial Assessment*, Washington, D.C.: National Academies Press, 2012. As of August 23, 2013:
http://www.nap.edu/openbook.php?record_id=13364

———, *Returning Home from Iraq and Afghanistan: Assessment of Readjustment Needs of Veterans, Service Members, and Their Families*, Washington, D.C.: National Academies Press, 2013. As of May 9, 2014:
http://www.nap.edu/catalog.php?record_id=13499

IOM—See Institute of Medicine.

Ipser, Jonathan C., and Dan J. Stein, “Evidence-Based Pharmacotherapy of Post-Traumatic Stress Disorder (PTSD),” *International Journal of Neuropsychopharmacology*, Vol. 15, No. 6, July 2012, pp. 825–840.

Jia, Haomiao, Matthew M. Zack, David G. Moriarty, and Dennis G. Fryback, “Predicting the EuroQol Group’s EQ-5D Index from CDC’s ‘Healthy Days’ in a US Sample,” *Medical Decision Making*, Vol. 31, No. 1, January–February 2011, pp. 174–185.

- Jonas, Daniel E., Karen Cusack, Catherine A. Forneris, Tania M. Wilkins, Jeffrey Sonis, Jennifer Cook Middleton, Cynthia Feltner, Dane Meredith, Jamie Cavanaugh, Kimberly A. Brownley, Kristine Rae Olmsted, Amy Greenblatt, Amy Weil, and Bradley N. Gaynes, *Psychological and Pharmacological Treatments for Adults with Posttraumatic Stress Disorder (PTSD)*, Comparative Effectiveness Review 92, Rockville, Md.: Agency for Healthcare Research and Quality, April 2013. As of May 9, 2014: <http://www.effectivehealthcare.ahrq.gov/ehc/products/347/1435/PTSD-adult-treatment-report-130403.pdf>
- Karlin, Bradley E., and Gerald Cross, "From the Laboratory to the Therapy Room: National Dissemination and Implementation of Evidence-Based Psychotherapies in the U.S. Department of Veterans Affairs Health Care System," *American Psychologist*, Vol. 69, No. 1, January 2014, pp. 19–33.
- Katon, Wayne, Jürgen Unützer, Ming-Yu Fan, John W. Williams Jr., Michael Schoenbaum, Elizabeth H. B. Lin, and Enid M. Hunkeler, "Cost-Effectiveness and Net Benefit of Enhanced Treatment of Depression for Older Adults with Diabetes and Depression," *Diabetes Care*, Vol. 29, No. 2, February 2006, pp. 265–270.
- Keane, Terence M., Juesta M. Caddell, and Kathryn L. Taylor, "Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: Three Studies in Reliability and Validity," *Journal of Consulting and Clinical Psychology*, Vol. 56, No. 1, February 1988, pp. 85–90.
- Kelly, Thomas M., and J. J. Mann, "Validity of DSM-III-R Diagnosis by Psychological Autopsy: A Comparison with Clinician Ante-Mortem Diagnosis," *Acta Psychiatrica Scandinavica*, Vol. 94, No. 5, November 1996, pp. 337–343.
- Kessler, Ronald C., "The Epidemiology of Dual Diagnosis," *Biological Psychiatry*, Vol. 56, No. 10, November 15, 2004, pp. 730–737.
- Kessler, Ronald C., Guilherme Borges, and Ellen E. Walters, "Prevalence of and Risk Factors for Lifetime Suicide Attempts in the National Comorbidity Survey," *Archives of General Psychiatry*, Vol. 56, No. 7, July 1999, pp. 617–626.
- Kessler, Ronald C., Amanda Sonnega, Evelyn Bromet, Michael Hughes, and Christopher B. Nelson, "Posttraumatic Stress Disorder in the National Comorbidity Survey," *Archives of General Psychiatry*, Vol. 52, No. 12, December 1995, pp. 1048–1060.
- Kilpatrick, Dean G., Lois J. Veronen, and Patricia A. Resick, "Psychological Sequelae to Rape: Assessment and Treatment Strategies," in Daniel M. Doleys, R. L. Meredith, and Anthony R. Ciminero, eds., *Behavioral Medicine: Assessment and Treatment Strategies*, New York: Plenum Press, 1982, pp. 473–497.
- King, Lynda A., Daniel W. King, Gregory Leskin, and David W. Foy, "The Los Angeles Symptom Checklist: A Self Report Measure of Posttraumatic Stress Disorder," *Assessment*, Vol. 2, No. 1, March 1995, pp. 1–17.
- Klerman, Gerald L., Myrna M. Weissman, Bruce J. Rounsaville, and Eve S. Chevron, *Interpersonal Psychotherapy of Depression*, New York: Basic Books, 1984.
- Kongstvedt, Peter R., *Managed Care: What It Is and How It Works*, Sudbury, Mass.: Jones and Bartlett Publishers, 2009.
- Kreyenbuhl, Julie, Ilana R. Nossel, and Lisa B. Dixon, "Disengagement from Mental Health Treatment Among Individuals with Schizophrenia and Strategies for Facilitating Connections to Care: A Review of the Literature," *Schizophrenia Bulletin*, Vol. 35, No. 4, July 2009, pp. 696–703.
- Kroenke, Kurt, Robert L. Spitzer, and Janet B. W. Williams, "The PHQ-9: Validity of a Brief Depression Severity Measure," *Journal of General Internal Medicine*, Vol. 16, No. 9, September 2001, pp. 606–613.
- , "The Patient Health Questionnaire–2: Validity of a Two-Item Depression Screener," *Medical Care*, Vol. 41, No. 11, November 2003, pp. 1284–1292.
- Kubany, Edward S., Mary Beth Leisen, Aaron S. Kaplan, and Martin P. Kelly, "Validation of a Brief Measure of Posttraumatic Stress Disorder: The Distressing Event Questionnaire (DEQ)," *Psychological Assessment*, Vol. 12, No. 2, June 2000, pp. 197–209.
- Lagomasino, Isabel, Robert Daly, and Alan Stoudemire, "Medical Assessment of Patients Presenting with Psychiatric Symptoms in the Emergency Setting," *Psychiatric Clinics of North America*, Vol. 22, No. 4, December 1, 1999, pp. 819–850.

- Landon, Bruce E., Ira B. Wilson, and Paul D. Cleary, "A Conceptual Model of the Effects of Health Care Organizations on the Quality of Medical Care," *JAMA*, Vol. 279, No. 17, May 6, 1998, pp. 1377–1382.
- Landon, Bruce E., Alan M. Zaslavsky, Nancy Dean Beaulieu, James A. Shaul, and Paul D. Cleary, "Health Plan Characteristics and Consumers' Assessments of Quality," *Health Affairs*, Vol. 20, No. 2, March 2001, pp. 274–286.
- Lang, Ariel J., and Murray B. Stein, "An Abbreviated PTSD Checklist for Use as a Screening Instrument in Primary Care," *Behaviour Research and Therapy*, Vol. 43, No. 5, May 2005, pp. 585–594.
- Lapham, Gwen T., Carol E. Achtmeyer, Emily C. Williams, Eric J. Hawkins, Daniel R. Kivlahan, and Katharine A. Bradley, "Increased Documented Brief Alcohol Interventions with a Performance Measure and Electronic Decision Support," *Medical Care*, Vol. 50, No. 2, February 2012, pp. 179–187.
- Lave, Judith R., Richard G. Frank, Herbert C. Schulberg, and Mark S. Kamlet, "Cost-Effectiveness of Treatments for Major Depression in Primary Care Practice," *Archives of General Psychiatry*, Vol. 55, No. 7, July 1998, pp. 645–651.
- Le Fauve, Charlene E., Raye Z. Litten, Carrie L. Randall, Darlene H. Moak, Ihsan M. Salloum, and Alan I. Green, "Pharmacological Treatment of Alcohol Abuse/Dependence with Psychiatric Comorbidity," *Alcoholism: Clinical and Experimental Research*, Vol. 28, No. 2, February 2004, pp. 302–312.
- Leskin, G. A., and D. A. Westrup, "Theoretical and Empirical Approaches to PTSD Screening Measurement Construction," paper presented at Screening for PTSD in Primary Care, 15th Annual Meeting of the International Society for Traumatic Stress Studies, Miami, Fla., 1999.
- Linehan, Marsha M., Katherine Anne Comtois, Angela M. Murray, Milton Z. Brown, Robert J. Gallop, Heidi L. Heard, Kathryn E. Korslund, Darren A. Tutek, Sarah K. Reynolds, and Noam Lindenboim, "Two-Year Randomized Controlled Trial and Follow-Up of Dialectical Behavior Therapy vs Therapy by Experts for Suicidal Behaviors and Borderline Personality Disorder," *Archives of General Psychiatry*, Vol. 63, No. 7, July 2006, pp. 757–766.
- Löwe, Bernd, Kurt Kroenke, Wolfgang Herzog, and Kerstin Gräfe, "Measuring Depression Outcome with a Brief Self-Report Instrument: Sensitivity to Change of the Patient Health Questionnaire (PHQ-9)," *Journal of Affective Disorders*, Vol. 81, No. 1, July 2004, pp. 61–66.
- Löwe, Bernd, Jürgen Unützer, Christopher M. Callahan, Anthony J. Perkins, and Kurt Kroenke, "Monitoring Depression Treatment Outcomes with the Patient Health Questionnaire-9," *Medical Care*, Vol. 42, No. 12, December 2004, pp. 1194–1201.
- Luppa, Melanie, Sven Heinrich, Matthias C. Angermeyer, Hans-Helmut König, and Steffi G. Riedel-Heller, "Cost-of-Illness Studies of Depression: A Systematic Review," *Journal of Affective Disorders*, Vol. 98, No. 1–2, February 2007, pp. 29–43.
- Lyness, Jeffrey M., Tamson Kelly Noel, Christopher Cox, Deborah A. King, Yeates Conwell, and Eric D. Caine, "Screening for Depression in Elderly Primary Care Patients: A Comparison of the Center for Epidemiologic Studies—Depression Scale and the Geriatric Depression Scale," *Archives of Internal Medicine*, Vol. 157, No. 4, 1997, pp. 449–454.
- MacDonald, Joanna, Nicola Brown, and Peter Ellis, "Using Telephone Prompts to Improve Initial Attendance at a Community Mental Health Center," *Psychiatric Services*, Vol. 51, No. 6, June 1, 2000, pp. 812–814.
- Management of Bipolar Disorder Working Group, *Clinical Practice Guideline: Management of Bipolar Disorder in Adults (BD)*, Version 2.0, May 2010. As of September 25, 2013:
http://www.healthquality.va.gov/bipolar/bd_305_full.pdf
- Management of MDD Working Group, *Clinical Practice Guideline: Management of Major Depressive Disorder (MDD)*, Version 2.0, May 2009. As of September 25, 2013:
http://www.healthquality.va.gov/mdd/MDD_FULL_3c1.pdf
- Management of Post-Traumatic Stress Working Group, *VA/DoD Clinical Practice Guideline: Management of Post-Traumatic Stress*, Version 2.0, 2010. As of May 7, 2014:
<http://www.healthquality.va.gov/PTSD-Full-2010c.pdf>

- Management of Substance Use Disorders Working Group, *Clinical Practice Guideline: Management of Substance Use Disorders*, Version 2.0, August 2009. As of May 7, 2014: http://www.dcoe.mil/content/navigation/documents/sud_full_601f.pdf
- Marks, Isaac, Karina Lovell, Homa Noshirvani, Maria Livanou, and Sian Thrasher, "Treatment of Posttraumatic Stress Disorder by Exposure and/or Cognitive Restructuring: A Controlled Study," *Archives of General Psychiatry*, Vol. 55, No. 4, April 1998, pp. 317–325.
- Martsof, Grant R., Jeffrey A. Alexander, Yunfeng Shi, Lawrence P. Casalino, Diane R. Rittenhouse, Dennis P. Scanlon, and Stephen M. Shortell, "The Patient-Centered Medical Home and Patient Experience," *Health Services Research*, Vol. 47, No. 6, December 2012, pp. 2273–2295.
- McHorney, Colleen A., John E. Ware Jr., and Anastasia E. Raczek, "The MOS 36-Item Short-Form Health Status Survey (SF-36): II. Psychometric and Clinical Tests of Validity in Measuring Physical and Mental Health Constructs," *Medical Care*, Vol. 31, No. 3, March 1993, pp. 247–263.
- McKeown, Robert E., Steven P. Cuffe, and Richard M. Schulz, "US Suicide Rates by Age Group, 1970–2002: An Examination of Recent Trends," *American Journal of Public Health*, Vol. 96, No. 10, October 2006, pp. 1744–1751.
- Means-Christensen, Adrienne J., Cathy D. Sherbourne, Peter P. Roy-Byrne, Michelle G. Craske, and Murray B. Stein, "Using Five Questions to Screen for Five Common Mental Disorders in Primary Care: Diagnostic Accuracy of the Anxiety and Depression Detector," *General Hospital Psychiatry*, Vol. 28, No. 2, March–April 2006, pp. 108–118.
- Meichenbaum, Donald, *Cognitive Behavior Modification*, Morristown, N.J.: General Learning Press, 1974.
- Melartin, Tarja K., Heikki J. Rytsälä, Ulla S. Leskelä, Paula S. Lestelä-Mielonen, T. Petteri Sokero, and Erkki T. Isometsä, "Continuity Is the Main Challenge in Treating Major Depressive Disorder in Psychiatric Care," *Journal of Clinical Psychiatry*, Vol. 66, No. 2, February 2005, pp. 220–227.
- Military Health System, "Quality Patient Care," date unknown; referenced August 23, 2013.
- Miller, Michael Craig, Douglas G. Jacobs, and Thomas G. Gutheil, "Talisman or Taboo: The Controversy of the Suicide-Prevention Contract," *Harvard Review of Psychiatry*, Vol. 6, No. 2, 1998, pp. 78–87.
- Mitchell, Alex J., and Thomas Selmes, "Why Don't Patients Attend Their Appointments? Maintaining Engagement with Psychiatric Services," *Advances in Psychiatric Treatment*, Vol. 13, No. 6, November 2007, pp. 423–434.
- Moncrieff, J., S. Wessely, and R. Hardy, "Active Placebos Versus Antidepressants for Depression," *Cochrane Database of Systematic Reviews*, No. 1, 2004, p. CD003012.
- Monson, Candice M., Jaimie L. Gradus, Yinong Young-Xu, Paula P. Schnurr, Jennifer L. Price, and Jeremiah A. Schumm, "Change in Posttraumatic Stress Disorder Symptoms: Do Clinicians and Patients Agree?" *Psychological Assessment*, Vol. 20, No. 2, June 2008, pp. 131–138.
- Montgomery, Stuart A., and Marie Asberg, "A New Depression Scale Designed to Be Sensitive to Change," *British Journal of Psychiatry*, Vol. 134, No. 4, April 1979, pp. 382–389.
- Montori, Victor M., Juan Pablo Brito, and M. Hassan Murad, "The Optimal Practice of Evidence-Based Medicine: Incorporating Patient Preferences in Practice Guidelines," *JAMA*, Vol. 310, No. 23, December 18, 2013, pp. 2503–2504.
- Moriarty, David G., Mathew M. Zack, and Rosemarie Kobau, "The Centers for Disease Control and Prevention's Healthy Days Measures: Population Tracking of Perceived Physical and Mental Health over Time," *Health and Quality of Life Outcomes*, Vol. 1, No. 1, September 2003, p. 37.
- Morrison, Alan S., *Screening in Chronic Disease*, 2nd ed., New York: Oxford University Press, 1992.
- Nacasch, Nitzah, Edna B. Foa, Jonathan D. Huppert, Dana Tzur, Leah Fostick, Yula Dinstein, Michael Polliack, and Joseph Zohar, "Prolonged Exposure Therapy for Combat- and Terror-Related Posttraumatic Stress Disorder: A Randomized Control Comparison with Treatment as Usual," *Journal of Clinical Psychiatry*, Vol. 72, No. 9, September 2011, pp. 1174–1180.

Nassir Ghaemi, S., Christopher J. Miller, Douglas A. Berv, Jeffrey Klugman, Klara J. Rosenquist, and Ronald W. Pies, "Sensitivity and Specificity of a New Bipolar Spectrum Diagnostic Scale," *Journal of Affective Disorders*, Vol. 84, No. 2–3, February 2005, pp. 273–277.

National Center for PTSD, "Using the PTSD Checklist," 2012.

———, "DSM-5 Validated Measures," referenced 2013 update January 30, 2014; last updated February 6, 2014. As of May 14, 2014:

http://www.ptsd.va.gov/professional/assessment/DSM_5_Validated_Measures.asp

National Committee for Quality Assurance, "HEDIS 2013," c. 2012a. As of August 26, 2013:

<http://www.ncqa.org/HEDISQualityMeasurement/HEDISMeasures/HEDIS2013.aspx>

———, "NQF Endorsed #0576," 2012b.

———, *Improving Quality and Patient Experience: The State of Health Care Quality 2013*, 2013a. As of December 20, 2014:

http://www.ncqa.org/Portals/0/Newsroom/SOHC/2013/SOHC-web_version_report.pdf

———, *Medicare Health Outcomes Survey*, 2013b. As of September 30, 2013:

http://www.hosonline.org/surveys/hos/download/HOS_2013_Survey.pdf

National Institute for Health and Care Excellence, "Depression: Management of Depression in Primary and Secondary Care—NICE Guidance," Clinical Guidelines CG23, December 2004.

National Quality Forum, "Find Measures," undated (a); referenced September 10, 2013. As of May 12, 2014:

<http://www.qualityforum.org/Qps/QpsTool.aspx>

———, "Measure Evaluation Criteria," undated (b); referenced August 26, 2013. As of May 12, 2014:

http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx

———, "Who We Are," undated (c); referenced September 10, 2013.

———, *National Voluntary Consensus Standards for Patient Outcomes 2009: A Consensus Report*, July 2011a. As of January 30, 2014:

<http://www.qualityforum.org/Publications/2011/07/>

[National_Voluntary_Consensus_Standards_for_Patient_Outcomes_2009.aspx](http://www.qualityforum.org/Publications/2011/07/National_Voluntary_Consensus_Standards_for_Patient_Outcomes_2009.aspx)

———, "NQF #0418 Screening for Clinical Depression," last update September 23, 2011b.

———, *National Voluntary Consensus Standards for Cost and Resource Use: Final Report*, April 2012a. As of September 13, 2013:

<http://www.qualityforum.org/Publications/2012/04/>

[National_Voluntary_Consensus_Standards_for_Cost_and_Resource_Use.aspx](http://www.qualityforum.org/Publications/2012/04/National_Voluntary_Consensus_Standards_for_Cost_and_Resource_Use.aspx)

———, "NQF #0008 Experience of Care and Health Outcomes (ECHO) Survey," last updated September 17, 2012.

———, *MAP Pre-Rulemaking Report: 2013 Recommendations on Measures Under Consideration by HHS, Final Report*, February 2013a. As of May 19, 2014:

http://www.qualityforum.org/Publications/2013/02/MAP_Pre-Rulemaking_Report_-_February_2013.aspx

———, "NQF #0109—Bipolar Disorder and Major Depression: Assessment for Manic or Hypomanic Behaviors," last updated February 6, 2013b.

———, "NQF #0110 Bipolar Disorder and Major Depression: Appraisal for Alcohol or Chemical Substance Use," last updated February 6, 2013c.

———, "NQF #0576 Follow-Up After Hospitalization for Mental Illness," last updated February 28, 2013d.

———, "NQF #0710 Depression Remission at 12 Months," last updated March 7, 2013e.

———, "NQF #1884 Measure Under Consideration: Depression Response at 6 Months—Progress Towards Remission," last updated March 7, 2013f.

———, "NQF #1885 Measure Under Consideration: Depression Response at 12 Months—Progress Towards Remission," last updated March 7, 2013g.

———, "NQF #0105: Antidepressant Medication Management," last updated March 9, 2013h.

———, “NQF #0104 Adult Major Depressive Disorder: Suicide Risk Assessment,” last updated March 15, 2013i.

———, “NQF #0711 Depression Remission at 6 Months,” last updated April 24, 2013j.

———, “NQF #0712 Depression Utilization of the PHQ-9 Tool,” last updated April 24, 2013k.

NCQA—See National Committee for Quality Assurance.

Nelson, David E., Deborah Holtzman, Julie Bolen, Carol A. Stanwyck, and Karin A. Mack, “Reliability and Validity of Measures from the Behavioral Risk Factor Surveillance System (BRFSS),” *Sozial- und Präventivmedizin*, Vol. 46, 2000, pp. S3–S42.

Newschaffer, Craig J., *Validation of Behavioral Risk Factor Surveillance System (BRFSS) HRQOL Measures in a Statewide Sample: Final Report*, Atlanta, Ga.: Centers for Disease Control and Prevention, 1998.

Nicholas, Will, Donna O. Farley, Mary E. Vaiana, and Shan Cretin, *Putting Practice Guidelines to Work in the Department of Defense Medical System: A Guide for Action*, Santa Monica, Calif.: RAND Corporation, MR-1267-A, 2001. As of May 12, 2014:

http://www.rand.org/pubs/monograph_reports/MR1267.html

Norris, Fran H., and Julia L. Perilla, “The Revised Civilian Mississippi Scale for PTSD: Reliability, Validity, and Cross-Language Stability,” *Journal of Traumatic Stress*, Vol. 9, No. 2, April 1996, pp. 285–298.

NQF—See National Quality Forum.

Obama, Barack, “Executive Order: Improving Access to Mental Health Services for Veterans, Service Members, and Military Families,” Washington, D.C.: White House, Executive Order 13625, August 31, 2012. As of May 12, 2014:

<http://www.gpo.gov/fdsys/pkg/FR-2012-09-05/pdf/2012-22062.pdf>

O’Donnell, Meaghan L., Mark C. Creamer, Ruth Parslow, Peter Elliott, Alexander C. Holmes, Steven Ellen, Rodney Judson, Alexander C. McFarlane, Derrick Silove, and Richard A. Bryant, “A Predictive Screening Index for Posttraumatic Stress Disorder and Depression Following Traumatic Injury,” *Journal of Consulting and Clinical Psychology*, Vol. 76, No. 6, December 2008, pp. 923–932.

O’Donnell, Meaghan L., Mark Creamer, and Phillipa Pattison, “Posttraumatic Stress Disorder and Depression Following Trauma: Understanding Comorbidity,” *American Journal of Psychiatry*, Vol. 161, No. 8, August 1, 2004, pp. 1390–1396.

Oldham, John, “Advanced Access in Primary Care,” January 2001.

Opsut, Robert, “Per Member Per Month (PMPM) for Dummies: What It Means and What to Do About It,” *2010 Military Health System Conference: Sharing Knowledge: Achieving Breakthrough Performance*, National Harbor, Md., 2010.

———, “Managing Per Member Per Month (PMPM): Findings Regarding Variation and Drivers of Cost,” *2011 Military Health Conference: The Quadruple Aim: Working Together, Achieving Success*, National Harbor, Md., January 25, 2011. As of May 12, 2014:

<http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA556294>

Oxman, Thomas E., Allen J. Dietrich, John W. Williams Jr., Charles C. Engel, Mathew Friedman, Paula Schnurr, Stanley Rosenberg, and Sheila L. Barry, *RESPECT-Mil Primary Care Clinician’s Manual: Three Component Model for Primary Care Management of Depression and PTSD (Military Version)*, 3CM, August 2008. As of May 12, 2014:

http://www.pdhealth.mil/respect-mil/downloads/PCC_Final.pdf

Parkerson, George R., W. Eugene Broadhead, and Chiu-Kit Tse, “Anxiety and Depressive Symptom Identification Using the Duke Health Profile,” *Journal of Clinical Epidemiology*, Vol. 49, No. 1, January 1996, pp. 85–93.

Perlman, Sharon E., Stephen Friedman, Sandro Galea, Hemanth P. Nair, Monika Erős-Sarnyai, Steven D. Stellman, Jeffrey Hon, and Carolyn M. Greene, “Short-Term and Medium-Term Health Effects of 9/11,” *Lancet*, Vol. 378, No. 9794, September 3–9, 2011, pp. 925–934.

Porter, Michael E., “What Is Value in Health Care?” *New England Journal of Medicine*, Vol. 363, No. 26, December 23, 2010, pp. 2477–2481.

Post-Deployment Health Guideline Expert Panel, *Recommendations for Monitoring Metrics: DoD/VA Practice Guideline for Post-Deployment Health Evaluation and Management*, July 6, 2001. As of September 13, 2013: http://www.pdhealth.mil/guidelines/downloads/view/3/2_recommendations_for_metrics.pdf

Powers, M. B., J. M. Halpern, M. P. Ferenschak, S. J. Gillihan, and E. B. Foa, "A Meta-Analytic Review of Prolonged Exposure for Posttraumatic Stress Disorder," *Clinical Psychology Review*, Vol. 30, No. 6, August 2010, pp. 635–641.

Prins, Annabel, Paige Ouimette, Rachel Kimerling, Rebecca P. Cameron, Daniela S. Hugelshofer, Jennifer Shaw-Hegwer, Ann Thraillkill, Fred D. Gusman, and Javaid I. Sheikh, "The Primary Care PTSD Screen (PC-PTSD): Development and Operating Characteristics," *Primary Care Psychiatry*, Vol. 9, No. 1, 2003, pp. 9–14. As of May 12, 2014: <http://www.ptsd.va.gov/professional/articles/article-pdf/id26676.pdf>

Quitkin, Frederic M., Eva Petkova, Patrick J. McGrath, Bonnie Taylor, Charles Beasley, Jonathan Stewart, Jay Amsterdam, Maurizio Fava, Jerrold Rosenbaum, Frederick Reimherr, Jan Fawcett, Ying Chen, and Donald Klein, "When Should a Trial of Fluoxetine for Major Depression Be Declared Failed?" *American Journal of Psychiatry*, Vol. 160, No. 4, 2003, pp. 734–740.

Rabin, Rosalind, and Frank de Charro, "EQ-5D: A Measure of Health Status from the EuroQol Group," *Annals of Medicine*, Vol. 33, No. 5, July 2001, pp. 337–343.

Radloff, Lenore Sawyer, "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population," *Applied Psychological Measurement*, Vol. 1, No. 3, June 1977, pp. 385–401.

Ramchand, Rajeev, Joie Acosta, Rachel M. Burns, Lisa H. Jaycox, and Christopher G. Pernin, *The War Within: Preventing Suicide in the U.S. Military*, Santa Monica, Calif.: RAND Corporation, MG-953-OSD, 2011. As of May 12, 2014: <http://www.rand.org/pubs/monographs/MG953.html>

Ramchand, Rajeev, Terry L. Schell, Benjamin R. Karney, Karen Chan Osilla, Rachel M. Burns, and Leah Barnes Caldarone, "Disparate Prevalence Estimates of PTSD Among Service Members Who Served in Iraq and Afghanistan: Possible Explanations," *Journal of Traumatic Stress*, Vol. 23, No. 1, February 2010, pp. 59–68.

Resick, Patricia A., Tara E. Galovski, Mary O'Brien Uhlmansiek, Christine D. Scher, Gretchen A. Clum, and Yinong Young-Xu, "A Randomized Clinical Trial to Dismantle Components of Cognitive Processing Therapy for Posttraumatic Stress Disorder in Female Victims of Interpersonal Violence," *Journal of Consulting and Clinical Psychology*, Vol. 76, No. 2, April 2008, pp. 243–258.

Resick, Patricia A., Pallavi Nishith, Terri L. Weaver, Millie C. Astin, and Catherine A. Feuer, "A Comparison of Cognitive-Processing Therapy with Prolonged Exposure and a Waiting Condition for the Treatment of Chronic Posttraumatic Stress Disorder in Female Rape Victims," *Journal of Consulting and Clinical Psychology*, Vol. 70, No. 4, August 2002, pp. 867–879.

Resick, Patricia A., Lauren F. Williams, Michael K. Suvak, Candice M. Monson, and Jaimie L. Gradus, "Long-Term Outcomes of Cognitive-Behavioral Treatments for Posttraumatic Stress Disorder Among Female Rape Survivors," *Journal of Consulting and Clinical Psychology*, Vol. 80, No. 2, April 2012, pp. 201–210.

Rittenhouse, Diane R., Lawrence P. Casalino, Robin R. Gillies, Stephen M. Shortell, and Bernard Lau, "Measuring the Medical Home Infrastructure in Large Medical Groups," *Health Affairs*, Vol. 27, No. 5, September–October 2008, pp. 1246–1258.

Rittenhouse, Diane R., Lawrence P. Casalino, Stephen M. Shortell, Sean R. McClellan, Robin R. Gillies, Jeffrey A. Alexander, and Melinda L. Drum, "Small and Medium-Size Physician Practices Use Few Patient-Centered Medical Home Processes," *Health Affairs*, Vol. 30, No. 8, June 2011, pp. 1575–1584.

Roberts, Neil P., Neil J. Kitchiner, Justin Kenardy, and Jonathan Bisson, "Multiple Session Early Psychological Interventions for the Prevention of Post-Traumatic Stress Disorder," *Cochrane Database of Systematic Reviews*, Vol. 3, July 8, 2009, p. CD006869.

Romano, Patrick, Peter Hussey, and Dominique Ritley, *Selecting Quality and Resource Use Measures: A Decision Guide for Community Quality Collaboratives*, Rockville, Md.: Agency for Healthcare Research and Quality, Publication 09(10)-0073, May 2010. As of September 10, 2013: <http://www.ahrq.gov/qual/perfmeasguide/perfmeas.pdf>

- Rona, Roberto J., Kenneth C. Hyams, and Simon Wessely, "Screening for Psychological Illness in Military Personnel," *JAMA*, Vol. 293, No. 10, April 2005, pp. 1257–1260.
- Rose, Suzanna C., Jonathan Bisson, Rachel Churchill, and Simon Wessely, "Psychological Debriefing for Preventing Post Traumatic Stress Disorder (PTSD)," *Cochrane Database of Systematic Reviews*, Vol. 2, No. 2, 2002, p. CD000560.
- Rosen, Craig S., Caryn DiLandro, Kim N. Corwin, Kent D. Drescher, Jean H. Cooney, and Fred Gusman, "Telephone Monitoring and Support for Veterans with Chronic Posttraumatic Stress Disorder: A Pilot Study," *Community Mental Health Journal*, Vol. 42, No. 5, October 2006, pp. 501–508.
- Royal Australian and New Zealand College of Psychiatrists Clinical Practice Guidelines Team for Depression, "Australian and New Zealand Clinical Practice Guidelines for the Treatment of Depression," *Australian and New Zealand Journal of Psychiatry*, Vol. 38, No. 6, 2004, pp. 389–407.
- Ruggiero, Kenneth J., Kevin Del Ben, Joseph R. Scotti, and Aline E. Rabalais, "Psychometric Properties of the PTSD Checklist—Civilian Version," *Journal of Traumatic Stress*, Vol. 16, No. 5, October 2003, pp. 495–502.
- Rush, A. John, Christina M. Gullion, Monica R. Basco, Robin B. Jarrett, and Madhukar H. Trivedi, "The Inventory of Depressive Symptomatology (IDS): Psychometric Properties," *Psychological Medicine*, Vol. 26, No. 3, May 1996, pp. 477–486.
- Rush, A. John, Madhukar H. Trivedi, Hicham M. Ibrahim, Thomas J. Carmody, Bruce Arnow, Daniel N. Klein, John C. Markowitz, Philip T. Ninan, Susan Kornstein, Rachel Manber, M. E. Thase, J. H. Kocsis, and M. B. Keller, "The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): A Psychometric Evaluation in Patients with Chronic Major Depression," *Biological Psychiatry*, Vol. 54, No. 5, September 1, 2003, pp. 573–583.
- Rush, A. John, Madhukar H. Trivedi, Stephen R. Wisniewski, Andrew A. Nierenberg, Jonathan W. Stewart, Diane Warden, George Niederehe, Michael E. Thase, Philip W. Lavori, Barry D. Lebowitz, Patrick J. McGrath, Jerrold F. Rosenbaum, Harold A. Sackeim, David J. Kupfer, James Luther, and Maurizio Fava, "Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report," *American Journal of Psychiatry*, Vol. 163, No. 11, November 1, 2006, pp. 1905–1917.
- Sareen, Jitender, Tanya Houlahan, Brian J. Cox, and Gordon J. Asmundson, "Anxiety Disorders Associated with Suicidal Ideation and Suicide Attempts in the National Comorbidity Survey," *Journal of Nervous and Mental Disease*, Vol. 193, No. 7, July 2005, pp. 450–454.
- Schell, Terry L., and Grant N. Marshall, "Survey of Individuals Previously Deployed for OEF/OIF," in Terri Tanielian and Lisa H. Jaycox, eds., *Invisible Wounds of War: Psychological and Cognitive Injuries, Their Consequences, and Services to Assist Recovery*, Santa Monica, Calif.: RAND Corporation, MG-720-CCF, 2008, pp. 87–116. As of May 12, 2014:
<http://www.rand.org/pubs/monographs/MG720.html>
- Schnurr, Paula P., and M. K. Jankowski, "Physical Health and Post-Traumatic Stress Disorder: Review and Synthesis," *Seminars in Clinical Neuropsychiatry*, Vol. 4, No. 4, October 1999, pp. 295–304.
- Schnurr, Paula P., Matthew J. Friedman, Charles C. Engel, Edna B. Foa, M. Tracie Shea, Bruce K. Chow, Patricia A. Resick, Veronica Thurston, Susan M. Orsillo, Rodney Haug, Carole Turner, and Nancy Bernardy, "Cognitive Behavioral Therapy for Posttraumatic Stress Disorder in Women: A Randomized Controlled Trial," *JAMA*, Vol. 297, No. 8, February 28, 2007, pp. 820–830.
- Seale, J. Paul, John M. Boltri, Sylvia Shellenberger, Mary M. Velasquez, Monica Cornelius, Monique Guyinn, Ike Okosun, and Heather Sumner, "Primary Care Validation of a Single Screening Question for Drinkers," *Journal of Studies on Alcohol*, Vol. 67, No. 5, September 2006, pp. 778–784.
- Shaul, James, Susan V. Eisen, Brian R. Clarridge, Vickie L. Stringfellow, F. Fowler, and Paul D. Cleary, *Experience of Care and Health Outcomes (ECHO) Survey Field Test Report: Survey Evaluation*, Boston, Mass.: Harvard Medical School, Department of Health Care Policy, 2001.
- Sheehan, D. V., K. Harnett-Sheehan, and B. A. Raj, "The Measurement of Disability," *International Clinical Psychopharmacology*, Vol. 11, June 1996, pp. 89–95.

Shih, Margaret, and Paul A. Simon, "Health-Related Quality of Life Among Adults with Serious Psychological Distress and Chronic Medical Conditions," *Quality of Life Research*, Vol. 17, No. 4, May 2008, pp. 521–528.

Shiner, Brian, Leonard W. D'Avolio, Thien M. Nguyen, Maha H. Zayed, Bradley V. Watts, and Louis Fiore, "Automated Classification of Psychotherapy Note Text: Implications for Quality Assessment in PTSD Care," *Journal of Evaluation in Clinical Practice*, Vol. 18, No. 3, June 2012, pp. 698–701.

Shugar, Gerald, Stephen Schertzer, Brenda B. Toner, and Ines Di Gasbarro, "Development, Use, and Factor Analysis of a Self-Report Inventory for Mania," *Comprehensive Psychiatry*, Vol. 33, No. 5, September–October 1992, pp. 325–331.

Simon, Gregory E., "Evidence Review: Efficacy and Effectiveness of Antidepressant Treatment in Primary Care," *General Hospital Psychiatry*, Vol. 24, No. 4, July–August 2002, pp. 213–224.

Simon, Gregory E., Willard G. Manning, David J. Katzelnick, Steven D. Pearson, Henry J. Henk, and Cindy P. Helstad, "Cost-Effectiveness of Systematic Depression Treatment for High Utilizers of General Medical Care," *Archives of General Psychiatry*, Vol. 58, No. 2, February 2001, pp. 181–187.

Sorbero, Melony E., Thomas E. Mannle, Brad Smith, Katherine E. Watkins, Abigail Woodroffe, Susan M. Paddock, Lisa R. Shugarman, E. Dela Cruz, Jacob Solomon, Q. Burkhart, Teryn Mattox, and Harold A. Pincus, *Program Evaluation of VHA Mental Health Services: Administrative Data Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Spates, C. Richard, Ellen Koch, Karen Cusack, Sherry Pagato, and Stacey Waller, "Eye Movement Desensitization and Reprocessing," in Edna B. Foa, Terence M. Keane, Matthew J. Friedman, and Judith A. Cohen, eds., *Effective Treatments for PTSD: Practice Guidelines from the International Society for Traumatic Stress Studies*, New York: Guilford Press, 2009.

Spearing, Melissa K., Robert M. Post, Gabriele S. Leverich, Diane Brandt, and Willem Nolen, "Modification of the Clinical Global Impressions (CGI) Scale for Use in Bipolar Illness (BP): The CGI-BP," *Psychiatry Research*, Vol. 73, No. 3, December 5, 1997, pp. 159–171.

Spitzer, Robert L., Kurt Kroenke, and Janet B. W. Williams, "Validation and Utility of a Self-Report Version of PRIME-MD: The PHQ Primary Care Study," *JAMA*, Vol. 282, No. 18, November 10, 1999, pp. 1737–1744.

Stein, Dan J., Jonathan C. Ipser, and Soraya Seedat, "Pharmacotherapy for Post Traumatic Stress Disorder (PTSD) (Review)," *Cochrane Database of Systematic Reviews*, No. 1, 2009, p. CD002795.

Svanborg, Pär, and Marie Åsberg, "A Comparison Between the Beck Depression Inventory (BDI) and the Self-Rating Version of the Montgomery Åsberg Depression Rating Scale (MADRS)," *Journal of Affective Disorders*, Vol. 64, No. 2, May 2001, pp. 203–216.

Tarrier, N., H. Pilgrim, C. Sommerfield, B. Faragher, M. Reynolds, E. Graham, and C. Barrowclough, "A Randomized Trial of Cognitive Therapy and Imaginal Exposure in the Treatment of Chronic Posttraumatic Stress Disorder," *Journal of Consulting and Clinical Psychology*, Vol. 67, No. 1, February 1999, pp. 13–18.

Taylor, Steven, Dana S. Thordarson, Louise Maxfield, Ingrid C. Fedoroff, Karina Lovell, and John Ogradniczuk, "Comparative Efficacy, Speed, and Adverse Effects of Three PTSD Treatments: Exposure Therapy, EMDR, and Relaxation Training," *Journal of Consulting and Clinical Psychology*, Vol. 71, No. 2, April 2003, pp. 330–338.

Tehrani, Noreen, Sara Jane Cox, and Tom Cox, "Assessing the Impact of Stressful Incidents in Organizations: The Development of an Extended Impact of Events Scale," *Counselling Psychology Quarterly*, Vol. 15, No. 2, 2002, pp. 191–200.

Thomas, Jeffrey L., Joshua E. Wilk, Lyndon A. Riviere, Dennis McGurk, Carl A. Castro, and Charles W. Hoge, "Prevalence of Mental Health Problems and Functional Impairment Among Active Component and National Guard Soldiers 3 and 12 Months Following Combat in Iraq," *Archives of General Psychiatry*, Vol. 67, No. 6, June 2010, pp. 614–623.

Trangle, M., B. Dieperink, T. Gabert, B. Haight, B. Lindvall, J. Mitchell, H. Novak, D. Rich, D. Rossmiller, L. Setterlund, and K. Somers, "Institute for Clinical Systems Improvement: Major Depression in Adults in Primary Care," May 2012.

TRICARE Management Activity and Office of the Chief Financial Officer, Defense Health Cost Assessment and Program Evaluation, *Evaluation of the Tricare Program: Access, Cost, and Quality—Fiscal Year 2012 Report to Congress*, Falls Church, Va.: Office of the Assistant Secretary of Defense for Health Affairs, February 28, 2012. As of August 23, 2013:

http://tricare.mil/tma/dhcape/program/downloads/TRICARE2012_02_28v5.pdf

Trivedi, Madhukar H., A. John Rush, Stephen R. Wisniewski, Andrew A. Nierenberg, Diane Warden, Louise Ritz, Grayson Norquist, Robert H. Howland, Barry Lebowitz, Patrick J. McGrath, Kathy Shores-Wilson, Melanie M. Biggs, G. K. Balasubramani, and Maurizio Fava, "Evaluation of Outcomes with Citalopram for Depression Using Measurement-Based Care in STAR*D: Implications for Clinical Practice," *American Journal of Psychiatry*, Vol. 163, No. 1, January 1, 2006, pp. 28–40.

U.S. Department of Veterans Affairs, Office of Management, Office of Performance Management, *Department of Veterans Affairs (VA) 2011 Performance and Accountability Report*, November 15, 2011. As of May 7, 2014: <http://www.va.gov/budget/report/>

———, *2012 Performance and Accountability Report: Honoring All Who Have Served*, 2012a. As of May 7, 2014:

<http://www.va.gov/budget/report/>

———, Office of Inspector General, Office of Audits and Evaluations and Healthcare Inspections, *Veterans Health Administration: Review of Veterans' Access to Mental Health Care*, 12-00900-168, April 23, 2012b. As of March 1, 2013:

<http://www.va.gov/oig/pubs/VAOIG-12-00900-168.pdf>

U.S. National Institutes of Health, "Veterans Telemedicine Outreach for PTSD Services (VTOPS),"

ClinicalTrials.gov, last updated October 25, 2013; referenced 2012. As of September 10, 2013:

<http://clinicaltrials.gov/ct2/show/NCT00645047>

U.S. Preventive Services Task Force, "Screening for Depression in Adults: U.S. Preventive Services Task Force Recommendation Statement," *Annals of Internal Medicine*, Vol. 151, No. 11, 2009, pp. 784–792.

VA—See U.S. Department of Veterans Affairs.

Vaughan, Kevin, "A Trial of Eye Movement Desensitization Compared to Image Habituation Training and Applied Muscle Relaxation in Post-Traumatic Stress Disorder," *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 25, No. 4, December 1994, pp. 283–291.

Vrana, Scott R., and Dean Lauterbach, "The Purdue Posttraumatic Stress Disorder Questionnaire—Revised," Purdue University, 1991.

Watkins, Katherine E., Harold Alan Pincus, Susan Paddock, Brad Smith, Abigail Woodroffe, Carrie Farmer, Melony E. Sorbero, Marcela Horvitz-Lennon, Thomas Mannle Jr., Kimberly A. Hepner, J. Solomon, and C. Call, "Care for Veterans with Mental and Substance Use Disorders: Good Performance, but Room to Improve on Many Measures," *Health Affairs*, Vol. 30, No. 11, November 2011, pp. 2194–2203.

Watkins, Katherine E., Harold Alan Pincus, Brad Smith, Susan M. Paddock, Thomas E. Mannle Jr., Abigail Woodroffe, Jake Solomon, Melony E. Sorbero, Carrie M. Farmer, Kimberly A. Hepner, David M. Adamson, Lanna Forrest, and Catherine Call, *Veterans Health Administration Mental Health Program Evaluation: Capstone Report*, Santa Monica, Calif.: RAND Corporation, TR-956-VHA, 2011. As of May 13, 2014: http://www.rand.org/pubs/technical_reports/TR956.html

Weathers, Frank, Jennifer A. Huska, and Terrence M. Keane, "The PTSD Checklist Military Version (PCL-M)," 1991.

Weathers, Frank, B. T. Litz, D. S. Herman, Jennifer A. Huska, and Terrence M. Keane, "The PTSD Checklist: Reliability, Validity, and Diagnostic Utility," paper presented at Annual Meeting of the International Society for Traumatic Stress Studies, San Antonio, Texas, 1993.

Williams, Edwin R., and Suzanne Moore Shepherd, "Medical Clearance of Psychiatric Patients," *Emergency Medicine Clinics of North America*, Vol. 18, No. 2, May 1, 2000, pp. 185–198.

Williams, Emily C., Chris L. Bryson, Haili Sun, Ryan B. Chew, Lisa D. Chew, David K. Blough, David H. Au, and Katharine A. Bradley, "Association Between Alcohol Screening Results and Hospitalizations for Trauma in Veterans Affairs Outpatients," *American Journal of Drug and Alcohol Abuse*, Vol. 38, No. 1, January 2012, pp. 73–80.

Williams, Janet B. W., "A Structured Interview Guide for the Hamilton Depression Rating Scale," *Archives of General Psychiatry*, Vol. 45, No. 8, 1988, pp. 742–747.

Williams, John W., Jr., Martha Gerrity, Tracey Holsinger, Steve Dobscha, Bradley Gaynes, and Allen Dietrich, "Systematic Review of Multifaceted Interventions to Improve Depression Care," *General Hospital Psychiatry*, Vol. 29, No. 2, March–April 2007, pp. 91–116.

Woodroffe, Abigail, Christine Vaughan, Brad Smith, Jacob Solomon, Andrew L. Crouter, and Robyn Gerdes, *Program Evaluation of VHA Mental Health Services: Facility Survey Report*, Alexandria, Va.: Altarum Institute and RAND–University of Pittsburgh Health Institute, 2010.

Work Group on Major Depressive Disorder, *Practice Guideline for the Treatment of Patients with Major Depressive Disorder*, 3rd ed., Arlington, Va.: American Psychiatric Association, 2010.

Work Group on Psychiatric Evaluation, *Practice Guideline for the Psychiatric Evaluation of Adults*, 2nd ed., Washington, D.C.: American Psychiatric Association, 2006.

Work Group on Suicidal Behaviors, *Practice Guideline for the Assessment and Treatment of Patients with Suicidal Behaviors*, Arlington, Va.: American Psychiatric Association, 2003.

World Health Organization, "Constitution of the World Health Organization," adopted July 22, 1946, New York, and entered into force April 7, 1948. As of June 7, 2014:

<http://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf?ua=1>

———, *Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies*, October 2010. As of May 13, 2014:

<http://www.who.int/healthinfo/systems/monitoring/en/>

Young, R. C., J. T. Biggs, V. E. Ziegler, and D. A. Meyer, "A Rating Scale for Mania: Reliability, Validity and Sensitivity," *British Journal of Psychiatry*, Vol. 133, 1978, pp. 429–435.

Zack, Matthew M., David G. Moriarty, Donna F. Stroup, Earl S. Ford, and Ali H. Mokdad, "Worsening Trends in Adult Health-Related Quality of Life and Self-Rated Health: United States, 1993–2001," *Public Health Reports*, Vol. 119, No. 5, September–October 2004, pp. 493–505.

Zahran, Hatice S., Rosemarie Kobau, David G. Moriarty, Matthew M. Zack, James Holt, and Ralph Donehoo, "Health-Related Quality of Life Surveillance—United States, 1993–2002," *Morbidity and Mortality Weekly Report Surveillance Summaries*, Vol. 54, No. SS04, October 28, 2005, pp. 1–35. As of May 13, 2014: <http://www.cdc.gov/Mmwr/preview/mmwrhtml/ss5404a1.htm>

Zlotnick, Caron, Jonathan Davidson, M. Tracie Shea, and Teri Pearlstein, "Validation of the Davidson Trauma Scale in a Sample of Survivors of Childhood Sexual Abuse," *Journal of Nervous and Mental Disease*, Vol. 184, No. 4, April 1996, pp. 255–257.

In recent years, the number of U.S. service members treated for psychological health conditions has increased substantially. In particular, at least two psychological health conditions—posttraumatic stress disorder (PTSD) and major depressive disorder (MDD)—have become more common, with prevalence estimates up to 20 percent for PTSD and 37 percent for MDD. Delivering quality care to service members with these conditions is a high-priority goal for the military health system (MHS). Meeting this goal requires understanding the extent to which the care the MHS provides is consistent with evidence-based clinical practice guidelines and its own standards for quality. To better understand these issues, RAND Corporation researchers developed a framework to identify and classify a set of measures for monitoring the quality of care provided by the MHS for PTSD and MDD. The goal of this project was to identify, develop, and describe a set of candidate quality measures to assess care for PTSD and MDD. To accomplish this goal, the authors performed two tasks: (1) developed a conceptual framework for assessing the quality of care for psychological health conditions and (2) identified a candidate set of measures for monitoring, assessing, and improving the quality of care for PTSD and MDD. This document describes their research approach and the candidate measure sets for PTSD and MDD that they identified. The current task did not include implementation planning but provides the foundation for future RAND work to pilot a subset of these measures.



NATIONAL DEFENSE RESEARCH INSTITUTE

www.rand.org

\$55.00

ISBN-10 0-8330-8656-1
ISBN-13 978-0-8330-8656-3

