



Measurement-Based Care in the Treatment of Mental Health and Substance Use Disorders

March 2021

MEADOWS
MENTAL HEALTH
POLICY INSTITUTE

 **The Path Forward**
for mental health and substance use
Health Equity for all Americans

Recommended Citation

Alter, C.L., Mathias, A., Zahniser, J., Shah, S., Schoenbaum, M., Harbin, H.T., McLaughlin, R, & Sieger-Walls, J. (2021, February). *Measurement-Based Care in the Treatment of Mental Health and Substance Use Disorders*. Dallas, TX: Meadows Mental Health Policy Institute. mmhpi.org

Contributing Authors

Carol L. Alter, MD, is the Senior Fellow for Medical Integration at the Meadows Mental Health Policy Institute. calter@mmhpi.org

Amanda Mathias, PhD, is the Senior Director of Innovation at the Meadows Mental Health Policy Institute. amathias@mmhpi.org

James Zahniser, PhD, is the Senior Director of Evaluation Design at the Meadows Mental Health Policy Institute. jzahniser@mmhpi.org

Seema Shah, MD, is the Senior Fellow of Pediatric Health Policy at the Meadows Mental Health Policy Institute. sshah@mmhpi.org

Michael Schoenbaum, PhD, is Senior Advisor for Mental Health Services, Epidemiology, and Economics in the Division of Services and Intervention Research at the National Institute of Mental Health. schoenbaumm@mail.nih.gov

Henry Harbin, MD, is a psychiatrist with over 40 years of experience in the mental health and substance use care field. He has held a number of senior positions in both public and private health care organizations. htharbin@aol.com

Rachael McLaughlin, MA, is the Project Coordinator for Policy Implementation at the Meadows Mental Health Policy Institute. rmclaughlin@mmhpi.org

Jesse Sieger-Walls, MSW, PhD, is the Evaluation and Data Analyst at the Meadows Mental Health Policy Institute. jsieger-walls@mmhpi.org

Additional Contributors

Catie Hilbelink, MPP, Chief of Staff for Policy Implementation, Meadows Mental Health Policy Institute

Melissa Rowan, MSW, MBA, Executive Vice President for Policy Implementation, Meadows Mental Health Policy Institute

Andy Keller, PhD, President, Chief Executive Officer, and Linda Perryman Evans Presidential Chair, Meadows Mental Health Policy Institute

Acknowledgements

This work was commissioned by the Path Forward and funded and produced by the Meadows Mental Health Policy Institute.

The views expressed here are those of the authors and not necessarily those of the National Institute of Mental Health or the federal government.

Contents

Executive Summary	1
Introduction	3
Measurement-Based Care	4
Recent Advances in Measurement-Based Care	4
Enabling and Incentivizing Measurement-Based Care for Mental Health and Substance Use Care	5
Quality Measurement and Paying for Value.....	6
Measurement-Based Care as a Component of Quality Measurement, Reporting, and Reimbursement	8
Accreditation Standards Now Requiring Use of Measurement-Based Care in Mental Health and Substance Use Care	9
Purchasers are Recognizing the Importance of Measurement-Based Care.....	11
Advocating for Measurement-Based Care for Mental Health and Substance Use Conditions	12
Moving Measurement-Based Care Forward: A Call to Action.....	13
Overview of Methods, Recommended Measures, and Measurement-Based Care in Real World Settings	14
Recommended Adult Measures	16
Recommended Pediatric Measures	19
Multi-Condition or Cross-Cutting Tools Used in Adults and Children	21
Appendix 1: Recommended Adult Measures	25
Tool: Patient Health Questionnaire – 9 Items (PHQ-9)	25
Tool: PROMIS Depression	26
Tool: Generalized Anxiety Disorder – 7 Items (GAD-7)	26
Tool: PROMIS Anxiety	26
Tool: Panic Disorder Severity Scale – Self Report (PDSS-SR)	26
Tool: PROMIS Alcohol Use	27
Tool: US-Adapted Alcohol Use Disorders Identification Test-Consumption (USAUDIT-C).....	27
Tool: Brief Addiction Monitor (BAM).....	27
Tool: Substance Abuse Outcomes Module.....	28
Tool: PTSD Checklist (PCL).....	28

Tool: Columbia-Suicide Severity Rating Scale (C-SSRS) 28

Tool: Ask Suicide Screening Questions (ASQ) 28

Tool: Brief Pain Inventory (BPI)..... 29

Tool: Positive and Negative Syndrome Scale (PANSS-6) 29

Tool: Brief Psychiatric Rating Scale (BPRS) 29

Tool: Altman Self-Rated Mania Scale (ASRM)..... 29

Tool: Eating Disorder Examination – Questionnaire Short (EDE-QS) 29

Tool: Eating Attitudes Test (EAT-26)..... 30

Tool: Florida Obsessive-Compulsive Inventory (FOCI)..... 30

Tool: Edinburgh Postnatal Depression Scale (EPDS)..... 30

Tool: Medical Outcomes Study Short-Form Health Survey (SF-12)..... 31

Tool: World Health Organization Disability Assessment Schedule (WHODAS II) 31

Appendix 2: Recommended Pediatric Measures 32

Tool: Patient Health Questionnaire for Adolescents (PHQ-A)..... 32

Tool: PROMIS Depression Scale 32

Tool: Suicide Behavior Questionnaire-Revised (SBQ-R) 32

Tool: Vanderbilt ADHD Rating Scale 32

Tool: Pediatric Symptom Checklist (PSC) 33

Tool: Screen for Child Anxiety Related Emotional Disorders (SCARED) 34

Tool: PROMIS Anxiety 34

Tool: Mood and Feelings Questionnaire (MFQ) 34

Tool: The Brief Addiction Monitor (BAM)..... 34

Tool: Altman Self-Rated Mania Scale (ASRM)..... 35

Tool: Children’s Version of Eating Attitudes Test (ChEAT-26) 35

Tool: Children’s Florida Obsessive-Compulsive Inventory (C-FOCI) 35

Tool: Other PROMIS Measures 35

Appendix 3: PROMIS Instruments and Development..... 36

Appendix 4: APA Protocol and Battery for Patient Assessment and Monitoring..... 39

Appendix 5: List of Abbreviations 41

Endnotes 43

Executive Summary

Abundant evidence has shown that use of repeated, validated rating scales will improve outcomes of mental health and substance use (MH/SU) treatment, just as use of repeated measurement of other health conditions such as blood pressure and blood sugar, for example, improves outcomes in care for other health conditions.¹ Potential outcome improvements are in the range of 20% to 60%.

The evidence base for this work is not new. In 2015, the Kennedy Forum published a report summarizing the data supporting use of measurement-based care (MBC) in MH/SU treatment and provided information on a number of self-report, validated rating scales that could be used in clinical settings for that purpose.² This paper expands upon the data in that Kennedy Forum report to include additional measures and to document additional research supporting the use of MBC for the treatment of MH/SU disorders in primary, specialty, and acute care settings. The report also addresses payment and policy strategies for expanding access to MBC, including development of and participation in value-based payment arrangements with payors and provider and health system accreditation standards created to support integration of MBC into clinical practice.

MBC is defined as the use of repeated, validated measures to track symptoms and functional outcomes in clinical settings. Examples of MBC for other health conditions can include clinician administered measures such as blood pressure or respiratory rate, the use of lab tests like liver function tests, monitoring cholesterol levels, or, in the case of one leading MH/SU condition (depression), repeated use of a patient-reported tool such as the Patient Health Questionnaire (PHQ) 9. Not only has use of MBC in MH/SU care been shown to improve outcomes dramatically, it also provides the foundation for measurement of quality, which impacts health plan accreditation and reimbursement.

Recently, standard setting organizations including the Joint Commission (TJC) and the Utilization Review Accreditation Commission (URAC) have begun to incorporate use of MBC into their accreditation standards. TJC now requires that specialty MH/SU facilities seeking accreditation must utilize MBC in the treatment of all common MH/SU conditions. URAC developed a voluntary standard that provider organizations in primary care, specialty MH/SU, and other specialty care settings are encouraged to implement, and that employers and payors are encouraged to consider when partnering with health systems and provider practices.

Quality measures used more broadly in pay for performance and other value-based care programs have been shown to improve outcomes through the incorporation of MBC data. Aggregate MBC data generated from validated rating scales such as those highlighted in this report, as well as those used to measure patient outcomes or other clinical care processes across many providers and health delivery settings, have been able to improve care outcomes

and reduce care costs (e.g., Medicare Star ratings). Evidence strongly suggests that if a particular outcome is included in a quality measure, those outcomes improve.

But the key to MBC is helping providers implement sufficiently reliable and valid measurements tools needed to accurately assess symptoms, conditions, treatment progress, and functional outcomes. In developing this paper, we conducted a survey of the literature as well as community use of such measures. The assessment tools featured in this report meet the standards of reliability and validity necessary for such use, and we have categorized them by adult and child/youth administration. We have included both measures for specific conditions, as well as tools that can be used to assess overall functioning. We also included measures that are able to screen or facilitate diagnosis and also are sensitive to repeat use in order to assess outcomes. Most importantly, we included measures that have evidence of use in real world settings and were not considered burdensome for clinicians or patients to complete and analyze. Thirty-six rating scales meeting these criteria are described – all of which can be used in primary care or specialty care settings for the MH/SU conditions they address. We also reference a number of proprietary tools and those that can be used to track multiple symptoms and conditions. Several examples of health plans and other providers that have successfully incorporated MBC into their routine care are also provided.

This paper presents important information that can support clinicians and health systems in implementing evidence-based MBC at scale across their entire practice or system for priority MH/SU conditions in order to improve care and outcomes for the patients they serve. Furthermore, it highlights that integration of MBC at scale into clinical practice can also contribute to accreditation, improve quality, and impact financial performance. While there is clear evidence for broad-based use of MBC, widespread adoption of these practices will need to be both required, and just as importantly, reimbursed in order to realize its full promise. All organizations accrediting both medical and behavioral providers and health plans can and should require use of MBC in treatment of MH/SU conditions.

Additional quality measures that are developed to include MBC and focus on clinical outcomes will help facilitate adoption. Development of reimbursement mechanisms which can facilitate use of MBC, as well as refinement of patient reporting and provider analysis of this data within electronic health records, are critical components of successful expansion. Currently there are a limited number of specific billing codes for behavioral MBC tools, and therefore, there is a need for additional reimbursement mechanisms. This is important not only to policy makers, but also to employers, payers, providers, and patients; consistent use of MBC will greatly improve MH/SU care for individuals, outcomes for health systems, and results for health payers.

Introduction

This paper builds on earlier work the Kennedy Forum published in 2015, which called for the nation to embrace measurement-based mental health and substance use (MH/SU) care, an approach to tracking the clinical status of people receiving evidence-based MH/SU interventions that is associated with improved outcomes.³ As use of repeated measures is a core component of the delivery of effective outcomes for patients with other health conditions (i.e., diabetes and hypertension), measurement-based care (MBC) should be a foundation of all MH/SU care. Among other things, the Kennedy Forum’s attention to the importance of MBC for MH/SU care included the following policy statement:

All primary care and MH/SU care providers treating mental health and substance use disorders should implement a system of measurement-based care whereby validated symptom rating scales are completed by patients and reviewed by clinicians during encounters. Measurement-based care will help providers determine whether the treatment is working and facilitate treatment adjustments, consultations, or referrals for higher intensity services when patients are not improving as expected.

As Fortney and colleagues summarized in their notable “tipping point” paper concerning use of MBC in MH/SU care, patients receiving usual care have far worse outcomes than patients who received MBC.⁴ Successful recovery rates are much higher when MBC is utilized. Fortney et al cited studies that found up to a nearly 75% improvement in remission rates between patients receiving MBC for MH/SU and those who received usual care.

Studies find up to a nearly 75% improvement in remission rates between patients receiving MBC for behavioral health and those who received usual care.

-Fortney et al

Along with a supplement that provided clinicians, payers, and quality improvement agencies with a list of commonly used and validated symptom rating scales, the 2015 Kennedy Forum publication defined the essential elements of MBC for MH/SU care, summarized research evidence supporting it, and provided guidance concerning its implementation and use. Since its publication, several opportunities and incentives for utilizing MBC specifically for MH/SU have emerged. There has also been additional development of rating scales that can be used to measure MH/SU outcomes. This paper updates and expands on the Kennedy Forum’s recommendations concerning the use of MBC for MH/SU and its list of standardized patient outcome tools that are currently available and offers examples of how MBC is being used in clinical settings. The paper also provides additional background information

which indicates that increased use of MBC can increase reimbursement opportunities for providers and health systems.

Measurement-Based Care

Measurement-based care (MBC) is defined as the use of repeated, validated measures to track symptoms and outcomes in the clinical setting. Across many fields of medicine, regular and repeated use of validated patient outcome measures – such as HbA1c for patients with diabetes and blood pressure for patients with hypertension – is standard practice to help identify if patients are progressing adequately and to inform treatment adjustment if they are not. Moreover, MBC can both inform clinical decision-making for individual patients, and – by aggregating data from repeated outcome measurement – be used to track and improve quality of care across patient panels, practices, systems, and plans.

Validated and standardized patient outcome measures are also available across a range of mental health and substance use (MH/SU) conditions, mainly based on patient report due to the general lack of clinically valid biomarkers in MH/SU conditions (which is also true, to varying degrees, in other fields). Yet MBC is not yet standard practice in MH/SU care, neither in primary nor in specialty care settings.

However, increasingly health plans, accrediting organizations, and public and private payers are prioritizing outcome-based data to drive access and payment. This paper provides information on tools that can be used to deliver MBC in MH/SU care. It includes a listing of rating scales that have demonstrated validity to both identify and monitor outcomes for common MH/SU conditions. Use of these measures can support quality reporting, help providers meet accreditation standards requiring MBC, and – most importantly – improve patient care.

Recent Advances in Measurement-Based Care

Rapid Growth of Patient-Reported Outcomes Measures

As the integration of healthcare delivery systems has shifted toward patient-centered models of care, the scientific community has responded to the need to quantify patients' experience towards health outcome endpoints. Patient reported outcome measures (PROMs) are surveys that are completed directly by the patient (or family member in some cases) and are designed specifically to capture self-reported (vs. provider-reported) symptoms or severity. These measures have been compared against provider-reported assessments of similar symptoms and syndromes and therefore are highly reliable and efficient ways of assessing mental health and substance use (MH/SU) conditions or other health experiences, like quality of life or physical functioning. Given the importance of patient reported input and the availability of validated PROMs, many organizations have published frameworks, guidance, and standards for the development of PROMs, including the Food and Drug Administration (FDA),^{5, 6} the National Quality Forum (NQF),⁷ the Patient-Reported Outcomes Measurement Information System

(PROMIS),⁸ and the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative.^{9, 10}

The application of these delineated methods for instrument development has led to a great assortment of PROMs that may be used across other health conditions and mental health conditions and care settings.

Electronic Health Records and Patient Portals

Health technology products (e.g., electronic health records and patient portals) have incorporated many standardized outcome measures for MH/SU conditions as well as other health conditions in order to improve patient adherence, clinical outcomes, and patient engagement. Electronic health records (EHRs) now have the capacity to incorporate repeated MH/SU measures into the health record and into patient care workflows. Algorithms that drive repeated assessments for other health outcomes like blood pressure, cholesterol testing, and need for health prevention activities can now be utilized to assess MH/SU outcomes and gaps of care. While most EHR systems contain a number of the tools we describe in their libraries, the ability to easily employ these tools with a user-friendly interface, supporting the use of the tool as a patient reported measure as it was designed, is less common. Importantly, nearly all of the major EHR platforms include a patient-facing portal that can be used to collect patient reported outcome data. Additionally, third-party technology developers have been actively engaged in developing this type of tool, which can be integrated into any EHR. Providers and health systems will need to work with their EHR provider to make sure these rating scales are available, accessible to patients, and can provide the reporting necessary for clinical and regulatory use.

Enabling and Incentivizing Measurement-Based Care for Mental Health and Substance Use Care

Not only does measurement-based care (MBC) lead to improved clinical care for the individual patient, it also provides the foundation for measurement of quality, which impacts health plan accreditation and reimbursement. Over the past ten years, there has been a shift in healthcare from paying for services delivered (paying for volume) to paying for the clinical outcomes achieved and/or the quality of care provided—in other words, paying for value. These “value-based payment programs” create a way for public and commercial payers to incorporate measures of quality care into payment strategies. Most recently, many organizations have begun to rely on the presence of MBC or direct outcome measurement, specifically for mental health and substance use (MH/SU) care delivery, therefore creating a significant incentive for providers and health plans to implement MBC.¹¹

Quality Measurement and Paying for Value

There are many ways that quality is measured and incentivized. Accreditation programs, like The Joint Commission (TJC, historically the Joint Commission on Accreditation of Hospitals) or the National Committee on Quality Assurance (NCQA) have surveyed hospital and health plan processes and procedures to determine whether they would meet standards of care. In addition, these accreditation bodies, as well as payers, are directly making assessments of quality which will impact the reimbursement of providers or health systems.

Quality measurement is the mechanism that is used to standardize the approach to measuring outcomes and quality across health systems and providers. Quality measures are developed and tested using large data sets that are linked to patient outcomes and informed by disease experts. Healthcare providers collect data that they then report to federal and commercial payers; the results of these measures (whether a provider or health system met the defined threshold for a particular measure) can then influence their reimbursement rates. The performance on these measures may also be used more broadly in public rankings of the quality of care delivered by a particular health plan or hospital (e.g., Medicare Star Ratings¹²). Quality measures that report directly on outcomes (e.g., the number of patients with diabetes in a health plan who have evidence of HBA1C levels within normal range, and therefore have well controlled diabetes) provide the most valuable information on whether good care is being delivered.

While clinicians or providers might use a validated rating tool like the Patient Health Questionnaire-9 (PHQ9) to measure a patient's progress toward achieving depression remission, the translation of that clinical activity into a quality measure aggregates the data from many providers in order to assess whether a provider or a group of providers is achieving a benchmark level of depression remission in their patients. Use of the PHQ9 in the clinical setting is an example of MBC, but examination of the scores on all of the PHQ9s over a population of patients leads to reporting on the quality measure (QM). That is, MBC is used to analyze individual patient care level data while quality measurement is used to analyze population health level data.

Various organizations develop QMs, including professional associations and accrediting agencies like the NCQA, TJC, Pharmacy Quality Alliance (PQA), and the University of Southern California (USC). Typically, QMs must undergo further review by an organization like the National Quality Forum (NQF), which may provide consensus-based endorsement for a measure. NQF-endorsed measures are considered the gold standard for healthcare measurement in the United States. Expert committees that are comprised of various stakeholders, including patients, providers, and payers, evaluate measures for NQF endorsement. The federal government and many private sector entities use NQF-endorsed measures above all others because of the rigor and consensus process behind them. Nearly all

NQF-endorsed measures are in use. To earn NQF endorsement designation, quality measures must meet certain standards and demonstrate an ability to provide reliable and accurate assessments of clinical performance.

Payers (including both federal and commercial payers) select QMs to be used by health plans, hospitals, other facilities, and individual providers based on endorsement by such groups as the NQF and whether they fit their specific needs for monitoring clinical care. QMs are also used to support accreditation activities (e.g., NCQA or TJC) or by the Centers for Medicare & Medicaid Services (CMS) to rate the quality of programs it funds. CMS uses such measures to incentivize performance to support its value-based payment arrangements. Importantly, while there are many opportunities for quality measurement reporting across a number of CMS and commercial programs, few specific measures are actually required. TJC and other accrediting organizations or payers may sometimes use measures that are not NQF-endorsed. However, CMS is still the largest user of QMs, and most payers, when requiring providers to report QMs, rely on measures used by CMS (many of which NQF endorses). There are over 1,000 different quality measures utilized across all CMS programs; 49 of them focus on MH/SU care.

However, it is often difficult to measure outcomes over large populations, and so measurement of processes that are closely aligned with health outcomes may be the basis for quality measures. Ideally, there should be robust data to link the care process to a clinical outcome.

In MH/SU care, 95% of quality measures that are used to assess quality in health plans, or that become the basis for reimbursement incentives, are process measures (e.g., percent of people screened, whether children with ADHD have a visit with a provider more than three times in six months for follow up) and do not measure outcomes (e.g., quality of life improvement, symptom reduction, etc.) at all. Furthermore, there is little evidence that the process measures used to rate MH/SU care lead to improved outcomes. Therefore, the reporting on these measures in many cases is not providing any meaningful information on either outcomes achieved or quality of care. One reason cited for the lack of outcome-based MH/SU quality measures is that valid tools do not exist and therefore there is no way to meaningfully measure MH/SU outcomes. However, the Kennedy Forum report and this report provide ample evidence of valid tools that are available for **and are being used in**

In mental health and substance use care, 95% of quality measures that are used to assess quality in health plans, or that become the basis for reimbursement incentives, are process measures and do not measure outcomes at all.

multiple settings for patients with MH/SU conditions – and that use of MBC will improve outcomes for patients.

Measurement-Based Care as a Component of Quality Measurement, Reporting, and Reimbursement

Currently Medicaid, Medicare, and many commercial payers either require or allow providers and health plans to report outcome data on depression treatment.

Most states use a combination of process and outcomes measures that measure satisfaction, quality of care and quality of life through multiple data collection methods. Outcome measures are high-level clinical or financial outcomes that are targeted for improvement. These include mortality rates, readmission rates, surgical site infection rates, and satisfaction and access to care. Process measures quantify the specific steps in a process that lead to outcome metrics. These include the time that it takes for an individual to be seen by a physician, the number of prescriptions that an individual has, or the percentage of individuals with a particular diagnosis receiving preventive tests.

Many states participate in the Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey. CAHPS surveys ask consumers and patients to report on and evaluate their experiences with healthcare. These surveys cover topics that are important to consumers and focus on aspects of quality that consumers are best qualified to assess, such as the communication skills of providers and ease of access to healthcare services.¹³

The depression screening measure and the depression remission measures are reported in several common settings, including Medicare Shared Savings Plan programs. Medicare and Medicare SSP -NQF 418- requires depression screening but does not specify a particular tool or require follow-up in the core set. The Healthcare Effectiveness Data and Information Set (HEDIS) includes depression screening and a 6- and 12-month remission measure, which includes the PHQ9 specifically.

These measures can also be reported through the CMS quality measurement program, Merit Based Incentive Payment System (MIPS). The MIPS program requires providers who deliver care to Medicare beneficiaries to report data on six quality measures. While there are numerous measures that providers can choose to report, the depression remission measures are included in that list. The depression screening quality measure requires providers to screen all patients for depression; if a patient exceeds the scoring threshold, they then should be treated or referred for treatment. While this measure does not measure outcomes of treatment, it does require the use of objective measures to identify and manage depression. Its use has led to widespread screening for depression; however, data continues to suggest that without requirement for repeated monitoring, it is difficult to effectively improve outcomes of depression.^{14, 15}

CMS has prioritized the development of quality measures in MH/SU care, and specifically is supporting development of outcome measures. It recently awarded funds to the American Psychiatric Association to develop a number of new quality measures that can be used in MIPS and other quality measurement programs. The majority of these measures will be focused on the delivery of MBC, covering a number of MH/SU conditions, as well as expanding measurement of outcomes to additional conditions beyond depression.¹⁶

Another example of the requirement for use of MBC is in the reimbursement for the psychiatric Collaborative Care Model (CoCM). CoCM is delivered in the primary care setting to patients with any MH/SU condition. Enrolled patients work with a behavioral health care manager who collects information of patient history and symptoms and then reviews that information with a psychiatric consultant. A key requirement for reimbursement also includes the use of a validated rating scales to track symptoms on a regular basis. This requirement facilitates treatment-to-target and high rates of improvement in clinical outcomes for participants. It also allows providers to report on depression remission quality measures utilizing outcome data collected through the clinical program.

Because the use of MBC has been linked so strongly with improved outcomes, it can serve as an important surrogate for actual improvement in outcomes. Therefore, evidence of use of MBC may soon become the basis for quality measure reporting tied to reimbursement.

Accreditation Standards Now Requiring Use of Measurement-Based Care in Mental Health and Substance Use Care

While quality measures can have a great impact on improving outcomes, development of a quality measure involves significant time and testing across a number of healthcare settings before it can be approved for use by CMS or other payers. However, accreditation standards provide another important opportunity to require elements of care that have been shown to improve outcomes.

Accreditation entities, like TJC, NCQA, URAC (formerly known as the Utilization Review Accreditation Commission), and others, conduct reviews of hospitals, health plans, pharmacies, and health provider organizations. Most hospitals and health systems participate in at least one of these programs and must meet these accreditation requirements to deliver care and remain competitive. Both TJC and URAC have recently added standards that address MBC.

In 2018, TJC, which provides accreditation to hospitals, outpatient MH/SU programs, and others, instituted a new standard which assesses whether MH/SU organizations are routinely using MBC in provision of care. Health systems and providers that are now TJC-accredited for MH/SU care are required to document how many patients with any MH/SU disorder have received screening and follow-up measurement to guide treatment decisions. The standard, TJC Standard CTS.03.01.09, requires that the MH/SU providers use standardized tools to monitor

patients' treatment progress, that the data obtained from repeated measurement is used in treatment planning and delivery, and that the organization compiles and analyzes this data in order to improve quality of care delivered.

Table 1. Joint Commission Standard for Mental Health and Substance Use Providers

Standard CTS.03.01.09 – The organization assesses the outcomes of care, treatment, or services provided to the individual served.
EP 1 – The organization uses a standardized tool or instrument to monitor the individual's progress in achieving his or her care, treatment, or service goals.
EP 2 – The organization gathers and analyzes the data generated through standardized monitoring, and the results are used to inform the goals and objectives of the individual's plan for care, treatment, or services as needed.
EP 3 – The organization evaluates the outcomes of care, treatment, or services provided to the population(s) it serves by aggregating and analyzing the data gathered through the standardized monitoring effort.

Compliance with this measure will impact an organization's accreditation status. Currently, TJC is requiring organizations that may not be in full compliance to provide a timeline and plan for compliance (adoption of MBC).

TJC also requires primary screening for suicide risk in patients admitted to a psychiatric hospital or those in general hospital settings who have are being treated or evaluated for a MH/SU condition. It also requires a suicide risk assessment for any patient who has screened positive for suicide risk and to document strategies to be used for risk mitigation if risk is identified.¹⁷

URAC, which accredits health plans, provider organizations, and mental health and substance use parity compliance, has recently released an accreditation standard that is available across all of their accreditation programs: Designation for Measurement-Based Care. Currently this measure is voluntary (as opposed to TJC), but provider organizations are encouraged to complete it, and employers and payors could be encouraged to consider it when partnering with health systems and providers.

Table 2. URAC Measurement-Based Health Care Designation Standards At-a-Glance¹⁸

1: Evidence-Based Self-Assessment The organization engages in timely evidence-based patient self-assessment at each clinical encounter.
<i>1-1: Self-Assessment Data</i> The self-assessment process includes:

- a. Gathering structured quantifiable data describing the patient's perceptions about psychiatric symptoms;
- b. Enabling the clinician to compare current symptom severity to past symptom severity;
- c. Informing the provider's evaluation of the clinical effectiveness of the current treatment; and
- d. Promoting accountability for treatment outcomes.

2: Symptom Rating Scale

The standardized symptom rating scale is designed to produce reliable symptom severity data.

2-1: Symptom Severity Data

The rating scale(s) in use are:

- a. Supplemental to clinical interviews;
- b. Current, interpretable, and readily available during the clinical encounter;
- c. Clinically actionable;
- d. Culturally validated in low-income and minority populations;
- e. Stored in electronic health records in such a way that it is easily extractable.

3: Classification of Symptom Severity

Changes in symptom severity are classified into clinically meaningful categories.

4: Treatment-to-Target

Guidelines are employed to enable development of individualized plans of care and enable identification of patients that achieve remission.

Purchasers are Recognizing the Importance of Measurement-Based Care

Employers, who are the predominant purchaser of healthcare (outside of the government), have also begun to advocate for greater use of MBC in MH/SU care delivery. The National Alliance of Healthcare Purchaser Coalitions (National Alliance) is a membership organization representing over 12,000 employers and 45 million Americans across the country. They provide expertise and resources to employers and employer coalitions on healthcare purchasing. Specifically, they have been focused on delivery of value-based care as a way to improve the health status of Americans. They conduct an ongoing survey of employer health plans, eValue8, which compiles information on attributes of health plans and benchmarks those benefits against evidence-based practices. The most recent eValue8 survey addressed MH/SU care coverage, including a review of network adequacy, access to care, and quality of care. Results related to quality of care revealed that most health plans were not engaged in MBC and were not routinely tracking outcomes.¹⁹

Among the many National Alliance recommendations to purchasers was that their health plans and providers should measure outcomes of care and use ongoing MBC to drive treatment

decisions. In addition, they highlighted the Collaborative Care Model as an approach to care of MH/SU conditions in primary care that requires MBC for reimbursement. The recommendations further stated that the Collaborative Care Model, which is paid for by Medicare and most commercial payers, should be covered by all health plans.

The work of the National Alliance highlights the important role that measurement can play in improving outcomes, as well as allowing providers to participate in value-based payment programs that tie reimbursement levels to delivering better outcomes.

Advocating for Measurement-Based Care for Mental Health and Substance Use Conditions

In November of 2019, the National Alliance, in partnership with American Psychiatric Association (APA), American Psychiatric Association Foundation Center for Workplace Mental Health, Bowman Family Foundation, and Meadows Mental Health Policy Institute, launched The Path Forward initiative to execute a disciplined, private sector approach to systematically and measurably improve five established best practices of mental health and substance use care.²⁰

The initiative is based on five priority strategies to positively transform MH/SU care at a population level, and to move the system forward in improving access to effective detection and treatment. The five priority strategies include: (1) improving network adequacy for MH/SU specialists, (2) expanding adoption of the Collaborative Care Model for delivering MH/SU care in primary care, (3) implementing MBC in both the MH/SU care and primary care systems, (4) expanding tele-behavioral health, and (5) ensuring mental health parity compliance.

The Path Forward proposes a market-driven implementation plan, leveraging the influence of employers and regional employer coalitions motivated for change, supported by the technical expertise and guidance of our nation's leading MH/SU care experts. It is centralized on clear and attainable process reforms and demonstrable outcomes, informed and empowered by engagement of key stakeholders at both the national and regional levels, combining nationwide efforts with a disciplined and intensive engagement focused on six regions most ready for change. The project will be assessed against both process and outcomes metrics, anchored by the Milliman Mental Health Substance Use (MHSU) Disparities Assessment, the National Alliance Mental Health Assessment, and the Bowman Family Foundation's Model Data Request Form for measuring disparities in access. Outcome goals include, but are not limited to, increased prevalence of MBC by adoption by at least 40% of patient centered medical homes and accountable care organizations, including all of the largest health systems in each region, and substantial adoption of the Collaborative Care Model, including 50% of primary care practices in the largest health systems in each region.²¹

Recommendations by National Alliance specifically for increasing the prevalence of MBC by adoption is that health plans, at the request of employers and employer coalitions, provide an

action plan requiring providers to use standardized measurement tools. Additionally, health plans should be expected to require that enrollees be screened for mental health and substance use disorder conditions, as well as reporting treatment outcomes. Finally, health plans and MH/SU organizations should provide incentive payments and minimize administrative requirements to in network providers (primary care, mental health, and substance use) who participate in quality improvement programs by integrating and requiring measurement-based care clinical practices.²²

Moving Measurement-Based Care Forward: A Call to Action

The movement toward value-based care has created important opportunities to improve the quality of healthcare and to incentivize use of evidence-based practices. It also provides an opportunity to advocate for the inclusion of evidence-based practices into those incentives. Use of MBC is a critical driver of improved care, across all MH/SU conditions, and therefore should be leveraged when paying for MH/SU care in all settings, and when accrediting providers, health plans, hospitals and health systems treat patients with MH/SU conditions.

Quality measures that are used to determine performance payments by commercial and government payers currently contain few measures requiring MBC. While CMS has prioritized development of outcome measures, specifically for treatment of MH/SU conditions, development of measures is costly and takes several years. In addition, the regulatory hurdles related to adoption of these measures are difficult and have often directly blocked the use of measures relying on MBC. Although the APA is currently engaged in developing a number of MBC and outcome-based measures that will likely overcome those hurdles, not all providers will be required to report on those MBC measures.

While quality measurement is a critical driver of practice patterns and has been shown to improve outcomes when certain practices are measured (e.g., screening for depression and achieving diabetes control), accreditation standards may provide a better way to assure use of MBC. TJC and URAC standards that include MBC as a component necessary to receive accreditation in MH/SU settings is an important example of how accreditation standards can be leveraged. However, these are limited programs, and therefore do not apply to the majority of care delivered to patients with MH/SU disorders, whether in primary care or in other specialty MH/SU care settings. Expansion of these accreditation standards to other programs both within TJC and URAC and across all other accrediting organizations would have a significant impact on creating a culture of MBC, and creating the expectation that all care delivered to patients with MH/SU conditions had the benefit of MBC. Such standards should apply across all care settings, including primary and MH/SU care and inpatient and outpatient care. Table 3 provides some examples of accrediting organizations and the care settings that they review.

Table 3. Accreditation Programs

Organization	Programs Accredited
NCQA ²³	Health plans, patient-centered medical home (PCMH), patient-centered specialty practice (PCSP), patient-centered connected care, MH/SU plans, and managed behavioral health organizations (MBHO)
URAC ²⁴	Health plans, healthcare management, healthcare operations, provider integration and coordination, mental health and substance use disorder parity, and MBC
Commission on Accreditation of Rehabilitation Facilities (CARF) ²⁵	MH/SU care, child and youth services, employment and community services, and opioid treatment programs
TJC ²⁶	MH/SU care, critical access hospitals, home care, hospitals, and nursing care centers

The Path Forward has been actively engaged with these accrediting organizations, communicating the potential impact on improvement in MH/SU care and working to establish MBC as a standard of care.

Overview of Methods, Recommended Measures, and Measurement-Based Care in Real World Settings

This paper summarizes longstanding and evolving data which support use of a robust set of validated patient-reported outcome measures (PROMs) that can be used for measurement-based care (MBC) of mental health and substance use (MH/SU) conditions treated in primary and specialty care settings, and it explicates the evidence behind their selection.

We identified measures that not only have demonstrated their positive psychometric properties, but also have either been used as outcome measures in real-world examples of MBC or have shown in research studies their sensitivity to clinical change over time. As such, the measures described in this paper can be utilized as tools for MBC and for reporting in accreditation and quality measurement programs.

Methodology

The Meadows Mental Health Policy Institute (MMHPI) explored validated patient-reported MH/SU measures across peer-reviewed academic and other authoritative literature (government, technical, and professional reports) to identify tools that are reliable and feasible to use as PROMs for MBC. In selecting measures that meet these criteria, we additionally reviewed the Kennedy Forum’s report and selected those measures that would meet the criteria we had defined. Specifically, the primary goal of this exploration was to identify PROMs that are (a) psychometrically validated, (b) sensitive to clinical change over time, and (c) feasible for implementing in primary and/or MH/SU specialty care practices.

Exploration of relevant information for clinical instruments was performed independently by six MMHPI reviewers, most of whom either reviewed adult or pediatric measures. Disagreements were resolved through discussion.

Within each section of age-related measures, and for each individual measure, we cite the evidence that it can be used as a repeated measure of outcomes, examples of its use, and, where notable, implementation issues for consideration. We sometimes also note other potential measures of the same MH/SU condition that could be considered for use.

Methodological Considerations

The recommended outcome measures in this report include tools that have demonstrated sensitivity to change over time. This may have been demonstrated in either of two ways: (1) the measure was validated through a clinical outcome measure study, included as an outcome measure within a peer-reviewed research study that shows sensitivity to change over time, or (2) the measure has been used in clinical practice as an outcome measure and found to be useful for making clinical decisions. Some measures have achieved both types of validation.

As noted in the above criteria, in recommending validated rating scales for use in MBC, we did take into consideration the heavy demands and the vast array of current burdens required of healthcare providers. In other words, we prioritized measures that had minimal administrative burden to patients and clinical staff. We recognize that implementing MBC can add new demands to clinical practices and so the MH/SU measures recommended below consider the length and feasibility of use for each measure. Rating scales included in this report can and should also be used in both the MH/SU specialty and the primary care setting.

Although we considered whether the tools had been recommended by other well-recognized professional organization and agencies, and whether there was evidence of the tool's use for MBC in non-research-based "real-life" practice settings, we did not necessarily exclude tools that had not received endorsements or for which we could not find examples of their ongoing use in clinical settings. Measures *were* excluded if there was no evidence for their sensitivity/responsiveness to change over time or if there was evidence that the tool was too lengthy or burdensome for use by patients or clinical providers.

For all these reasons, then, we emphasize tools that not only have demonstrated utility as outcome measures in real world settings or in research but also are feasible for use (are not too lengthy, too costly, or otherwise overly burdensome to use).

Since more attention in MH/SU care has been paid to services for adults, it is not surprising that MH/SU outcome measures for adults are better-developed and more widely used than

We emphasize tools that not only *have demonstrated utility as outcome measures in real world settings, but also are feasible for use.*

comparable measures for children and youth. As a result, the evidence for use of some of the child and youth outcome measures may be less robust than those that undergird our recommendations concerning adult measures. However, in areas where no measures were available, we used our expert judgment (rooted in a thorough literature review, as well as consulting and clinical experience with integrated care programs) to recommend pediatric measures that we believe clinicians will find are useful for MBC. In most cases, such measures are currently being used in several health systems.

Additionally, as the field of patient reported outcome measurement in MBC evolves in MH/SU care, new or emerging tools may demonstrate superiority to some of the tools described within this summary. Similar to the tracking of best practices, stakeholders should remain attuned to “best” tools for practicing MBC.

Our recommendations are predominately focused on condition-specific instruments because of the body of research that suggests they may be more responsive to change over time than other types of measures.²⁷ This is likely due to their close association with the diagnostic criteria and symptomology of the specific conditions themselves. However, one important exception to this relates to the measurement of suicidality. Risk for suicide is present in a number of MH/SU conditions and should be assessed in addition to other symptoms associated with a particular disease state. Also, as suicide risk is either present or absent, tools that are effective for screening and assessing risk are also by definition valid as “outcome” measures since they are designed to assess level and degree of risk.

Most of the PROMs we have recommended are already commonly, if not widely, used in primary care and specialty practices, either as screening tools or as outcome assessment tools, or both. All *can* be used for both purposes. Therefore, the next step for many practices will be to connect the instrument to clinical monitoring practices with repeated instrument administrations across the treatment phases respective to the condition.

Recommended Adult Measures

The table below provides an overview of adult mental health and substance use (MH/SU) tools we recommend for use in measurement-based care (MBC). The name of the instrument, the condition(s) for which it represents an outcome measure, and a brief description of the instrument are all included in the table. Please note that neither the adult measures nor the pediatric measures represent exclusive lists of measures that can be used for MBC. Our point

was not to list every conceivable measure, but to identify a comprehensive set that could serve as a starting point for decision-makers. Thus, the measures listed in the table do not necessarily constitute all measures that are recommended for use in MBC; other measures may also have demonstrated their validity and utility. In Appendix 1, we provide more detail concerning utilization of the outcome measures, as well as real-world uses that indicate the feasibility of implementing the outcome measures, when such examples are available.

Table 4. Recommended Adult Measures

Instrument	Condition/Type	Brief Description
Patient Health Questionnaire-9 (PHQ-9)*	Depression	Number of items: nine Dimension(s) measured: Depression severity based on diagnostic symptoms derived from the Diagnostic and Statistical Manual of Mental Disorders (DSM-5).
PROMIS Depression	Depression	Number of items: four, six, or eight (short forms); four to 12 items using computer-adapted test. Dimension(s) measured: Negative mood, views of self and social cognition, decreased positive affect, and engagement.
General Anxiety Disorder-7 (GAD-7)*	Anxiety	Number of items: seven Dimension(s) measured: Anxiety severity based on diagnostic symptoms derived from the DSM-5.
PROMIS Anxiety	Anxiety	Number of items: four, six, seven, or eight (short forms); four to 12 items using computer-adapted test. Dimension(s) measured: Fear, anxious misery, hyperarousal, and somatic symptoms related to arousal.
Panic Disorder Severity Scale – Self Report (PDSS-SR)*	Panic attacks	Number of items: seven Dimension(s) measured: Frequency of panic attacks, distress during panic attacks, anticipatory anxiety, agoraphobic fear and avoidance, interoceptive fear and avoidance, impairment of or interference in work functioning, and impairment of or interference in social functioning.
PROMIS Alcohol	Alcohol use disorders	Number of items: seven (short form); four to 12 items using computer-adapted test. Dimension(s) measured: Drinking patterns, cue-based drinking, cravings to drink, and efforts to control drinking that indicate problematic drinking.

Instrument	Condition/Type	Brief Description
US-Alcohol Use Disorders Identification Test-Consumption (USAUDIT-C)*	Alcohol use disorders	Number of items: three (short screen) and 10 (full measure) Dimension(s) measured: Drinking frequency and quantity.
Brief Addiction Monitor (BAM-R)*	Substance use disorders	Number of items: 17 Dimension(s) measured: Type and frequency of substance use, indicators of relapse risk (risk factors), and recovery-oriented behaviors (protective factors).
Substance Abuse Outcomes Module*	Substance abuse	Number of items: 22 Dimension(s) measured: Patient characteristics, including diagnosis and prognosis, patient outcomes, and process of care.
Post-Traumatic Stress Disorder (PTSD) Checklist (PCL)*	Trauma	Number of items: 17 Dimension(s) measured: PTSD symptoms based on the DSM.
Columbia–Suicide Severity Rating Scale (C-SSRS)	Suicide	Number of items: 17 Dimension(s) measured: Suicidal ideation, intensity of ideation, and suicidal behavior.
Ask Suicide Screening Questions (ASQ)	Suicide	Number of items: four Dimension(s) measured: Acute suicidal ideation and intent.
Brief Pain Inventory (BPI)	Pain	Number of items: 11 Dimension(s) measured: Pain severity and pain-related interference.
Positive and Negative Syndrome Scale-6 (PANSS-6)	Psychosis	Number of items: six Dimension(s) measured: Symptoms of psychosis: delusions, conceptual disorganization, hallucinations, blunted affect, passive/apathetic, social withdrawal, and lack of spontaneity and flow of conversation.
Brief Psychiatric Rating Scale (BPRS)	Psychiatric severity	Number of items: 24 Dimension(s) measured: Mood disturbance, reality distortion, activation, apathy disorganization, and somatization.
Altman Self-Rated Mania Scale (ASRM)*	Mania	Number of items: five Dimension(s) measured: Elevated mood, increased self-esteem, decreased need for sleep, pressured speech, and psychomotor agitation.

Instrument	Condition/Type	Brief Description
Eating Disorder Examination – Questionnaire Short (EDE-QS)	Eating disorder pathology	Number of items: 12 Dimension(s) measured: Concerns about dietary restraint, eating, weight and shape.
Eating Attitudes Test (EAT-26)	Eating disorder pathology	Number of items: 26 Dimensions measured: Dieting, bulimia and food preoccupation, and oral control.
Florida Obsessive-Compulsive Inventory (FOCI) ²⁸ C-FOCI (child version) ²⁹	Obsessive compulsive symptomology	Number of items: 20 Dimensions measured: Used to assess presence of obsessions and compulsions; additional five item severity scale if so needed.
Edinburgh Post Natal Depression Screen	Maternal depression	Number of items: 10 Dimension(s) measured: Frequency of depressive symptoms and indicators of positive emotions.
Medical Outcomes Study Short-Form Health Survey (SF-12)*	Health-related quality of life/functional status	Number of items: 12 Dimension(s) measured: Physical functioning, role limitations resulting from physical health problems, bodily pain, general health, vitality (energy/fatigue), social functioning, role limitation resulting from emotional problems, and mental health.
World Health Organization Disability Assessment Schedule (WHODAS II)	Functional status	Number of items: 12 and 36-item Dimension(s) measured: Cognition (understanding and communicating); mobility; self-care (activities of daily living); getting along (interacting with other people); life activities (domestic responsibilities, leisure, work, and school); and participation (joining in community activities, participating in society).

*Kennedy Forum recommended measure.

Recommended Pediatric Measures

The table below summarizes the pediatric mental health and substance use (MH/SU) measures that have either been implemented as MH/SU measurement-based care (MBC) tools or are good candidates to be used in MBC. Descriptions of the measures, endorsements of them (where applicable), and examples of their use are further detailed in Appendix 2.

Table 5. Recommended Pediatric Measures

Instrument	Condition	Brief Description
Patient Health Questionnaire for Adolescents (PHQ-A)*	Depression	Number of items: nine Dimension(s) measured: Anxiety severity based on diagnostic symptoms derived from the DSM.
PROMIS Depression	Depression	Number of items: six (parent-reported), eight (adolescent-reported), four to 12 items using computer-adapted test. Dimension(s) measured: Negative mood, views of self, and social cognition, as well as decreased positive affect and engagement.
Suicide Behavior Questionnaire-Revised (SBQ-R)	Suicide risk	Number of items: four, self-report Dimension(s) measured: Lifetime and current suicidal ideation and history of actual events.
Vanderbilt Attention Deficit Hyperactivity Disorder (ADHD) Rating Scale*	ADHD	Number of items: 55 – parent; teacher – 43 Dimension(s) measured: ADHD symptoms, as well as other symptoms related to other conditions like oppositional defiant disorder (ODD), conduct disorder, anxiety, depression, and learning disorders.
Pediatric Symptom Checklist*	Psychosocial functioning	Number of items: 17 and 35 Dimension(s) measured: Behavioral health-related health and functioning, including aspects of attention and symptoms of internalizing and externalizing problems. ³⁰
Screen for Child Anxiety Related Emotional Disorders (SCARED)*	Anxiety disorders	Number of items: 41 Dimension(s) measured: Symptoms of generalized anxiety disorder in addition to several specific phobias, including separation anxiety disorder, panic disorder, social phobia, and school-related phobia.
PROMIS Anxiety	Anxiety disorders	Number of items: eight (parent-reported), eight (adolescent-reported), four to 12 items using computer-adapted test Dimension(s) measured: Fear, anxious misery, hyperarousal, and somatic symptoms related to arousal.
Mood and Feelings Questionnaire (MFQ)*	Depression, dysthymia	Number of items: long form – 33; short form – 13 Dimension(s) measured: symptoms of depression based on DSM symptom criteria

Instrument	Condition	Brief Description
Brief Addiction Monitor (BAM)	Substance use disorders	Number of items: 17 Dimension(s) measured: Type and frequency of substance use, in addition to indicators of relapse risk (risk factors) and recovery-oriented behaviors (protective factors).
PROMIS Anger	Anger	Number of items: Parent-reported measures were not available; eight (adolescent-reported), four to 12 items using computer-adapted test Dimension(s) measured: Type and frequency of substance use, in addition to indicators of relapse risk (risk factors) and recovery-oriented behaviors (protective factors).
Altman Self-Rated Mania Scale (ASRM)	Mania	Number of items: five Dimension(s) measured: Elevated mood, increased self-esteem, decreased need for sleep, pressured speech, and psychomotor agitation.
ChEAT (Children’s version of Eating Attitudes Test-26)	Eating Disorder Pathology	Number of items: 26 Dimensions measured: Dieting, bulimia and food preoccupation, and oral control.
Children’s Florida Obsessive-Compulsive Inventory (C-FOCI)	Obsessive Compulsive Symptomology	Number of items: 17 Dimension(s) measured: Time occupied, interference, distress, degree of avoidance, and degree of control.

*Kennedy Forum recommended measures.

Multi-Condition or Cross-Cutting Tools Used in Adults and Children

In recent years, software platforms have been developed to provide screening and monitoring capabilities across a number of symptom domains and mental health and substance use (MH/SU) conditions. These platforms contain measures described in this paper as well as others, treatment planning facility, and treatment decision support and can be integrated into existing electronic health records.

Table 6. Proprietary Tools

Multi-Condition Screening and Outcome Measurement (Proprietary) Tools	
Tools	Link/Reference
Behavioral Health Checkup	http://dpbh.ucla.edu/behavioral-health-checkup

Multi-Condition Screening and Outcome Measurement (Proprietary) Tools	
Tools	Link/Reference
VitalSign6	https://www.utsouthwestern.edu/education/medical-school/departments/psychiatry/research/center/vital-sign6/
Behavioral Health-Works (BH-Works)	https://mdlogix.com/bhworks-page/
Tridium ONE	https://tridium.com/capabilities/tridium-one/
M-3 Checklist	Whatsmym3.com (public domain for individual use) m3information.com
Behavioral Health Lad (BHL)	http://www.capitolsolutiondesign.com/
Outcome Questionnaire (OQ-45.2) and Outcome Rating Scale (ORS)	http://www.oqmeasures.com/oq-45-2/
Total Brain	https://www.totalbrain.com/

Health System Utilization of Measurement-Based Care for Mental Health and Substance Use Care

A number of health systems have begun to adopt MBC programs, which have established an expectation that providers of mental health and substance use (MH/SU) care should utilize MBC to manage conditions, just as other health providers rely on MBC to deliver routine care. For illustration, we have selected case examples of providers or health systems delivering the Collaborative Care Model. Collaborative Care, an evidence-based model of care delivered in the primary care setting to patients with any MH/SU condition, is reimbursed by public and private payers and requires the use of repeated measures to assess improvement in the MH/SU condition being treated.

Table 7. Measurement-based Care Utilization Examples

Case Examples of MBC Utilization				
Case Example	Population	Disorders	Tool	Reference
The University of California – Los Angeles (UCLA) Behavioral Health Associates	Primary care patients	Psychiatric disorders	Behavioral Health Checkup	UCLA Behavioral Health Associates ³¹
JPS Health System	Primary care patients, MH/SU patients	Depression, suicide risk	Patient Health Questionnaire-9 (PHQ-9) with utilization of Institute for Healthcare Improvement Model (Plan-Do-Study-Act)	Agovi, A. M. A., Anikpo, I., Cvitanovich, M. J., Yan, L., Chu, T. C., MacDonald, B., ... & Ojha, R. P. (2019). Sociodemographic and Health-Related Characteristics of a Safety-Net Patient Population.
Department of Veteran Affairs, nationwide	Primary care patients	Depression, panic, generalized anxiety, PTSD, alcohol misuse	Behavioral Health Laboratory (BHL)	The Kennedy Forum (2015) ³² ; A tipping point for measurement-based care (2017) ³³
Department of Defense (Army Branch), nationwide	Specialty mental health patients	Depression, panic, generalized anxiety, PTSD, bipolar disorder, alcohol misuse	Behavioral Health Data Portal (BHDP)	The Kennedy Forum (2015) ³⁴ ; A tipping point for measurement-based care (2017) ³⁵
Sheppard Pratt	Community	Mental health, special education, substance use, developmental disability, and social services	Track responses from validated questionnaires based on diagnosis and treatment needs, including PROMIS, WHODAS, and PHQ9	Dr. Harsh Trivedi, President and CEO, Sheppard Pratt (permission granted) www.sheppardpratt.org
Penn State Psychiatry	Youth	Mental health treatment	PCARES-Youth	Developing Measurement-Based Care for Youth in an Outpatient Psychiatry Clinic: The Penn State Psychiatry Clinical Assessment and Rating Evaluation System for Youth (PCARES-Youth) (2020) ³⁶

Case Examples of MBC Utilization				
Case Example	Population	Disorders	Tool	Reference
Collaborative Care Implementation: University of Pennsylvania	Adult and youth	Primary Care: program includes not only management of depression and anxiety, but other MH/SU conditions and may include repeated measures	PHQ9, GAD7 and others	Livesey, C., & Wolk, C. B. (2019). Innovative Program Provides MH Care to Thousands. Psychiatric News. https://doi.org/10.1176/appi.pn.2020.1a11
Cohen Veterans Network	Veterans and families, mental health patients	Variety of mental health issues including depression, anxiety, post-traumatic stress, adjustment issues, anger, grief and loss, family issues, transition challenges, relationship problems, and children's behavioral problems	PCL-5, PHQ-9, GAD-7, QLES	Communication with leadership of Cohen Veterans Network. Also found at https://www.cohenveteransnetwork.org/about-us/getthefacts/

Appendix 1: Recommended Adult Measures

The following section provides further detail of each adult measure. We did not attempt to incorporate every known piece of evidence that could support a measure's use or every example of its use in real-world settings. For example, although we reviewed significant evidence for the Patient Health Questionnaire's (PHQ's) use as a measurement-based care (MBC) measure, we did not attempt to summarize all available evidence. That would have been beyond the scope of this paper. Rather, we are attempting to make payers, providers, and other stakeholders aware of a variety of suitable measures for MBC, some of which are not as well-known as the PHQ.

Indeed, some of the measures we recommend are not necessarily widely used in the real world as MBC measures, at least not as far as we know. However, we have included several such measures because they are brief (meeting our feasibility criterion), have successfully been used in research studies as valid outcome measures, and in our judgment could easily be appropriated in real-world settings.

Tool: Patient Health Questionnaire – 9 Items (PHQ-9)

The PHQ has several versions focused on depression,³⁷ including the brief PHQ, PHQ-8,³⁸ PHQ-4, and the PHQ-2. The PHQ-9 has been validated as both a diagnostic tool,^{39, 40, 41, 42} and as depression monitoring/management tool because of its responsiveness to measure depression severity over time.^{43, 44, 45, 46, 47, 48} The PHQ-9 is also required for reporting the only outcome-based mental health and substance use (MH/SU) quality measure used in federal accountability programs (NQF 710, 711). This measure is also included in the Healthcare Effectiveness Data and Information Set (HEDIS) measures that are often used for Medicaid and Medicare Advantage plans and other health plans receiving accreditation by the National Committee on Quality Assurance (NCQA). The Minnesota Statewide Quality Reporting and Measurement System tracks depression remission rates by requiring all providers to report data on this measure. Other organizations that recommend the PHQ-9 as a monitoring and outcome measure for depression include the Institute for Clinical Systems Improvement (ICSI),⁴⁹ the International Consortium for Health Outcome Measurement (ICHOM),^{50, 51} the Substance Abuse and Mental Health Services Administration (SAMHSA), and the Interagency Task Force on Military and Veterans Mental Health (ITF), which is a collaboration between the United States Department of Defense (DoD) and Department of Veterans Affairs (VA).⁵² Furthermore, the United States Preventative Services Task Force has recommended depression screening and use of the PHQ.

The PHQ-2, the other most commonly used PHQ variation for depression, is a brief two-item version that has been validated as a depression screener, and it is typically used within a "two-step" process in concert with the PHQ-9.^{53, 54, 55} In instances when patients score high on the

PHQ-2 (a score of two or three or higher),⁵⁶ the clinician will subsequently administer the PHQ-9 to measure depression severity.

The PHQ-2 and 9 have also been used or required in a number of additional settings including:

- the VA's patient-centered medical home (PCMH) and evidence-based programs,
- the DoD for pre- and post-deployment mental health assessments and monitoring,⁵⁷ and
- most Collaborative Care implementations. The University of Washington AIMS Center (www.uwaims.edu) describes how it can be used in the required case review registry.

Tool: PROMIS Depression

The PROMIS Depression scale is a validated instrument that measures negative mood, views of self and social cognition, decreased positive affect, and engagement. The PROMIS Depression scale has demonstrated responsiveness to treatment in clinical research.^{58, 59, 60, 61, 62, 63} For a brief review of PROMIS instruments and their instrument development process, see Appendix 3.

Tool: Generalized Anxiety Disorder – 7 Items (GAD-7)

The GAD-7 is one of the most common and widely used measures for anxiety. The GAD-7 has also demonstrated validity to assess specific anxiety conditions, including post-traumatic stress disorder (PTSD), social anxiety disorder (SAD), and panic disorder (PD).^{64, 65} Several studies have used the GAD-7 as an outcome measure for testing the difference between clinical interventions. It is also responsive to change over time.^{66, 67, 68, 69, 70} The GAD-2 is a brief two-item version that has been validated as an anxiety screener and can be used within a “two-step” process in concert with the GAD-7.⁷¹ The GAD-7 has been recommended by the ICHOM⁷² and by the ITF.⁷³ Like the PHQ-9, it is also used in MBC programs in the VA, DoD and most Collaborative Care programs.

Tool: PROMIS Anxiety

The PROMIS Anxiety scale is a validated instrument that measures fear, anxious misery, hyperarousal, and somatic symptoms related to arousal. The PROMIS Anxiety scale has demonstrated responsiveness to treatment in clinical research.^{74, 75, 76, 77, 78, 79} For a brief review of PROMIS instruments and their instrument development process, see Appendix 3.

Tool: Panic Disorder Severity Scale – Self Report (PDSS-SR)

The PDSS-SR is a reliable and valid instrument that measures panic disorder severity. The tool is known to be useful in clinical and research settings with the promise of becoming a standard global rating scale for panic disorder. The tool is most appropriate for rating severity and treatment progress in patients with established diagnoses of panic disorder. It is modeled after

the Yale-Brown Obsessive-Compulsive Scale and contains items that assess the severity of the seven dimensions of panic disorder and associated symptoms.^{80, 81, 82}

Tool: PROMIS Alcohol Use

The PROMIS Alcohol Use scale is a validated instrument that measures drinking patterns, cue-based drinking, cravings to drink, and efforts to control drinking that indicate problematic drinking. The PROMIS Alcohol Use scale has demonstrated responsiveness to treatment in clinical research.^{83, 84} For a brief review of PROMIS instruments and their instrument development process, see Appendix 3.⁸⁵

Tool: US-Adapted Alcohol Use Disorders Identification Test-Consumption (USAUDIT-C)

The Alcohol Use Disorders Identification Test-Consumption (AUDIT-C) is the most widely implemented *screening* instrument for harmful alcohol use; however, it does not have adequate psychometric evidence to strongly support its use as a monitoring instrument. A version of the AUDIT-C which has shown evidence of responsiveness to change is the USAUDIT-C. The USAUDIT-C shares the same items with the AUDIT-C, but uses a slight variation in the wording of the last item in the three-item screening version, as well as adjusted response options across all items.^{86, 87} While the Brief Addiction Monitor might be a superior tool (in part, because it incorporates both alcohol and illicit drugs – see below), the USAUDIT-C is a viable option for MBC of alcohol disorders. A provider could screen with the three-item version and then use the 10-item version for MBC of those patients who screened positive and entered treatment. (A score of seven for women and a score of eight for men indicates a positive screen.) The 10-item version has demonstrated sensitivity in measuring responsiveness to treatment.⁸⁸

The AUDIT-C is one of many self-reported screening and outcomes tools included in the University of California, Los Angeles (UCLA) Behavioral Health Checkup Platform (BHC).⁸⁹ While we recommend using the USAUDIT-C, the fact that the AUDIT-C has been used for MBC speaks to the likely real-world utility of the AUDIT-C.

Tool: Brief Addiction Monitor (BAM)

The BAM is a validated measurement tool that has undergone testing for responsiveness in measuring substance use severity and it has also been used in a clinical research study to examine response to treatment.⁹⁰ The BAM has an advantage over most alcohol and drug abuse/dependence screens because of its multi-factorial structure. Most other self-reported alcohol and drug screening tools focus on consumption levels or abstinence, but the BAM also includes other aspects of health and mental health that are often affected by substance abuse and dependence and therefore yields additional, clinically useful information that can inform treatment.^{91, 92}

In 2016, the VA implemented the Measurement-Based Care in Mental Health Initiative. Among the identified measurements tools utilized in this initiative, the BAM-R is the designated tool for monitoring substance use symptoms as well as risk and recovery factors.⁹³ The BAM-R adopts a continuous item response format, which is the preferred format for use in practices, as the construct validity is not consistent across response formats.⁹⁴

Tool: Substance Abuse Outcomes Module

The Substance Abuse Outcomes Module is considered a routine outcomes assessment that examines a patient's characteristics, processes of care, and patient outcomes for substance abuse treatment. The tool has both clinician and patient baseline assessment forms, with patient follow-up assessment completed three- and six-months following baseline. It uses a "tracer condition" approach where a single disorder is closely examined in a given treatment setting. Therefore, it is most appropriate to use in a substance treatment setting and/or with patients who have primary substance problems in general healthcare settings. It has appropriate reliability and validity for use in routine substance treatment settings.⁹⁵

Tool: PTSD Checklist (PCL)

The PCL is a validated instrument of PTSD severity that has demonstrated responsiveness for measuring PTSD symptoms as described in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5).^{96, 97} Dissemination and utilization of the PCL is recommended by the VA's National Center for PTSD,⁹⁸ ITF,⁹⁹ and the DoD.¹⁰⁰

Tool: Columbia-Suicide Severity Rating Scale (C-SSRS)

The Columbia-Suicide Severity Rating Scale is a validated instrument for measuring suicide risk severity as described in the DSM.^{101, 102, 103} The Joint Commission has also include the C-SSRS in its set of recommended measures that can be used to meet the NPSG 15.01.01.¹⁰⁴ The US Food and Drug Administration (FDA) includes the C-SSRS as the "gold standard" measure for tracking suicidal ideation and behavior over time in treatment studies seeking to decrease suicide risk among patients.^{105, 106}

Tool: Ask Suicide Screening Questions (ASQ)

The ASQ was developed and validated by the National Institute of Mental Health (NIMH). It includes a four-item screening tool that can be used for all patients in primary care settings. It is recommended by the NIMH, the National Action Alliance for Suicide Prevention, the Zero Suicide Initiative and The Joint Commission. The ASQ has established validity in adult patients¹⁰⁷ and demonstrates high sensitivity, specificity, and negative predictive value with pediatric populations.¹⁰⁸

Tool: Brief Pain Inventory (BPI)

The BPI short form, an instrument that includes 11 numerically scored items from zero to 10, is a validated outcome measure for pain¹⁰⁹ across dimensions of severity and interference. Additionally, the BPI has demonstrated superior responsiveness to change over time compared to other outcome measures of pain, and may satisfy most criteria for FDA guidance for PRO measures.^{110, 111, 112, 113} The NQF endorses process measures that track pain assessment and follow-up, and it specified the BPI as one of the standardized pain measurement tools.^{114, 115}

Tool: Positive and Negative Syndrome Scale (PANSS-6)

The PANSS-6 is a frequently used validated clinician-rated instrument for measuring symptoms of psychosis, which has also demonstrated responsiveness to treatment.^{116, 117, 118, 119, 120}

We were not able to find examples of the PANSS-6 being used for MBC, only for its utility in research studies. However, like the C-SSRS above as a measure of suicidality, the PANSS-6 is the FDA's "gold standard" for anti-psychotic medication trials and we believe that it could and should be used for MBC.

Tool: Brief Psychiatric Rating Scale (BPRS)

The BPRS is a validated clinician-rated instrument for measuring psychosis severity in clinical research that has demonstrated sensitivity to treatment response.^{121, 122, 123} Moreover, the BPRS has been shown to be a sensitive measure of psychiatric severity among adults with unipolar depression in outpatient settings.¹²⁴ The BPRS is a longer, more comprehensive instrument than the PANSS-6 and, although it has the advantages of addressing a wider array of symptoms and yielding a wider range of scores, some practitioners might be put off by its length. However, the validity and responsiveness of a six-item version of the BPRS is under investigation and soon will become available for use in MBC.¹²⁵

Tool: Altman Self-Rated Mania Scale (ASRM)

The ASRM is a validated measure of mania-related symptoms and has also demonstrated sensitivity to treatment response.¹²⁶ The American Psychiatric Association (APA) designated the ASRM as an "emerging measure" for further research and evaluation for initial assessment and monitoring of treatment progress.¹²⁷

Tool: Eating Disorder Examination – Questionnaire Short (EDE-QS)

The EDE-QS is a brief, reliable and valid measure of eating disorder symptom severity that performs similarly to the measure from which it was derived, the EDE-Q, a 28-item tool. The EDE-QS has shown excellent internal consistency, test-retest reliability, and convergent validity with the long version, both for people with and without an eating disorder. Because of its conciseness and revised response categories, the EDE-QS can be used as a weekly measure permitting ongoing progress monitoring.¹²⁸

Tool: Eating Attitudes Test (EAT-26)

The EAT-26 is likely the most widely used screening tool for identification of symptoms and concerns characteristic of eating disorders. It is highly reliable and valid and was designed as a screening tool to be used with at-risk populations (as well as non-clinical populations).^{129, 130, 131, 132} The EAT-26 is an instrument to help identify individuals who might be at risk for an eating disorder – not for diagnostic purposes.¹³³ It can be used to screen eating disturbances in general as well as to measure change over time. The EAT-26 can be used in both clinical and non-clinical settings, with both adolescents and adults. A version of this instrument (ch-EAT) was developed for children ages eight to 13.¹³⁴

Tool: Florida Obsessive-Compulsive Inventory (FOCI)

The FOCI is a symptom checklist that consists of 25 items, which includes 20 commonly occurring obsessions and compulsions derived from the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) and a five-item severity scale that captures symptom severity and impairment of functioning over the past month. The severity scale examines five dimensions of severity, including time occupied, interference, distress, degree of avoidance, and degree of control. The FOCI demonstrate strong psychometric properties, although further research is warranted to confirm results. While there is limited availability of psychometric data compared to other obsessive-compulsive disorder (OCD) questionnaires, the severity scale corresponds strongly with clinician's ratings of OCD severity, suggesting that this measure can serve as a useful parent-report assessment of symptom severity. In addition, the FOCI has shown to have treatment sensitivity to cognitive behavioral therapy (CBT) and therefore could be a useful tool for outcoming monitoring.^{135, 136, 137} A version of this instrument has been developed for children (C-FOCI).

Tool: Edinburgh Postnatal Depression Scale (EPDS)

The EPDS is a widely used, validated depression screen.^{138, 139, 140} Perinatal depression, which includes depression experienced before, during, and up to one year after pregnancy, can have a negative impact on the family system. However, treating maternal depression until remission can improve psychosocial outcomes in children.^{141, 142} The EPDS has been used in multiple studies to measure symptoms of depression during pregnancy and after childbirth over time.^{143, 144, 145}

The National Academy for State Health Policy (NASHP) has recently shown that several states, including Massachusetts, have recommended the EPDS as a screening and outcome measurement tool.¹⁴⁶ EPDS has been used in a population-based model of integrated care in Massachusetts that supports pediatric and maternal outcomes (Massachusetts Child Psychiatry Access Program [MCPAP] for Moms).^{147, 148, 149}

Tool: Medical Outcomes Study Short-Form Health Survey (SF-12)

The SF-12 is one of the most widely used, validated, and responsive instruments for measuring the health-related quality of life across many medical conditions in research studies.^{150, 151, 152, 153, 154} Additionally, the SF-12 has been validated with special populations, including veterans and people living with serious mental health conditions.^{155, 156}

The SF-12 has been used to demonstrate the effectiveness of collaborative care in treating depression in a primary care setting.¹⁵⁷ Moreover, the SF-12 has utility to support multiple population health management practice efforts. For instance, the SF-12 can be used as depression screener¹⁵⁸ and it has demonstrated effectiveness in predicting medical expenditures.¹⁵⁹

The SF-12 has been endorsed by the European Medical Agency for testing pharmaceutical label claims of improved quality of life.¹⁶⁰

The SF-12 takes approximately ten minutes for patients to complete. Clinical practices and provider systems should consider their target population and their orientation to clinical practice when selecting a version of SF-12 and its corresponding scoring system. Clinicians and administrators should consult the SF-36 User Manual for selecting the most appropriate SF version for their practice.¹⁶¹

Tool: World Health Organization Disability Assessment Schedule (WHODAS II)

The WHODAS II is a validated instrument that has demonstrated sensitivity to treatment response across a broad spectrum of health conditions, cultural contexts, and geographic regions.^{162, 163, 164, 165, 166, 167} The International Consortium for Health Outcome Measurement (ICHOM) endorses the WHODAS II as one of the MBC instruments for measuring the level of physical and social functioning among patients with depression or anxiety.¹⁶⁸ Similarly, the Kennedy Forum recommends the WHODAS II for measuring level of functioning in patients with a wider variety of general medical and mental health conditions.¹⁶⁹ The APA has designated the WHODAS II as an “emerging measure” for further research and evaluation for monitoring treatment progress.¹⁷⁰ Given its demonstrated validity and brief nature, however, we believe it could be used for MBC.

Appendix 2: Recommended Pediatric Measures

The following section provides further detail of each pediatric measure.

Tool: Patient Health Questionnaire for Adolescents (PHQ-A)

The PHQ, modified for adolescents (PHQ-A), is a validated instrument for measuring depression severity,^{171, 172, 173, 174} and has been used in clinical studies to demonstrate the effectiveness of collaborative care in treating depression.^{175, 176}

The PHQ-A has minor changes from the original Patient Health Questionnaire-9 (PHQ-9) that incorporate characteristics of depression among adolescents and add age-appropriate language, including items related to irritability, weight loss, self-harm, and suicide.

The PHQ-A is a measure for tracking depression outcomes in adolescents age 12 years and older in Healthcare Effectiveness Data and Information Set (HEDIS). Additionally, the American Psychiatric Association (APA) recommends the PHQ-A as an outcome measure for clinical practice with children and youth ages 11 to 17 years.^{177 178, 179}

Tool: PROMIS Depression Scale

The PROMIS Depression scale is a validated instrument that measures negative mood, views of self and social cognition, decreased positive affect, and engagement. The PROMIS Depression scale adult version has demonstrated sensitivity to treatment response in clinical research.^{180, 181, 182, 183, 184, 185} Although we have not yet found studies that examined the pediatric version's sensitivity to treatment response, the APA included it in its assessment and monitoring battery. For a brief review of PROMIS instruments and their instrument development process, see Appendix 3.

Tool: Suicide Behavior Questionnaire-Revised (SBQ-R)

The SBQ-R is a four item self-report questionnaire, to be used in patients 13-18 years old that asks about suicidal thoughts or behaviors either in the past, or in anticipation of the future. It is recommended for use by TJC and the National Action Alliance for Suicide Prevention. This tool demonstrates established internal consistency in clinical and non-clinical settings with high test-retest reliability.¹⁸⁶ In addition, concurrent validity has been established.^{187, 188}

Tool: Vanderbilt ADHD Rating Scale

The Vanderbilt Assessment Scale (VAS) includes two (initial and follow-up) assessment components for both parents and teachers that include 55 and 43 items, respectively. Follow-up assessments are shorter in length compared to the initial assessment scale; the follow-up step makes this tool conducive for clinicians to track the change in prevalent symptoms over time.

The National Institute for Children’s Health Quality (NICHQ) VAS is a validated instrument for measuring symptoms of attention deficit hyperactivity disorder (ADHD)^{189, 190} and has been used in clinical research to measure ADHD outcomes.¹⁹¹

When treating children with ADHD, the American Academy of Pediatrics (AAP) recommends that clinicians objectively monitor core symptoms and target goals, which may include use of the VAS follow-up assessment.¹⁹² In 2002, the AAP and NICHQ published a toolkit for ADHD management in primary care settings. This toolkit includes the Vanderbilt rating scales, which can facilitate its effective incorporation into clinical practice and MBC.¹⁹³ The American Academy of Child and Adolescent Psychiatry (AACAP) has included the instrument in its “Toolbox of Forms” for baseline and repeat monitoring of clinical symptoms.¹⁹⁴

The VAS has been identified as a validated assessment tool for diagnosing, treating, and managing ADHD symptoms.¹⁹⁵ The VAS was also studied as an outcome measure in a collaborative care setting across eight pediatric practices in Pittsburgh, PA, of which seven were Children’s Community Pediatric practices and one was an academic pediatric practice affiliated with Children’s Hospital of Pittsburgh.¹⁹⁶

Tool: Pediatric Symptom Checklist (PSC)

The PSC is a validated parent report questionnaire that is used to assess emotional and behavioral-related symptoms and psychosocial functioning in children and youth.^{197, 198} The PSC has been used in clinical effectiveness studies.^{199, 200, 201, 202, 203, 204}

The PSC is so widely used, even many state welfare systems have incorporated the PSC as a routine screening tool.²⁰⁵ Using the PCS as an outcome measure and for MBC may provide opportunities to communicate and benchmark outcomes across providers – and even service systems – without having to “recalibrate” the results. In addition, because it includes physical and psychosocial symptoms and behaviors, the PSC may be well suited for a care system that is becoming more oriented to integrated physical and mental health and substance use (MH/SU) care. It also has high potential to help facilitate communication between physical and mental health professionals.²⁰⁶

The PSC is being used by the Massachusetts General Hospital²⁰⁷ and the California Department of Health and Child Services. The National Quality Forum (NQF) has endorsed it for use in behavioral health program performance measurement (#0722).²⁰⁸ In 2017, California’s Department of Health Care Services standardized the PSC for outcome tracking across all children and youth receiving state-sponsored services.²⁰⁹

The PSC demonstrated sensitivity to change in psychosocial problems over time for children treated for psychiatric disorders.^{210, 211, 212, 213, 214 215, 216, 217, 218, 219}

Tool: Screen for Child Anxiety Related Emotional Disorders (SCARED)

The SCARED is a validated measure of childhood anxiety disorders based on *DSM-IV* criteria.^{220, 221, 222, 223} The California Evidence-Based Clearinghouse for Child Welfare (CEBC) graded the SCARED with an “A” rating, meaning its psychometrics were well-demonstrated.²²⁴ SCARED is incorporated as an outcome measure in the University of California, Los Angeles (UCLA) program (see Table 5). Data from the self-reported completion of the SCARED are used not only to provide regular feedback on clinical status and change on focal outcomes, but also to inform interpretations of outcomes and, when expected outcomes are not achieved, guidance concerning potentially advantageous treatment changes.

Tool: PROMIS Anxiety

The PROMIS Anxiety scale is a validated instrument that measures fear, anxious misery, hyperarousal, and somatic symptoms related to arousal. The PROMIS Anxiety scale adult version demonstrated responsiveness to treatment in clinical research.^{225, 226, 227, 228, 229, 230} Although no studies were found that examined the instrument’s responsiveness with the pediatric versions, the APA recommend their use in their assessment and monitoring battery. For a brief review of PROMIS instruments and their instrument development process, see Appendix 3.

Tool: Mood and Feelings Questionnaire (MFQ)

The MFQ is a valid instrument to assess depression in children. There are short- and long-form versions with 13-items and 33-items, respectively. It was originally developed for children and youth ages eight to 18, based on the *DSM-III-R* symptoms criteria and has been recommended by the National Institute for Health and Clinical Excellence to screen for depression in children and adolescents. There are both parent and youth versions available for use. Previous studies have validated it in clinical settings^{231, 232, 233} and nonclinical settings^{234, 235, 236} and it is currently widely used in clinical and research settings.²³⁷ In addition, MFQ has demonstrated diagnostic accuracy and sensitivity to change, highlighting its use to identify depression and measure symptom change.²³⁸

Tool: The Brief Addiction Monitor (BAM)

The BAM is a brief, validated instrument for measuring addiction behaviors related to alcohol and illicit drugs (previously described within the adult section, above). Although no studies have validated the use of the BAM among youth, it has been utilized in two studies examining change in addiction behaviors among youth receiving post-treatment, mobile-based texting interventions.^{239, 240} The BAM may have sufficient face validity to apply it in clinical settings – with the caveat that additional research is needed – and clinicians are advised to not rely solely upon this instrument’s results for monitoring clinical change.

Tool: Altman Self-Rated Mania Scale (ASRM)

See the adult section above for a brief description of the ASRM. Although the authors did not find studies examining the ASRM's responsiveness to treatment among pediatric populations, the APA encourages its use for pediatric populations within its assessment and monitoring battery (see Appendix 4).

Tool: Children's Version of Eating Attitudes Test (ChEAT-26)

The EAT-26 is likely the most widely used screening tool for identification of symptoms and concerns characteristic of eating disorders. It is highly reliable and valid and was designed as a screening tool to be used with at-risk populations (as well as non-clinical populations).^{241, 242, 243,}
²⁴⁴ The EAT-26 is an instrument to help identify individuals who might be at risk for an eating disorder – not for diagnostic purposes.²⁴⁵ It can be used to screen eating disturbances in general as well as to measure change over time. The EAT-26 can be used in both clinical and non-clinical settings, with both adolescents and adults. The ChEAT is the children's version of the EAT-26 and was developed for children ages eight to 13.²⁴⁶ Test-retest and internal reliability of the ChEAT is comparable to the adult version, according to previous research.²⁴⁷

Tool: Children's Florida Obsessive-Compulsive Inventory (C-FOCI)

The C-FOCI is the child-report version of the FOCI, with some minor distinctions. Its symptom checklist consists of 17 obsessions and compulsions, rated as absent or present over the past month. The symptoms that are endorsed on the symptom checklist are further assessed on the severity scale, which collectively rates obsessions and compulsions on a six-point scale across the same five items in the FOCI. The questionnaire demonstrates adequate psychometric properties with support noted for its treatment sensitivity to cognitive behavioral therapy. Similar to the Obsessive Compulsive Inventory – Child Version OCI-CV, the C-FOCI could also serve as an acceptable screening tool to identify OCD symptoms in children and youth and can be considered for assessment of children and youth's symptom severity and functional impairment.^{248, 249}

Tool: Other PROMIS Measures

PROMIS has publicly released many mental health and other health instruments that have demonstrated responsiveness to treatment in clinical research, including instruments for measuring depression,^{250, 251, 252, 253} anxiety,^{254, 255, 256} anger,²⁵⁷ and peer relationships.^{258, 259} In the DSM-5, the APA has recommended specific PROMIS instruments for assessing and monitoring some pediatric mental health-related conditions (depression, anxiety, anger, sleep disturbance, and inattention). An advantage of the PROMIS measures for pediatric patients is that parent-reported and adolescent-reported instruments are both available.²⁶⁰

Appendix 3: PROMIS Instruments and Development

Funded by the National Institute of Mental Health (NIMH), the Patient-Reported Outcomes Measurement Information System (PROMIS) is a non-profit organization that seeks to develop, maintain, improve and apply patient-reported outcome (PRO) measures across clinical research and practice settings. Since 2004, PROMIS has released over 300 psychometrically validated measures of other health, mental health, and social health into the public domain for further evaluation and development. Moreover, PROMIS measures are often required in federally funded clinical studies, as well as studies funded by the Patient-Centered Outcomes Research Institute (PCORI), which was established under the Patient Protection and Affordable Care Act (ACA) to advance comparative effectiveness research.

As the tables below indicate, the PROMIS addresses dozens of clinical conditions or targets (called “domains” in the PROMIS) of both other health conditions and mental health conditions, as well as aspects of social health. Most domains include one or more paper-based “short form” versions, and the forms vary in length between four and 12 items. MBC practitioners can choose the particular short form to be used, based on their needs to reduce administrative burden or to maximize their ability to measure the various components of a domain (clinical condition).

Among PROMIS’s most innovative contributions is the development of psychometrically validated item banks designed for computer-adapted testing (CAT) across all domains used in MBC. For example, an item bank for a particular condition might have 34 items, but when patients use the CAT system (typically on a hand-held tablet), the particular items to which they respond will be selected based on their early responses on the domain test. This method allows for the possibility of addressing a number of domain elements in patients who have more severe symptoms, while reducing administrative burden for patients who do not have severe symptoms in the domain.

The following tables summarize PROMIS measures for adult and pediatric populations. Note that the Adult and Pediatric “Profile of Global Health” instruments each include 10 items for adults and there are both 7 and 9 item versions for children and youth. The global instruments cover the domains (e.g., Fatigue, Anxiety, etc.) listed below their respective headings in the two tables below. Only one or two items are used per domain on the global tests. But if the patient’s response is positive on an item, then a short form – either paper-based or CAT-based – can then be completed.

Adult PROMIS Instruments		
Physical Health	Mental Health	Social Health
Domains (Clinical Conditions/Targets) in the Adult Profile of Global Health (10 Items)		
<ul style="list-style-type: none"> • Fatigue • Pain Intensity • Pain Interference • Physical Function 	<ul style="list-style-type: none"> • Anxiety • Depression 	<ul style="list-style-type: none"> • Ability to Participate in Social Roles & Activities
All Other PROMIS Instruments by Domain (Clinical Condition)		
<ul style="list-style-type: none"> • Dyspnea • Gastrointestinal Symptoms • Itch • Pain Behavior • Pain Quality • Sexual Function • Sleep-Related Impairment 	<ul style="list-style-type: none"> • Alcohol • Anger • Cognitive Function • Life Satisfaction • Meaning & Purpose • Positive Affect • Psychosocial Illness Impact Self-Efficacy for Managing Chronic Conditions: <ul style="list-style-type: none"> – Smoking – Substance use 	<ul style="list-style-type: none"> • Companionship • Satisfaction with Social Roles & Activities • Social Isolation • Social Support

Pediatric PROMIS Instruments		
Physical Health	Mental Health	Social Health
Domains (Clinical Conditions/Targets) in the Pediatric Profile of Global Health (7, 9 Item versions)		
<ul style="list-style-type: none"> • Fatigue • Mobility • Pain Intensity • Pain Interference • Upper Extremity Function 	<ul style="list-style-type: none"> • Anxiety • Depression Symptoms 	<ul style="list-style-type: none"> • Peer Relationships
All Other PROMIS Instruments by Domain (Clinical Condition/Target)		
<ul style="list-style-type: none"> • Asthma Impact • Pain Behavior • Pain Quality • Physical Activity • Physical Stress Experiences • Sleep • Strength Impact 	<ul style="list-style-type: none"> • Anger • Cognitive Function • Life Satisfaction • Meaning & Purpose • Positive Affect • Psychological Stress Experiences 	<ul style="list-style-type: none"> • Family Relationships

All PROMIS measures follow a five-stage process of instrument maturity:

1. Conceptualization & Item Pool Development (Developmental),
2. Calibration Phase (Developmental),
3. Calibrated and Preliminary Validation Completed (Public Release),
4. Responsiveness and Expansion (Maturing), and
5. Fully Mature User Support.

All publicly available PROMIS instruments have completed stage 3, in which they have satisfied industry-standard psychometric testing (reliability and validity). Further development and testing of the instruments' responsiveness to treatment or sensitivity to change over time are expected to be conducted by the broader scientific community and published in peer-reviewed journals. For a full description of the PROMIS instrument development process see PROMIS Instrument Development and Validation Scientific Standards at

http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf.²⁶¹

Appendix 4: APA Protocol and Battery for Patient Assessment and Monitoring

Upon releasing the *DSM-5*, the American Psychiatric Association (APA) published adult and pediatric assessment protocols and a battery of select patient-reported outcome instruments labeled as “emerging measures” for patient assessment and monitoring. The recommended protocol includes administering a 23-item global screener for mental health and substance use (MH/SU) conditions (labeled “Level-1 Cross-Cutting Symptom Measure”) and then following up with Level-2 condition-specific instruments, as needed. The APA recommends using the World Health Organization Disability Assessment Schedule (WHODAS) II as a measure of impairment/disability. Additional “disorder-specific severity measures” are also noted.

The following table summarizes the conditions and domains examined within the Level-1 Cross-Cutting Symptom Measure, and the available condition-specific and disorder-specific measures suggested by the APA, for both adults and children/youth.

APA Protocol and Battery of Suggested Instruments for Patient Assessment and Monitoring		
Level-1: “Cross-Cutting Symptom Measure” Domains ¹	Level-2: Condition-Specific Measures	Disorder-Specific Severity Measures
Patient-Reported – Adults		
<ul style="list-style-type: none"> • Depression • Anger • Mania • Anxiety • Somatic Symptoms • <i>Suicidal Ideation</i> • <i>Psychosis</i> • Sleep Problems • <i>Memory</i> • Repetitive Thoughts and Behaviors • <i>Dissociation</i> • <i>Personality Functioning</i> • Substance Use 	<ul style="list-style-type: none"> • PROMIS Depression • PROMIS Anger • Altman Self-Rating Mania Scale • PROMIS Anxiety • PHQ-15 (Somatic symptom severity) • PROMIS Sleep Disturbance • Repetitive Thoughts and Behaviors² • ASSIST (Substance Use) 	<ul style="list-style-type: none"> • PHQ-9 (Depression) • APA Published Severity Measures • Separation Anxiety Disorder • Specific Phobia • Social Anxiety Disorder • Panic Disorder • Agoraphobia • Generalized Anxiety Disorder • National Stressful Events Survey PTSD Short Form (NSESS) • National Stressful Events Survey Acute Stress Short Form (NSESS)

¹ *Italicized* domains do not have a corresponding Level-2 condition-specific measure.

² Adapted from the Florida Obsessive-Compulsive Inventory (FOCI) Severity Scale (part b).

APA Protocol and Battery of Suggested Instruments for Patient Assessment and Monitoring		
Level-1: “Cross-Cutting Symptom Measure” Domains ¹	Level-2: Condition-Specific Measures	Disorder-Specific Severity Measures
Parent/Guardian-Reported and Clinician-Rated Measures for Children/Youth Age 6 to 17 Years		
<ul style="list-style-type: none"> • Somatic Symptoms • Sleep Problems • Inattention • Depression • Anger • Irritability • Mania • Anxiety • <i>Psychosis</i> • <i>Repetitive Thoughts and Behaviors</i> • Substance Use • <i>Suicidal Ideation/Suicide Attempts</i> 	<ul style="list-style-type: none"> • PHQ-15 (Somatic Symptom Severity) • PROMIS Sleep Disturbance • PROMIS Depression • PROMIS Anger • Affective Reactivity Index (Irritability) • Altman Self-Rating Mania Scale • PROMIS Anxiety • Repetitive Thoughts and Behaviors³ • ASSIST (Substance Use) 	(Clinician-Rated) <ul style="list-style-type: none"> • APA Published Severity Measures • Autism Spectrum and Social Communication • Psychosis Symptom Severity • Somatic Symptom Disorder • Oppositional Defiant Disorder • Conduct Disorder • Non-Suicidal Self-Injury
Patient-Reported – Youth Ages 11 to 17 Years		
<ul style="list-style-type: none"> • Somatic Symptoms • Sleep Problems • Inattention • Depression • Anger • Irritability • Mania • Anxiety • <i>Psychosis</i> • <i>Repetitive Thoughts and Behaviors</i> • Substance Use • <i>Suicidal Ideation/Suicide Attempts</i> 	<ul style="list-style-type: none"> • PHQ-15 (Somatic Symptom Severity) • PROMIS Sleep Disturbance • PROMIS Depression • PROMIS Anger • Affective Reactivity Index (Irritability) • Altman Self-Rating Mania Scale • PROMIS Anxiety • Repetitive Thoughts and Behaviors⁴ • ASSIST (Substance Use) 	<ul style="list-style-type: none"> • PHQ-A (Depression) • APA Published Severity Measures • Separation Anxiety Disorder • Specific Phobia • Social Anxiety Disorder • Agoraphobia • Generalized Anxiety Disorder • National Stressful Events Survey PTSD Short Form (NSESS) • National Stressful Events Survey Acute Stress Short Form (NSESS) • Brief Dissociative Experiences Scale (DES-B)

³ Adapted from the Florida Obsessive-Compulsive Inventory (FOCI) Severity Scale (part b).

⁴ Adapted from the Florida Obsessive-Compulsive Inventory (FOCI) Severity Scale (part b).

Appendix 5: List of Abbreviations

AAP	American Academy of Pediatrics
AACAP	American Academy of Child and Adolescent Psychiatry
ACOG	American College of Obstetricians and Gynecologists
BAM	Brief Addictions Monitor
BPRS	Brief Psychiatric Rating Scale
CMS	Centers for Medicare & Medicaid Services
CMTS	Care Management Tracking System
DoD	Department of Defense
EHR	Electronic health record
FQHC	Federally qualified health center
GAD-7	Generalized Anxiety Disorder 7-Item
HHS	Department of Health and Human Services
ITF	Interagency Task Force on Military and Veterans Mental Health
ICHOM	International Consortium for Health Outcome Measurement
ICSI	Institute for Clinical Systems Improvement
MBC	Measurement-based care
MCPAP	Massachusetts Child Psychiatry Access Program
MGH	Massachusetts General Hospital
NICHQ	National Institute for Children’s Health Quality
NIDA	National Institute on Drug Abuse
NQF	National Quality Forum
PANSS	Positive and Negative Syndrome Scale

PCMH	Patient-centered medical homes
PHQ	Patient Health Questionnaire
PRO	Patient-reported outcomes
PQ	Prodromal Questionnaire
PSC	Pediatric Symptoms Checklist
SAMSHA	Substance Abuse and Mental Health Services Administration
SF	Medical Outcomes Study Short-Form Health Survey
SLC-20	Hopkins Symptom Checklist
KF	Kennedy Forum
USPSTF	U.S. Preventive Services Task Force
VA	Department of Veterans Affairs
WHODAS	World Health Organization Disability Assessment Schedule

Endnotes

¹ Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A Tipping Point for Measurement-Based Care. *Psychiatric Services (Washington, D.C.)*, 68(2), 179–188.

<https://doi.org/10.1176/appi.ps.201500439>

² Fortney, J., Sladek, R., Unützer, J., Kennedy, P., Harbin, H., Emmet, B., Alfred, L., & Carneal, G. (2015). *Fixing behavioral health care in America: A national call for measurement-based care in the delivery of behavioral health services*. The Kennedy Forum. www.thekennedyforum.org

³ Fortney, J., Sladek, R., Unützer, J., Kennedy, P., Harbin, H., Emmet, B., Alfred, L., & Carneal, G. (2015). Previously cited.

See also the supplement to that paper, entitled, *A Core Set of Outcome Measures for Behavioral Health Across Service Settings*, which is also available through The Kennedy Forum website.

⁴ Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). Previously cited.

⁵ Food and Drug Administration. (2009). *Guidance for industry patient-reported outcome measures: Use in medical product development to support labeling claims*. U.S. Department of Health and Human Services. Retrieved from <https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf>

⁶ Minsk, A., Blakely, J.S., & Cohen, D.R. (2010). FDA Issues final guidance on patient-reported outcome measures used to support labeling claims, *Regulatory Focus*, 5.

⁷ Cella, D., Hahn, E., Jensen, S., Butt, Z., Nowinski, C., & Rothrock, N. (2012). Methodological Issues in The Selection, Administration and Use of Patient-Reported Outcomes In Performance Measurement In Health Care Settings. *National Quality Forum*, 63.

⁸ PROMIS® (2013). *Instrument development and validation scientific standards version 2.0*. <http://www.healthmeasures.net/explore-measurement-systems/promis/measure-development-research>

⁹ Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737–745.

¹⁰ Mokkink, L. B., Prinsen, C. A. C., Bouter, L. M., Vet, H. C. W. de, & Terwee, C. B. (2016). The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian Journal of Physical Therapy*, 20(2), 105–113.

¹¹ Mauri, A., Harbin, H., Unützer, J., Carlo, A., Ferguson, R., & Schoenbaum, M. (2017). Payment Reform and Opportunities for Behavioral Health: Alternative Payment Model Examples. Thomas Scattergood Behavioral Health Foundation, Peg’s Foundation and The Kennedy Forum.

¹² Centers for Medicare & Medicaid Services. (n.d.). Five-Star Quality Rating System. <https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandCompliance/FSQRS>

¹³ More information on CAHPS can be found at: <https://www.ahrq.gov/cahps/about-cahps/index.html>

¹⁴ Dejesus, R. S., Vickers, K. S., Melin, G. J., & Williams, M. D. (2007). A system-based approach to depression management in primary care using the Patient Health Questionnaire-9. *Mayo Clinic Proceedings*, 82(11), 1395–1402.

¹⁵ Yeung, A. S., Jing, Y., Brenneman, S. K., Chang, T. E., Baer, L., Hebden, T., Kalsekar, I., McQuade, R. D., Kurlander, J., Siebenaler, J., & Fava, M. (2012). Clinical Outcomes in Measurement-based Treatment (Comet): a trial of depression monitoring and feedback to primary care physicians. *Depression and Anxiety*, 29(10), 865–873.

¹⁶ The American Psychiatric Association. (2018, September 21). *APA Awarded CMS Funding to Develop Quality Measures*. <https://www.psychiatry.org/newsroom/news-releases/apa-awarded-cms-funding-to-develop-quality-measures>

¹⁷ Joint Commission’s National Patient Safety Goal 15.01.01 (see <https://www.jointcommission.org/resources/patient-safety-topics/suicide-prevention/>).

¹⁸ Measurement Based Care Designation. (n.d.). Retrieved from <https://www.urac.org/programs/measurement-based-care>

¹⁹ National Alliance of Healthcare Purchaser Coalitions. (2018, August). *Achieving Value in Mental Health Support: A Deep Dive Powered by eValue8*. <https://www.nationalalliancehealth.org/initiatives/initiatives-national/workplace-mental-health/evaluate8-deepdive>

This report was issued by the National Alliance to assist their employer member organizations in selecting health plans for their employees summarized findings from this initiative.

²⁰ National Alliance of Healthcare Purchaser Coalitions. (2019, November 12). *To Combat the Mental Health & Substance Use Public Health Crisis Employer, Physician and Policy Groups Partner*. <https://www.nationalalliancehealth.org/www/news/news-press-releases/path-forward>

²¹ National Alliance of Healthcare Purchaser Coalitions. (2019). *The Path Forward for Mental Health and Substance Use Improving Access to Effective Treatment*. https://higherlogicdownload.s3.amazonaws.com/NAHPC/3d988744-80e1-414b-8881-aa2c98621788/UploadedImages/BH_PF_Action_Plan_Executive_Summary_FINAL_1219.pdf

²² National Alliance of Healthcare Purchaser Coalitions. (2018, August). Previously cited.

²³ *Health Care Accreditation, Health Plan Accreditation Organization—NCQA*. (n.d.). NCQA. <https://www.ncqa.org/>

²⁴ *URAC Accreditation Programs*. (n.d.). URAC. <https://www.urac.org/accreditation-programs>

²⁵ *CARF Accreditation*. (n.d.). CARF. <http://www.carf.org/Accreditation/>

²⁶ The Joint Commission Accreditation & Certification. (n.d.). TJC. <https://www.jointcommission.org/accreditation-and-certification>

-
- ²⁷ Cella, D., Hahn, E., Jensen, S., Butt, Z., Nowinski, C., & Rothrock, N. (2012). Previously cited.
- ²⁸ Storch, E. A., Kaufman, D. A. S., Bagner, D., Merlo, L. J., Shapira, N. A., Geffken, G. R., Murphy, T. K., & Goodman, W. K. (2007). Florida Obsessive-Compulsive Inventory: development, reliability, and validity. *Journal of Clinical Psychology, 63*(9), 851–859.
- ²⁹ Storch, E. A., Khanna, M., Merlo, L. J., Loew, B. A., Franklin, M., Reid, J. M., Goodman, W. K., & Murphy, T. K. (2009). Children’s Florida Obsessive Compulsive Inventory: psychometric properties and feasibility of a self-report measure of obsessive-compulsive symptoms in youth. *Child Psychiatry and Human Development, 40*(3), 467–483.
- ³⁰ For information on interpreting clinically meaningful change from the PCS results see Murphy, J. M., Blais, M., Baer, L., McCarthy, A., Kamin, H., Masek, B., & Jellinek, M. (2015). Measuring outcomes in outpatient child psychiatry: reliable improvement, deterioration, and clinically significant improvement. *Clinical Child Psychology and Psychiatry, 20*(1), 39–52.
- ³¹ *UCLA Behavioral Health Associates*. (n.d.). UCLA Division of Population Behavioral Health. <http://dphb.ucla.edu/Behavioral-Health-Associates>
- ³² Fortney, J., Sladek, R., Unützer, J., Kennedy, P., Harbin, H., Emmet, B., Alfred, L., & Carneal, G. (2015). Previously cited.
- ³³ Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). Previously cited.
- ³⁴ Fortney, J., Sladek, R., Unützer, J., Kennedy, P., Harbin, H., Emmet, B., Alfred, L., & Carneal, G. (2015). Previously cited.
- ³⁵ Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). Previously cited.
- ³⁶ Waschbusch, D. A., Pearl, A., Babinski, D. E., Essayli, J. H., Koduvayur, S. P., Liao, D., Mukherjee, D., & Saunders, E. F. H. (2020). Developing Measurement-Based Care for Youth in an Outpatient Psychiatry Clinic: The Penn State Psychiatry Clinical Assessment and Rating Evaluation System for Youth (PCARES-Youth). *Evidence-Based Practice in Child and Adolescent Mental Health, 5*(1), 67–82.
- ³⁷ There are other variations of the PHQ that are intended for measuring anxiety, somatic symptoms, or depression and anxiety combinations.
- ³⁸ The PHQ-8 is primarily used for research and excludes the last item of the PHQ-9, related to self-harm.
- ³⁹ Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613.
- ⁴⁰ Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA, 282*(18), 1737–1744.
- ⁴¹ Spitzer, R. L., Williams, J. B., Kroenke, K., Hornyak, R., & McMurray, J. (2000). Validity and utility of the PRIME-MD patient health questionnaire in assessment of 3000 obstetric-gynecologic patients: the PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *American Journal of Obstetrics and Gynecology, 183*(3), 759–769.

-
- ⁴² Löwe, B., Gräfe, K., Zipfel, S., Witte, S., Loerch, B., & Herzog, W. (2004). Diagnosing ICD-10 depressive episodes: superior criterion validity of the Patient Health Questionnaire. *Psychotherapy and Psychosomatics*, *73*(6), 386–390.
- ⁴³ Löwe, B., Kroenke, K., Herzog, W., & Gräfe, K. (2004). Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of Affective Disorders*, *81*(1), 61–66.
- ⁴⁴ Malpass, A., Shaw, A., Kessler, D., & Sharp, D. (2010). Concordance between PHQ-9 scores and patients' experiences of depression: a mixed methods study. *The British Journal of General Practice: The Journal of the Royal College of General Practitioners*, *60*(575).
- ⁴⁵ Moore, M., Ali, S., Stuart, B., Leydon, G. M., Ovens, J., Goodall, C., & Kendrick, T. (2012). Depression management in primary care: an observational study of management changes related to PHQ-9 score for depression monitoring. *The British Journal of General Practice*, *62*(599), e451–457.
- ⁴⁶ Ell, K., Xie, B., Quon, B., Quinn, D. I., Dwight-Johnson, M., & Lee, P.-J. (2008). Randomized controlled trial of collaborative care management of depression among low-income patients with cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *26*(27), 4488–4496.
- ⁴⁷ Patten, S.B., & Schopflocher, B. (2009). Longitudinal epidemiology of major depression as assessed by the Brief Patient Health Questionnaire (PHQ-9). *Comprehensive Psychology*, *50*(1): 26–33.
- ⁴⁸ Hammond, G. C., Croudace, T. J., Radhakrishnan, M., Lafortune, L., Watson, A., McMillan-Shields, F., & Jones, P. B. (2012). Comparative effectiveness of cognitive therapies delivered face-to-face or over the telephone: an observational study using propensity methods. *PloS One*, *7*(9), e42916.
- ⁴⁹ Institute for Clinical Systems Improvement. (2016, March). *Depression, Adult in Primary Care*. <https://www.icsi.org/guideline/depression/>
- ⁵⁰ International Consortium for Health Outcomes Measurement. (2016). *Standard set for depression and anxiety*. <https://www.ichom.org/portfolio/depression-anxiety/>
- ⁵¹ ICHOM is an international nonprofit whose mission is to promote health outcomes practices across health care specialties. ICHOM includes the PHQ-9 in its “Standard (measurement) Set” for treating depression and anxiety. In 2016, the Danish health authorities endorsed the PHQ-9 for depression-related MBC.
- ⁵² Department of Defense, Department of Veterans Affairs & Department of Health and Human Services. (2016). *Interagency task force on military and veteran’s mental health: 2016 annual report*. https://www.mentalhealth.va.gov/docs/ITF_2016_Annual_Report_November_2016.pdf
- ⁵³ Kroenke, K., Spitzer, R.L., & Williams, J.B. (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. *Medical Care*, *41*(11):1284–1292.
- ⁵⁴ Arroll, B., Goodyear-Smith, F., Crengle, S., Gunn, J., Kerse, N., Fishman, T., Falloon, K., & Hatcher, S. (2010). Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. *Annals of Family Medicine*, *8*(4), 348–353.
- ⁵⁵ Arroll, B., Khin, N., & Kerse, N. (2003). Screening for depression in primary care with two verbally asked questions: cross sectional study. *BMJ (Clinical Research Ed.)*, *327*(7424), 1144–1146.

⁵⁶ According to Arroll et al. (2010), the PHQ-2 has good sensitivity but poor specificity, meaning it is better at accurately identifying those with depression than accurately identifying those without depression. With all of this in mind, the authors recommended decreasing the PHQ-2 positive score threshold to two or higher from three or higher, as well as following up by using the PHQ-9 and incorporating clinical observation and wisdom.

⁵⁷ Department of Defense, Department of Veterans Affairs & Department of Health and Human Services. (2016). Previously cited.

⁵⁸ Cook, K. F., Jensen, S. E., Schalet, B. D., Beaumont, J. L., Amtmann, D., Czajkowski, S., Dewalt, D. A., Fries, J. F., Pilkonis, P. A., Reeve, B. B., Stone, A. A., Weinfurt, K. P., & Cella, D. (2016). PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *Journal of Clinical Epidemiology*, *73*, 89–102.

⁵⁹ Jensen, R. E., Moinpour, C. M., Potosky, A. L., Lobo, T., Hahn, E. A., Hays, R. D., Cella, D., Smith, A. W., Wu, X.-C., Keegan, T. H. M., Paddock, L. E., Stroup, A. M., & Eton, D. T. (2017). Responsiveness of 8 Patient-Reported Outcomes Measurement Information System (PROMIS) measures in a large, community-based cancer study cohort. *Cancer*, *123*(2), 327–335.

⁶⁰ Lee, A. C., Driban, J. B., Price, L. L., Harvey, W. F., Rodday, A. M., & Wang, C. (2017). Responsiveness and Minimally Important Differences for 4 Patient-Reported Outcomes Measurement Information System Short Forms: Physical Function, Pain Interference, Depression, and Anxiety in Knee Osteoarthritis. *The Journal of Pain: Official Journal of the American Pain Society*, *18*(9), 1096–1110.

⁶¹ Purvis, T. E., Andreou, E., Neuman, B. J., Riley, L. H., & Skolasky, R. L. (2017). Concurrent Validity and Responsiveness of PROMIS Health Domains Among Patients Presenting for Anterior Cervical Spine Surgery. *Spine*, *42*(23), E1357–1365.

⁶² Schalet, B. D., Pilkonis, P. A., Yu, L., Dodds, N., Johnston, K. L., Yount, S., Riley, W., & Cella, D. (2016). Clinical validity of PROMIS Depression, Anxiety, and Anger across diverse clinical samples. *Journal of Clinical Epidemiology*, *73*, 119–127.

⁶³ Wohlfahrt, A., Bingham, C. O., Marder, W., Phillips, K., Bolster, M. B., Moreland, L. W., Zhang, Z., Neogi, T., & Lee, Y. C. (2019). Responsiveness of Patient-Reported Outcomes Measurement Information System Measures in Rheumatoid Arthritis Patients Starting or Switching a Disease-Modifying Antirheumatic Drug. *Arthritis Care & Research*, *71*(4), 521–529.

⁶⁴ Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*(5), 317–325.

⁶⁵ Beard, C., & Björgvinsson, T. (2014). Beyond generalized anxiety disorder: psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *Journal of Anxiety Disorders*, *28*(6), 547–552.

⁶⁶ Dear, B. F., Titov, N., Sunderland, M., McMillan, D., Anderson, T., Lorian, C., & Robinson, E. (2011). Psychometric comparison of the generalized anxiety disorder scale-7 and the Penn State Worry Questionnaire for measuring response during treatment of generalised anxiety disorder. *Cognitive Behaviour Therapy*, *40*(3), 216–227.

⁶⁷ Mewton, L., Wong, N., & Andrews, G. (2012). The effectiveness of internet cognitive behavioural therapy for generalized anxiety disorder in clinical practice. *Depression and Anxiety*, *29*(10), 843–849.

- ⁶⁸ Robinson, E., Titov, N., Andrews, G., McIntyre, K., Schwencke, G., & Solley, K. (2010). Internet treatment for generalized anxiety disorder: a randomized controlled trial comparing clinician vs. technician assistance. *PLoS One*, *5*(6), e10942.
- ⁶⁹ Titov, N., Gibson, M., Andrews, G., & McEvoy, P. (2009). Internet treatment for social phobia reduces comorbidity. *Australian & New Zealand Journal of Psychiatry*, *43*(8), 754–759.
- ⁷⁰ Hammond, G. C., Croudace, T. J., Radhakrishnan, M., Lafortune, L., Watson, A., McMillan-Shields, F., & Jones, P. B. (2012). Previously cited.
- ⁷¹ Plummer, F., Manea, L., Trepel, D., & McMillan, D. (2016). Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *General Hospital Psychiatry*, *39*, 24–31.
- ⁷² International Consortium for Health Outcomes Measurement (ICHOM). (2016). *ICHOM*. <http://www.ichom.org/>
- ⁷³ Department of Defense, Department of Veterans Affairs & Department of Health and Human Services. (2016). Previously cited.
- ⁷⁴ Cook, K. F., Jensen, S. E., Schalet, B. D., Beaumont, J. L., Amtmann, D., Czajkowski, S., Dewalt, D. A., Fries, J. F., Pilkonis, P. A., Reeve, B. B., Stone, A. A., Weinfurt, K. P., & Cella, D. (2016). Previously cited.
- ⁷⁵ Jensen, R. E., Moinpour, C. M., Potosky, A. L., Lobo, T., Hahn, E. A., Hays, R. D., Cella, D., Smith, A. W., Wu, X.-C., Keegan, T. H. M., Paddock, L. E., Stroup, A. M., & Eton, D. T. (2017). Previously cited.
- ⁷⁶ Lee, A. C., Driban, J. B., Price, L. L., Harvey, W. F., Rodday, A. M., & Wang, C. (2017). Previously cited.
- ⁷⁷ Purvis, T. E., Andreou, E., Neuman, B. J., Riley, L. H., & Skolasky, R. L. (2017). Previously cited.
- ⁷⁸ Schalet, B. D., Pilkonis, P. A., Yu, L., Dodds, N., Johnston, K. L., Yount, S., Riley, W., & Cella, D. (2016). Previously cited.
- ⁷⁹ Wohlfahrt, A., Bingham, C. O., Marder, W., Phillips, K., Bolster, M. B., Moreland, L. W., Zhang, Z., Neogi, T., & Lee, Y. C. (2019). Previously cited.
- ⁸⁰ Houck, P. R., Spiegel, D. A., Shear, M. K., & Rucci, P. (2002). Reliability of the self-report version of the panic disorder severity scale. *Depression and Anxiety*, *15*(4), 183–185.
- ⁸¹ Shear, M. K., Brown, T. A., Barlow, D. H., Money, R., Sholomskas, D. E., Woods, S. W., Gorman, J. M., & Papp, L. A. (1997). Multicenter collaborative panic disorder severity scale. *The American Journal of Psychiatry*, *154*(11), 1571–1575.
- ⁸² Furukawa, T. A., Shear, M. K., Barlow, D. H., Gorman, J. M., Woods, S. W., Money, R., Etschel, E., Engel, R. R., & Leucht, S. (2009). Evidence-based Guidelines for Interpretation of the Panic Disorder Severity Scale. *Depression and Anxiety*, *26*(10), 922–929.
- ⁸³ Pilkonis, P.A., Yu, L., Colditz, J., Dodds, N., Johnston, K.L., Maihoefer, C., Stover, A.M., Daley, D.C., & McCarty, D. (2013). Item banks for alcohol use from the Patient-Reported Outcomes Measurement Information System (PROMIS®): Use, consequences, and expectancies. *Drug and Alcohol Dependence*, *130*(1):167–177.

-
- ⁸⁴ Johnston, K. L., Lawrence, S. M., Dodds, N. E., Yu, L., Daley, D. C., & Pilonis, P. A. (2016). Evaluating PROMIS[®] instruments and methods for patient-centered outcomes research: Patient and provider voices in a substance use treatment setting. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, 25(3), 615–624.
- ⁸⁵ Marsden, J., Tai, B., Ali, R., Hu, L., Rush, A. J., & Volkow, N. (2019). Measurement-based care using DSM-5 for opioid use disorder: can we make opioid medication treatment more effective? *Addiction (Abingdon, England)*, 114(8), 1346–1353.
- ⁸⁶ Higgins-Biddle, J. C., & Babor, T. F. (2018). A review of the Alcohol Use Disorders Identification Test (AUDIT), AUDIT-C, and USAUDIT for screening in the United States: Past issues and future directions. *The American Journal of Drug and Alcohol Abuse*, 44(6), 578–586.
- ⁸⁷ Babor, T. F., Higgins-Biddle, J. C., & Robaina, K. (2017). USAUDIT - The Alcohol Use Disorders Identification Test, Adapted for Use in the United States: A Guide for Primary Care Practitioners. Substance Abuse and Mental Health Services Administration. <https://www.ct.gov/dmhas/lib/dmhas/publications/USAUDIT-2017.pdf>
- ⁸⁸ Bradley, K. A., McDonell, M. B., Bush, K., Kivlahan, D. R., Diehr, P., & Fihn, S. D. (1998). The AUDIT alcohol consumption questions: reliability, validity, and responsiveness to change in older male primary care patients. *Alcoholism, Clinical and Experimental Research*, 22(8), 1842–1849.
- ⁸⁹ Jeffrey, J., Sinclair, M., Linonis, R., Semaan, A., Hsiao, T., Chiu, W., Grossman, M., & Lester, P. (2018). Integration of a Web-based Behavioral Health Assessment within a Collaborative Care Setting. *Journal of Mobile Technology in Medicine*, 7(1), 9–15.
- ⁹⁰ Cacciola, J.S., Alterman, A.I., DePhilippis, D., & Drapkin, M.L. (2013). Development and initial evaluation of the Brief Addiction Monitor (BAM). *Journal of Substance Abuse Treatment*, 44(3), 256–263.
- ⁹¹ Witkiewitz, K. (2019). Double standards and gold standards in the evaluation of how a person feels and functions in substance use disorder pharmacotherapy trials. *Addiction*, 114(1), 17–18.
- ⁹² Kiluk, B.D., Fizmaurice, G.M., Strain, E.C., & Weiss, R.D. (2018). What defines a clinically meaningful outcome in the treatment of substance use disorders: Reductions in direct consequences of drug use or improvement in overall functioning? *Addiction*, 114(1), 9–15.
- ⁹³ Department of Defense, Department of Veterans Affairs & Department of Health and Human Services. (2016). Previously cited.
- ⁹⁴ Note that the BAM has undergone changes in its item response format from continuous to discrete Likert items, and then again revised back to continuous items (now called the BAM-R). This change was attributed to limitations in the VA's electronic health record and it was made to accommodate the original response format. Whenever possible, practices should use the versions of the BAM that include continuous scoring formats, as the factor structure (construct validity) is not consistent across response formats. For more details regarding permutations of the BAM, see Gaddy, M., Casner, H.G., & Rosinski, J. (2018). Factor structure and measurement invariance of the Brief Addiction Monitor. *Journal of Substance Abuse Treatment*, 90, 29–37.
- ⁹⁵ Smith, G. R., Burnam, M. A., Mosley, C. L., Hollenberg, J. A., Mancino, M., & Grimes, W. (2006). Reliability and validity of the Substance Abuse Outcomes Module. *Psychiatric Services*, 57(10), 1452–1460.

⁹⁶ Wilkins, K. C., Lang, A. J., & Norman, S. B. (2011). Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. *Depression and Anxiety, 28*(7), 596–606.

⁹⁷ Monson, C. M., Gradus, J. L., Young-Xu, Y., Schnurr, P. P., Price, J. L., & Schumm, J. A. (2008). Change in posttraumatic stress disorder symptoms: do clinicians and patients agree? *Psychological Assessment, 20*(2), 131–138.

⁹⁸ Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). *The PTSD checklist for DSM-5 (PCL-5)*. Retrieved from www.ptsd.va.gov

⁹⁹ Department of Defense, Department of Veterans Affairs & Department of Health and Human Services. (2016). Previously cited.

¹⁰⁰ Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). Previously cited.

¹⁰¹ The Columbia Lighthouse Project/Center for Suicide Risk Assessment. (2018, July 25). *The Columbia Suicide Severity Rating Scale (C-SSRS): Supporting evidence*. http://cssrs.columbia.edu/wp-content/uploads/CSSRS_Supporting-Evidence_Book_2018-08-13.pdf

¹⁰² Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., Currier, G. W., Melvin, G. A., Greenhill, L., Shen, S., & Mann, J. J. (2011). The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *The American Journal of Psychiatry, 168*(12), 1266–1277.

¹⁰³ Ionescu, D. F., Swee, M. B., Pavone, K. J., Taylor, N., Akeju, O., Baer, L., Nyer, M., Cassano, P., Mischoulon, D., Alpert, J. E., Brown, E. N., Nock, M. K., Fava, M., & Cusin, C. (2016). Rapid and Sustained Reductions in Current Suicidal Ideation Following Repeated Doses of Intravenous Ketamine: Secondary Analysis of an Open-Label Study. *The Journal of Clinical Psychiatry, 77*(6), e719-725.

¹⁰⁴ The Joint Commission. (2020, March 19). *Ligatures and/or Suicide Risk Reduction—Use of an Evidence-based Process to Assess Risk*. <https://www.jointcommission.org/standards/Standard-FAQs/Critical Access Hospital/National Patient Safety Goals NPSG/000002237>

¹⁰⁵ Center for Drug Evaluation and Research. (2012, August). *Guidance for industry – suicidality: Prospective assessment of occurrence in clinical trials, draft guidance* [Revision 1]. United States Food and Drug Administration, United States Department of Health and Human Services. <http://www.fda.gov/downloads/Drugs/Guidances/UCM225130.pdf>

¹⁰⁶ Some researchers have challenged the psychometric validity of the C-SSRS and claimed it has not received the same level of scrutiny other PRO measures have undergone (Giddens et al., 2014; Sheehan et al., 2014). For instance, researchers have suggested that the C-SSRS is not comprehensive in its probes of suicidal ideation and behaviors, uses ambiguous and imprecise language, and lacks consistency in terms and definitions throughout the measure. Also, it is constructed in such a way that it is burdensome and confusing to recipients and composed of too much ambiguous language. These limitations may lead to misclassifications of patients in that it might falsely intensify a high risk of suicide in some patients or fail to identify other patients who have a high risk of suicide.

- ¹⁰⁷ Horowitz, L. M., Snyder, D. J., Boudreaux, E. D., He, J. P., Harrington, C. J., Cai, J., Claassen, C. A., Salhany, J. E., Dao, T., Chaves, J. F., Jobes, D. A., Merikangas, K. R., Bridge, J. A., & Pao, M. (2020). Validation of the Ask Suicide-Screening Questions for Adult Medical Inpatients: A Brief Tool for All Ages. *Psychosomatics*.
- ¹⁰⁸ Horowitz, L. M., Bridge, J. A., Teach, S. J., Ballard, E., Klima, J., Rosenstein, D. L., Wharff, E. A., Ginnis, K., Cannon, E., Joshi, P., & Pao, M. (2012). Ask Suicide-Screening Questions (ASQ): a brief instrument for the pediatric emergency department. *Archives of Pediatrics & Adolescent Medicine*, *166*(12), 1170–1176.
- ¹⁰⁹ Atkinson, T. M., Mendoza, T. R., Sit, L., Passik, S., Scher, H. I., Cleeland, C., & Basch, E. (2010). The Brief Pain Inventory and its “pain at its worst in the last 24 hours” item: clinical trial endpoint considerations. *Pain Medicine (Malden, Mass.)*, *11*(3), 337–346.
- ¹¹⁰ Atkinson, T. M., Mendoza, T. R., Sit, L., Passik, S., Scher, H. I., Cleeland, C., & Basch, E. (2010). Previously cited.
- ¹¹¹ Keller, S., Bann, C. M., Dodd, S. L., Schein, J., Mendoza, T. R., & Cleeland, C. S. (2004). Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. *The Clinical Journal of Pain*, *20*(5), 309–318.
- ¹¹² Kean, J., Monahan, P. O., Kroenke, K., Wu, J., Yu, Z., Stump, T. E., & Krebs, E. E. (2016). Comparative Responsiveness of the PROMIS Pain Interference Short Forms, Brief Pain Inventory, PEG, and SF-36 Bodily Pain Subscale. *Medical Care*, *54*(4), 414–421.
- ¹¹³ Krebs, E. E., Bair, M. J., Damush, T. M., Tu, W., Wu, J., & Kroenke, K. (2010). Comparative responsiveness of pain outcome measures among primary care patients with musculoskeletal pain. *Medical Care*, *48*(11), 1007–1014.
- ¹¹⁴ National Quality Forum (2018). *Quality ID #131 (NQF 0420): Pain assessment and follow-up – National Quality Strategy Domain: Communication and care coordination*.
https://qpp.cms.gov/docs/QPP_quality_measure_specifications/Claims-Registry-Measures/2018_Measure_131_Claims.pdf
- ¹¹⁵ Additional measures that have been included in quality measures endorsed by the NQF include: Faces Pain Scale (FPS), McGill Pain Questionnaire (MPQ), Multidimensional Pain Inventory (MPI), Neuropathic Pain Scale (NPS), Numeric Rating Scale (NRS), Oswestry Disability Index (ODI), Roland Morris Disability Questionnaire (RMDQ), Verbal Descriptor Scale (VDS), Verbal Numeric Rating Scale (VNRS), Visual Analog Scale (VAS), and Patient-Reported Outcomes Measurement Information System (PROMIS).
- ¹¹⁶ Kozielska, M., Reddy, V. P., Johnson, M., Ridder, F. de, Vermeulen, A., Liu, J., Groothuis, G. M. M., Danhof, M., & Proost, J. H. (2013). Sensitivity of individual items of the Positive and Negative Syndrome Scale (PANSS) and items subgroups to differentiate between placebo and drug treatment in schizophrenia. *Schizophrenia Research*, *146*(1–3), 53–58.
- ¹¹⁷ Østergaard, S. D., Lemming, O. M., Mors, O., Correll, C. U., & Bech, P. (2016). PANSS-6: a brief rating scale for the measurement of severity in schizophrenia. *Acta Psychiatrica Scandinavica*, *133*(6), 436–444.
- ¹¹⁸ Østergaard, S. D., Foldager, L., Mors, O., Bech, P., & Correll, C. U. (2018). The Validity and Sensitivity of PANSS-6 in the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Study. *Schizophrenia Bulletin*, *44*(2), 453–462.

-
- ¹¹⁹ Østergaard, S. D., Opler, M. G. A., & Correll, C. U. (2017). Bridging the Measurement Gap Between Research and Clinical Care in Schizophrenia. *Innovations in Clinical Neuroscience*, 14(11–12), 68–72.
- ¹²⁰ Zanello, A., Berthoud, L., Ventura, J., & Merlo, M.C.G. (2013). The Brief Psychiatric Rating Scale (version 4.0) factorial structure and its sensitivity in the treatment of outpatients with unipolar depression. *International Journal of Methods in Psychiatric Research*, 3, 227-244.
- ¹²¹ Zanello, A., Berthoud, L., Ventura, J., & Merlo, M.C.G. (2013). Previously cited.
- ¹²² Correll, C. U., Galling, B., Pawar, A., Krivko, A., Bonetto, C., Ruggeri, M., Craig, T. J., Nordentoft, M., Srihari, V. H., Guloksuz, S., Hui, C. L. M., Chen, E. Y. H., Valencia, M., Juarez, F., Robinson, D. G., Schooler, N. R., Brunette, M. F., Mueser, K. T., Rosenheck, R. A., ... Kane, J. M. (2018). Comparison of Early Intervention Services vs Treatment as Usual for Early-Phase Psychosis: A Systematic Review, Meta-analysis, and Meta-regression. *JAMA Psychiatry*, 75(6), 555–565.
- ¹²³ Zanello, A., Berthoud, L., Ventura, J., & Merlo, M.C.G. (2013). Previously cited.
- ¹²⁴ Correll, C. U., Galling, B., Pawar, A., Krivko, A., Bonetto, C., Ruggeri, M., Craig, T. J., Nordentoft, M., Srihari, V. H., Guloksuz, S., Hui, C. L. M., Chen, E. Y. H., Valencia, M., Juarez, F., Robinson, D. G., Schooler, N. R., Brunette, M. F., Mueser, K. T., Rosenheck, R. A., ... Kane, J. M. (2018). Previously cited.
- ¹²⁵ Bech, P., Austin, S.F., Timmerby, N., Ban, T.A., & Møller, S.B. (2018). A clinimetric analysis of a BPRS-6 scale for schizophrenia severity. *Acta Neuropsychiatrica*, 30(4): 187–191.
- ¹²⁶ Altman, E.G., Hedecker, D., Peterson, J.L. & Davis, J.M. (1997). The Altman Self-Rating Mania Scale. *Biological Psychology*, 42(10): 948-955.
- ¹²⁷ Schmit, E. L., & Balkin, R. S. (2014). Evaluating Emerging Measures in the DSM-5 for Counseling Practice. *Professional Counselor*, 4(3): 216-231.
- ¹²⁸ Gideon, N., Hawkes, N., Mond, J., Saunders, R., Tchanturia, K., & Serpell, L. (2018). Correction: Development and Psychometric Validation of the EDE-QS, a 12 Item Short Form of the Eating Disorder Examination Questionnaire (EDE-Q). *PLOS ONE*, 13(11), e0207256.
- ¹²⁹ Garner, D. M., & Garfinkel, P. E. (1979). The Eating Attitudes Test: an index of the symptoms of anorexia nervosa. *Psychological Medicine*, 9(2), 273–279.
- ¹³⁰ Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The eating attitudes test: psychometric features and clinical correlates. *Psychological Medicine*, 12(4), 871–878.
- ¹³¹ Lee, S., Kwok, K., Liau, C., & Leung, T. (2002). Screening Chinese patients with eating disorders using the eating attitudes test in Hong Kong. *International Journal of Eating Disorders*, 32(1), 91–97.
- ¹³² Mintz, L. B., & O'Halloran, M. S. (2000). The Eating Attitudes Test: validation with DSM-IV eating disorder criteria. *Journal of Personality Assessment*, 74(3), 489–503.
- ¹³³ Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). Previously cited.
- ¹³⁴ Maloney, M. J., McGuire, J. B., & Daniels, S. R. (1988). Reliability Testing of a Children's Version of the Eating Attitude Test. *Journal of the American Academy of Child & Adolescent Psychiatry*, 27(5), 541–543.

-
- ¹³⁵ Storch, E. A., Kaufman, D. A. S., Bagner, D., Merlo, L. J., Shapira, N. A., Geffken, G. R., Murphy, T. K., & Goodman, W. K. (2007). Previously cited.
- ¹³⁶ Aldea, M. A., Geffken, G. R., Jacob, M. L., Goodman, W. K., & Storch, E. A. (2009). Further psychometric analysis of the Florida obsessive-compulsive inventory. *Journal of Anxiety Disorders*, *23*(1), 124-129.
- ¹³⁷ Rapp, A. M., Bergman, R. L., Piacentini, J., & McGuire, J. F. (2016). Evidence-Based Assessment of Obsessive-Compulsive Disorder. *Journal of Central Nervous System Disease*, *8*, 13–29.
- ¹³⁸ Hirst, K. P., & Moutier, C. Y. (2010). Postpartum Major Depression. *American Academy of Family Physicians*, *82* (8), 926-933.
- ¹³⁹ Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry: The Journal of Mental Science*, *150*, 782–786.
- ¹⁴⁰ Martin, C. R., & Redshaw, M. (2018). Establishing a coherent and replicable measurement model of the Edinburgh Postnatal Depression Scale. *Psychiatry Research*, *264*, 182–191.
- ¹⁴¹ Pilowsky, D. J., Wickramaratne, P., Talati, A., Tang, M., Hughes, C. W., Garber, J., Malloy, E., King, C., Cerda, G., Sood, A. B., Alpert, J. E., Trivedi, M. H., Fava, M., Rush, A. J., Wisniewski, S., & Weissman, M. M. (2008). Children of depressed mothers 1 year after the initiation of maternal treatment: findings from the STAR*D-Child Study. *The American Journal of Psychiatry*, *165*(9), 1136–1147.
- ¹⁴² Foster, C. E., Webster, M. C., Weissman, M. M., Pilowsky, D. J., Wickramaratne, P. J., Talati, A., Rush, A. J., Hughes, C. W., Garber, J., Malloy, E., Cerda, G., Kornstein, S. G., Alpert, J. E., Wisniewski, S. R., Trivedi, M. H., Fava, M., & King, C. A. (2008). Remission of Maternal Depression: Relations to Family Functioning and Youth Internalizing and Externalizing Symptoms. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *37*(4), 714–724.
- ¹⁴³ Evans, J., Heron, J., Francomb, H., Oke, S., & Golding, J. (2001). Cohort study of depressed mood during pregnancy and after childbirth. *BMJ (Clinical Research Ed.)*, *323*(7307), 257–260.
- ¹⁴⁴ Martin, C. R., & Redshaw, M. (2018). Previously cited.
- ¹⁴⁵ Evans, J., Heron, J., Francomb, H., Oke, S., & Golding, J. (2018). Previously cited.
- ¹⁴⁶ National Academy for State Health Policy. (n.d.) *Medicaid policies for maternal depression screening (MDS) during well-child visits, by state*. <https://healthychild.nashp.org/wp-content/uploads/2018/04/NASHP-MDS-Table-4.24.18.pdf>
- ¹⁴⁷ Byatt, N., Biebel, K., Cox J, Chaudron L, Ravech, M., & Straus, J. (2018). *MCPAP for Moms: A Primer for Massachusetts Pediatric Providers*. <https://www.mcpapformoms.org/Docs/MCPAP%20for%20Moms%20Primer%203-14-18.pdf>
- ¹⁴⁸ Byatt, N., Simas, T. A. M., Biebel, K., Sankaran, P., Pbert, L., Weinreb, L., Ziedonis, D., & Allison, J. (2018). PRogram In Support of Moms (PRISM): a pilot group randomized controlled trial of two approaches to improving depression among perinatal women. *Journal of Psychosomatic Obstetrics & Gynecology*, *39*(4), 297–306.

- ¹⁴⁹ National Academy for State Health Policy (NASHP). (2018) *Medicaid Policies for Maternal Depression Screening (MDS) During Well-Child Visits, By State*. <https://nashp.org/wp-content/uploads/2019/09/NASHP-MDS-Table-4.24.18.pdf>
- ¹⁵⁰ Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., Bullinger, M., Kaasa, S., Lepage, A., Prieto, L., & Sullivan, M. (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. *International Quality of Life Assessment. Journal of Clinical Epidemiology*, *51*(11), 1171–1178.
- ¹⁵¹ Hayes, C. J., Bhandari, N. R., Kathe, N., & Payakachat, N. (2017). Reliability and Validity of the Medical Outcomes Study Short Form-12 Version 2 (SF-12v2) in Adults with Non-Cancer Pain. *Healthcare (Basel, Switzerland)*, *5*(2).
- ¹⁵² Riddle, D. L., Lee, K. T., & Stratford, P. W. (2001). Use of SF-36 and SF-12 health status measures: a quantitative comparison for groups versus individual patients. *Medical Care*, *39*(8), 867–878.
- ¹⁵³ Alenzi, E. O., & Sambamoorthi, U. (2016). Depression treatment and health-related quality of life among adults with diabetes and depression. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *25*(6), 1517–1525.
- ¹⁵⁴ Wan, E.Y.F., Choi, E.P.H., Yu, E.Y.T., Chin, W.Y., Fung, C.S.C., Chan, A.K.C., & Lam, C.L.K. (2018). Evaluation of the internal and external responsiveness of Short Form-12 Health Survey version 2 (SF-12v2) in patients with type 2 diabetes mellitus. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *27*(9), 2459–2469.
- ¹⁵⁵ Huo, T., Guo, Y., Shenkman, E., & Muller, K. (2018). Assessing the reliability of the short form 12 (SF-12) health survey in adults with mental health conditions: a report from the wellness incentive and navigation (WIN) study. *Health and Quality of Life Outcomes*, *16*(1), 34.
- ¹⁵⁶ Iqbal, S. U., Rogers, W., Selim, A., Qian, S., Lee, A., Ren, X. S., Rothlander, J., Miller, d., & Kazis, L. (n.d.). *The Veterans RAND 12-Item Health Survey (VR-12): What it is and how it is used*. https://www.bu.edu/sph/files/2015/01/veterans_rand_12_item_health_survey_vr-12_2007.pdf
- ¹⁵⁷ Ell, K., Xie, B., Quon, B., Quinn, D. I., Dwight-Johnson, M., & Lee, P.-J. (2008). Previously cited.
- ¹⁵⁸ Vilagut, G., Forero, C. G., Pinto-Meza, A., Haro, J. M., Graaf, R. de, Bruffaerts, R., Kovess, V., Girolamo, G. de, Matschinger, H., Ferrer, M., & Alonso, J. (2013). The Mental Component of the Short-Form 12 Health Survey (SF-12) as a Measure of Depressive Disorders in the General Population: Results with Three Alternative Scoring Methods. *Value in Health*, *16*(4), 564–573.
- ¹⁵⁹ Fleishman, J.A., Cohen, J.W., Manning, W.G., & Kosinski, M. (2006). Using the SF-12 Health Status measure to improve predictions of medical expenditures: *Medical Care*, *44*(1), I-54-I-63.
- ¹⁶⁰ DeMuro, C., Clark, M., Doward, L., Evans, E., Mordin, M., & Gnanasakthy, A. (2013). Assessment of PRO Label Claims Granted by the FDA as Compared to the EMA (2006–2010). *Value in Health*, *16*(8), 1150–1155.
- ¹⁶¹ For more information regarding the selection of forms and versions of the SF instruments, see Optum, Inc’s® guidance on form selection at:
<http://campaign.optum.com/content/dam/optum/resources/Manual%20Excerpts/Which-Survey-To-Use.pdf>

-
- ¹⁶² Chopra, P. K., Chopra, P. K., Couper, J. W., & Herrman, H. (2004). The Assessment of Patients with Long-Term Psychotic Disorders: Application of the who Disability Assessment Schedule II. *Australian & New Zealand Journal of Psychiatry*, *38*(9), 753–759.
- ¹⁶³ Chwastiak, L. A., & Korff, M. V. (2003). Disability in depression and back pain: Evaluation of the World Health Organization Disability Assessment Schedule (WHO DAS II) in a primary care setting. *Journal of Clinical Epidemiology*, *56*(6), 507–514.
- ¹⁶⁴ Luciano, J. V., Ayuso-Mateos, J. L., Fernández, A., Serrano-Blanco, A., Roca, M., & Haro, J. M. (2010). Psychometric properties of the twelve item World Health Organization Disability Assessment Schedule II (WHO-DAS II) in Spanish primary care patients with a first major depressive episode. *Journal of Affective Disorders*, *121*(1), 52–58.
- ¹⁶⁵ Ustun, T. B., Kostanjsek, N., Chatterji, S., Rehm, J., & Organization, W. H. (2010). *Measuring health and disability : manual for WHO Disability Assessment Schedule (WHODAS 2.0)*. World Health Organization. <https://apps.who.int/iris/handle/10665/43974>
- ¹⁶⁶ Perini, S. J., Slade, T., & Andrews, G. (2006). Generic effectiveness measures: Sensitivity to symptom change in anxiety disorders. *Journal of Affective Disorders*, *90*(2), 123–130.
- ¹⁶⁷ Pösl, M., Cieza, A., & Stucki, G. (2007). Psychometric properties of the WHODASII in rehabilitation patients. *Quality of Life Research*, *16*(9), 1521–1531.
- ¹⁶⁸ International Consortium for Health Outcomes Measurement. (2016). *Standard set for depression and anxiety*. <https://www.ichom.org/medical-conditions/depression-anxiety/>
- ¹⁶⁹ Department of Defense, Department of Veterans Affairs & Department of Health and Human Services. (2016). Previously cited.
- ¹⁷⁰ American Psychological Association. (n.d). *WHODAS 2.0*. *American Psychiatric Association*. https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM5_WHODAS-2-Proxy-Administered.pdf
- ¹⁷¹ Johnson, J. G., Harris, E. S., Spitzer, R. L., & Williams, J. B. W. (2002). The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *30*(3), 196–204.
- ¹⁷² Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., Rockhill, C., & Katon, W. (2010). Evaluation of the Patient Health Questionnaire-9 Item for detecting major depression among adolescents. *Pediatrics*, *126*(6), 1117–1123.
- ¹⁷³ Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(1), 28–38.
- ¹⁷⁴ Spitzer, R.L., Kroenke, L., & Williams, J.B. (1999). Previously cited.

-
- ¹⁷⁵ Ginsburg, A.D., Stadem, P.S., Takala, C.R., et al. (2018). An examination of screening tools for collaborative care of adolescent depression. *Journal of Clinical Psychiatry*, 79(4):17m11543.
- ¹⁷⁶ Shippee, N. D., Mattson, A., Brennan, R., Huxsahl, J., Billings, M. L., & Williams, M. D. (2018). Effectiveness in Regular Practice of Collaborative Care for Depression Among Adolescents: A Retrospective Cohort Study. *Psychiatric Services (Washington, D.C.)*, 69(5), 536–541.
- ¹⁷⁷ American Psychological Association. (n.d). *Severity measure for depression—child age 11–17 (adapted from PHQ9 modified for Adolescents [PHQ-A])*. American Psychiatric Association. https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM5_Severity-Measure-For-Depression-Child-Age-11-to-17.pdf
- ¹⁷⁸ Richardson, L. P., McCauley, E., Grossman, D. C., McCarty, C. A., Richards, J., Russo, J. E., Rockhill, C., & Katon, W. (2010). Previously cited.
- ¹⁷⁹ Shippee, N. D., Mattson, A., Brennan, R., Huxsahl, J., Billings, M. L., & Williams, M. D. (2018). Previously cited.
- ¹⁸⁰ Cook, K. F., Jensen, S. E., Schalet, B. D., Beaumont, J. L., Amtmann, D., Czajkowski, S., Dewalt, D. A., Fries, J. F., Pilkonis, P. A., Reeve, B. B., Stone, A. A., Weinfurt, K. P., & Cella, D. (2016). Previously cited.
- ¹⁸¹ Jensen, R. E., Moinpour, C. M., Potosky, A. L., Lobo, T., Hahn, E. A., Hays, R. D., Cella, D., Smith, A. W., Wu, X.-C., Keegan, T. H. M., Paddock, L. E., Stroup, A. M., & Eton, D. T. (2017). Previously cited.
- ¹⁸² Lee, A. C., Driban, J. B., Price, L. L., Harvey, W. F., Rodday, A. M., & Wang, C. (2017). Previously cited.
- ¹⁸³ Purvis, T. E., Andreou, E., Neuman, B. J., Riley, L. H., & Skolasky, R. L. (2017). Previously cited.
- ¹⁸⁴ Schalet, B. D., Pilkonis, P. A., Yu, L., Dodds, N., Johnston, K. L., Yount, S., Riley, W., & Cella, D. (2016). Previously cited.
- ¹⁸⁵ Wohlfahrt, A., Bingham, C. O., Marder, W., Phillips, K., Bolster, M. B., Moreland, L. W., Zhang, Z., Neogi, T., & Lee, Y. C. (2019). Previously cited.
- ¹⁸⁶ Cotton, C. R., Peters, D. K., & Range, L. M. (1995). Psychometric properties of the Suicidal Behaviors Questionnaire. *Death Studies*, 19(4), 391–397.
- ¹⁸⁷ Brown, G. K. (2001). *A review of suicide assessment measures for intervention research with adults and older adults*. National Institute of Mental Health. <https://www.sprc.org/resources-programs/review-suicide-assessment-measures-intervention-research-adults-older-adults>
- ¹⁸⁸ Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment*, 8(4), 443–454.
- ¹⁸⁹ Wolraich, M. L., Lambert, W., Doffing, M. A., Bickman, L., Simmons, T., & Worley, K. (2003). Psychometric properties of the Vanderbilt ADHD diagnostic parent rating scale in a referred population. *Journal of Pediatric Psychology*, 28(8), 559–567.

- ¹⁹⁰ Wolraich, M. L., Bard, D. E., Neas, B., Doffing, M., & Beck, L. (2013). The psychometric properties of the Vanderbilt attention-deficit hyperactivity disorder diagnostic teacher rating scale in a community population. *Journal of Developmental and Behavioral Pediatrics, 34*(2), 83–93.
- ¹⁹¹ Kolko, D. J., Campo, J., Kilbourne, A. M., Hart, J., Sakolsky, D., & Wisniewski, S. (2014). Collaborative Care Outcomes for Pediatric Behavioral Health Problems: A Cluster Randomized Trial. *Pediatrics, 133*(4), e981–e992.
- ¹⁹² Subcommittee on Attention-Deficit/Hyperactivity Disorder, Steering Committee on Quality Improvement and Management. (2011). ADHD: Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics, 128*(5):1007–1022.
- ¹⁹³ American Academy of Pediatrics and National Initiative for Children’s Healthcare Quality. (2002). *Caring for children with ADHD: A resource toolkit for clinicians, 2nd edition*. <https://www.aap.org/en-us/pubserv/adhd2/Pages/kit/data/introframe.html>
- ¹⁹⁴ American Academy of Child & Adolescent Psychiatry. (n.d.). *Toolbox of Forms*. [Webpage]. https://www.aacap.org/AACAP/Member_Resources/AACAP_Toolbox_for_Clinical_Practice_and_Outcomes/Forms.aspx
- ¹⁹⁵ Leslie, L. K., Weckerly, J., Plemmons, D., Landsverk, J., & Eastman, S. (2004). Implementing the American Academy of Pediatrics attention-deficit/hyperactivity disorder diagnostic guidelines in primary care settings. *Pediatrics, 114*(1), 129–140.
- ¹⁹⁶ Kolko, D. J., Campo, J., Kilbourne, A. M., Hart, J., Sakolsky, D., & Wisniewski, S. (2014). Previously cited.
- ¹⁹⁷ Jacobson, J. H., Pullmann, M. D., Parker, E. M., & Kerns, S. E. U. (2019). Measurement Based Care in Child Welfare-Involved Children and Youth: Reliability and Validity of the PSC-17. *Child Psychiatry & Human Development, 50*(2), 332–345.
- ¹⁹⁸ Murphy, J. M., Bergmann, P., Chiang, C., Sturner, R., Howard, B., Abel, M. R., & Jellinek, M. (2016). The PSC-17: Subscale Scores, Reliability, and Factor Structure in a New National Sample. *Pediatrics, 138*(3).
- ¹⁹⁹ Andreas, J. B., & O’Farrell, T. J. (2017). Psychosocial Problems in Children of Women Entering Substance Use Disorder Treatment: A Longitudinal Study. *Addictive Behaviors, 65*, 193–197.
- ²⁰⁰ Kamin, H. S., McCarthy, A. E., Abel, M. R., Jellinek, M. S., Baer, L., & Murphy, J. M. (2015). Using a brief parent-report measure to track outcomes for children and teens with internalizing disorders. *Child Psychiatry and Human Development, 46*(6), 851–862.
- ²⁰¹ McCarthy, A., Asghar, S., Wilens, T., Romo, S., Kamin, H., Jellinek, M., & Murphy, M. (2016). Using a Brief Parent-Report Measure to Track Outcomes for Children and Teens with ADHD. *Child Psychiatry & Human Development, 47*(3), 407–416.
- ²⁰² Murphy, M., Kamin, H., Masek, B., Vogeli, C., Caggiano, R., Sklar, K., Drubner, S., Buonopane, R., Picciotto, M., Gold, J., & Jellinek, M. (2012). Using brief clinician and parent measures to track outcomes in outpatient child psychiatry: longer term follow-up and comparative effectiveness. *Child and Adolescent Mental Health, 17*(4), 222–230.

-
- ²⁰³ Stein, B. D., Jaycox, L. H., Kataoka, S. H., Wong, M., Tu, W., Elliott, M. N., & Fink, A. (2003). A mental health intervention for schoolchildren exposed to violence: a randomized controlled trial. *JAMA*, *290*(5), 603–611.
- ²⁰⁴ Murphy, J. M., Blais, M., Baer, L., McCarthy, A., Kamin, H., Masek, B., & Jellinek, M. (2015). Previously cited.
- ²⁰⁵ Hayek, M., Mackie, T. I., Mulé, C. M., Bellonci, C., Hyde, J., Bakan, J. S., & Leslie, L. K. (2014). A multi-state study on mental health evaluation for children entering foster care. *Administration and Policy in Mental Health and Mental Health Services Research*, *41*(4), 552–567.
- ²⁰⁶ Pourat, N., Zima, B., Marti, A., & Lee, C. (2017). *California Child Mental Health Performance Outcomes System: Recommendation report*. UCLA Center for Health Policy Research: Health Economics and Evaluation Research. [Report].
http://healthpolicy.ucla.edu/publications/Documents/PDF/2017/California_Child_Mental_Health_Performance_Outcomes_System_Recommendation_Report.pdf
- ²⁰⁷ Extensive information about the utilization of the PSC is available on the Massachusetts General Hospital Department of Psychiatry webpage is available here:
<https://www.massgeneral.org/psychiatry/services/treatmentprograms.aspx?id=2088&display=pediatric>
- ²⁰⁸ National Quality Forum. (2013, April). *NQF: NQF Removes Time-Limited Endorsement for Four Measures; Measures Now Have Endorsed Status*.
https://www.qualityforum.org/News_And_Resources/Press_Releases/2013/NQF_Removes_Time-Limited_Endorsement_for_Four_Measures;_Measures_Now_Have_Endorsed_Status.aspx
- ²⁰⁹ State of California Health and Human Services Agency, Department of Health Care Services. (2017). *MHSUDS Information Notice NO: 17-052: Early and Periodic Screening, Diagnostic and Treatment (EPSDT) – specialty mental health services performance outcomes system functional assessment tools for children and youth*. Retrieved from https://www.dhcs.ca.gov/services/MH/Documents/FMORB/Info_Notice_17-052_POS_Functional_Assessment_Tool.pdf
- ²¹⁰ Andreas, J. B., & O’Farrell, T. J. (2017). Previously cited.
- ²¹¹ Kamin, H. S., McCarthy, A. E., Abel, M. R., Jellinek, M. S., Baer, L., & Murphy, J. M. (2015). Previously cited.
- ²¹² McCarthy, A., Asghar, S., Wilens, T., Romo, S., Kamin, H., Jellinek, M., & Murphy, M. (2016). Previously cited.
- ²¹³ Murphy, M., Kamin, H., Masek, B., Vogeli, C., Caggiano, R., Sklar, K., Drubner, S., Buonopane, R., Picciotto, M., Gold, J., & Jellinek, M. (2012). Previously cited.
- ²¹⁴ Stein, B. D., Jaycox, L. H., Kataoka, S. H., Wong, M., Tu, W., Elliott, M. N., & Fink, A. (2003). Previously cited.
- ²¹⁵ Andreas, J. B., & O’Farrell, T. J. (2017). Previously cited.
- ²¹⁶ Kamin, H. S., McCarthy, A. E., Abel, M. R., Jellinek, M. S., Baer, L., & Murphy, J. M. (2015). Previously cited.
- ²¹⁷ McCarthy, A., Asghar, S., Wilens, T., Romo, S., Kamin, H., Jellinek, M., & Murphy, M. (2016). Previously cited.
- ²¹⁸ Murphy, M., Kamin, H., Masek, B., Vogeli, C., Caggiano, R., Sklar, K., Drubner, S., Buonopane, R., Picciotto, M., Gold, J., & Jellinek, M. (2012). Previously cited.

- ²¹⁹ Stein, B. D., Jaycox, L. H., Kataoka, S. H., Wong, M., Tu, W., Elliott, M. N., & Fink, A. (2003). Previously cited.
- ²²⁰ Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(4), 545–553.
- ²²¹ Birmaher, B., Brent, D. A., Chiappetta, L., Bridge, J., Monga, S., & Baugher, M. (1999). Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(10), 1230–1236.
- ²²² Boyd, R. C., Ginsburg, G. S., Lambert, S. F., Cooley, M. R., & Campbell, K. D. M. (2003). Screen for Child Anxiety Related Emotional Disorders (SCARED): Psychometric Properties in an African American Parochial High School Sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(10), 1188–1196.
- ²²³ Runyon, K., Chesnut, S. R., & Burley, H. (2018). Screening for childhood anxiety: A meta-analysis of the screen for child anxiety related emotional disorders. *Journal of Affective Disorders*, 240, 220–229.
- ²²⁴ Birmaher, B., Brent, D. A., Chiapetta, L., Bridge, J., Monga, S., & Baugher, M. (2015). *Screen for Childhood Anxiety Related Emotional Disorders (SCARED)*. The California Evidence-Based Clearinghouse for Child Welfare. <http://www.cebc4cw.org/assessment-tool/screen-for-childhood-anxiety-related-emotional-disorders-scared/>
- ²²⁵ Cook, K. F., Jensen, S. E., Schalet, B. D., Beaumont, J. L., Amtmann, D., Czajkowski, S., Dewalt, D. A., Fries, J. F., Pilkonis, P. A., Reeve, B. B., Stone, A. A., Weinfurt, K. P., & Cella, D. (2016). Previously cited.
- ²²⁶ Jensen, R. E., Moinpour, C. M., Potosky, A. L., Lobo, T., Hahn, E. A., Hays, R. D., Cella, D., Smith, A. W., Wu, X.-C., Keegan, T. H. M., Paddock, L. E., Stroup, A. M., & Eton, D. T. (2017). Previously cited.
- ²²⁷ Lee, A. C., Driban, J. B., Price, L. L., Harvey, W. F., Rodday, A. M., & Wang, C. (2017). Previously cited.
- ²²⁸ Purvis, T. E., Andreou, E., Neuman, B. J., Riley, L. H., & Skolasky, R. L. (2017). Previously cited.
- ²²⁹ Schalet, B. D., Pilkonis, P. A., Yu, L., Dodds, N., Johnston, K. L., Yount, S., Riley, W., & Cella, D. (2016). Previously cited.
- ²³⁰ Wohlfahrt, A., Bingham, C. O., Marder, W., Phillips, K., Bolster, M. B., Moreland, L. W., Zhang, Z., Neogi, T., & Lee, Y. C. (2019). Previously cited.
- ²³¹ Daviss, W. B., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Criterion validity of the Mood and Feelings Questionnaire for depressive episodes in clinic and non-clinic subjects. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 47(9), 927–934.
- ²³² Kent, L., Vostanis, P., & Feehan, C. (1997). Detection of major and minor depression in children and adolescents: evaluation of the Mood and Feelings Questionnaire. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 38(5), 565–573.
- ²³³ Wood, A., Kroll, L., Moore, A., & Harrington, R. (1995). Properties of the mood and feelings questionnaire in adolescent psychiatric outpatients: a research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 36(2), 327–334.

-
- ²³⁴ Banh, M. K., Crane, P. K., Rhew, I., Gudmundsen, G., Stoep, A. V., Lyon, A., & McCauley, E. (2012). Measurement Equivalence Across Racial/Ethnic Groups of the Mood and Feelings Questionnaire for Childhood Depression. *Journal of Abnormal Child Psychology, 40*(3), 353–367.
- ²³⁵ Daviss, W. B., Birmaher, B., Melhem, N. A., Axelson, D. A., Michaels, S. M., & Brent, D. A. (2006). Previously cited.
- ²³⁶ Sund, A. M., Larsson, B., & Wichstrøm, L. (2001). Depressive symptoms among young Norwegian adolescents as measured by The Mood and Feelings Questionnaire (MFQ). *European Child & Adolescent Psychiatry, 10*(4), 222–229.
- ²³⁷ Thabrew, H., Stasiak, K., Bavin, L.-M., Frampton, C., & Merry, S. (2018). Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents. *International Journal of Methods in Psychiatric Research, 27*(3), e1610.
- ²³⁸ Thabrew, H., Stasiak, K., Bavin, L.-M., Frampton, C., & Merry, S. (2018). Previously cited.
- ²³⁹ Gonzales, R., Ang, A., Murphy, D. A., Glik, D. C., & Anglin, M. D. (2014). Substance use recovery outcomes among a cohort of youth participating in a mobile-based texting aftercare pilot program. *Journal of Substance Abuse Treatment, 47*(1), 20–26.
- ²⁴⁰ Gonzales, R., Hernandez, M., Murphy, D. A., & Ang, A. (2016). Youth Recovery Outcomes at 6 and 9 Months Following Participation in a Mobile Texting Recovery Support Aftercare Pilot Study. *The American Journal on Addictions / American Academy of Psychiatrists in Alcoholism and Addictions, 25*(1), 62–68.
- ²⁴¹ Garner, D. M., & Garfinkel, P. E. (1979). Previously cited.
- ²⁴² Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). Previously cited.
- ²⁴³ Lee, S., Kwok, K., Liau, C., & Leung, T. (2002). Previously cited.
- ²⁴⁴ Mintz, L. B., & O'Halloran, M. S. (2000). Previously cited.
- ²⁴⁵ Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). Previously cited.
- ²⁴⁶ Maloney, M. J., McGUIRE, J. B., & Daniels, S. R. (1988). Previously cited.
- ²⁴⁷ Maloney, M. J., McGUIRE, J. B., & Daniels, S. R. (1988). Previously cited.
- ²⁴⁸ Storch, E. A., Khanna, M., Merlo, L. J., Loew, B. A., Franklin, M., Reid, J. M., Goodman, W. K., & Murphy, T. K. (2009). Previously cited.
- ²⁴⁹ Rapp, A. M., Bergman, R. L., Piacentini, J., & McGuire, J. F. (2016). Previously cited.
- ²⁵⁰ Dampier, C., Jaeger, B., Gross, H. E., Barry, V., Edwards, L., Lui, Y., DeWalt, D. A., & Reeve, B. B. (2016). Responsiveness of PROMIS® Pediatric Measures to Hospitalizations for Sickle Pain and Subsequent Recovery. *Pediatric Blood & Cancer, 63*(6), 1038–1045.
- ²⁵¹ Howell, C. R., Thompson, L. A., Gross, H. E., Reeve, B. B., DeWalt, D. A., & Huang, I.-C. (2016). Responsiveness to Change in PROMIS(®) Measures among Children with Asthma: A Report from the PROMIS(®) Pediatric Asthma

Study. *Value in Health: The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 19(2), 192–201.

²⁵² Reeve, B. B., Edwards, L. J., Jaeger, B. C., Hinds, P. S., Dampier, C., Gipson, D. S., Selewski, D. T., Troost, J. P., Thissen, D., Barry, V., Gross, H. E., & DeWalt, D. A. (2018). Assessing responsiveness over time of the PROMIS[®] pediatric symptom and function measures in cancer, nephrotic syndrome, and sickle cell disease. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*, 27(1), 249–257.

²⁵³ Selewski, D. T., Troost, J. P., Cummings, D., Massengill, S. F., Gbadegesin, R. A., Greenbaum, L. A., Shatat, I. F., Cai, Y., Kapur, G., Hebert, D., Somers, M. J., Trachtman, H., Pais, P., Seifert, M. E., Goebel, J., Sethna, C. B., Mahan, J. D., Gross, H. E., Herreshoff, E., ... Gipson, D. S. (2017). Responsiveness of the PROMIS[®] measures to changes in disease status among pediatric nephrotic syndrome patients: a Midwest pediatric nephrology consortium study. *Health and Quality of Life Outcomes*, 15.

²⁵⁴ Dampier, C., Jaeger, B., Gross, H. E., Barry, V., Edwards, L., Lui, Y., DeWalt, D. A., & Reeve, B. B. (2016). Previously cited.

²⁵⁵ Reeve, B. B., Edwards, L. J., Jaeger, B. C., Hinds, P. S., Dampier, C., Gipson, D. S., Selewski, D. T., Troost, J. P., Thissen, D., Barry, V., Gross, H. E., & DeWalt, D. A. (2018). Previously cited.

²⁵⁶ Selewski, D. T., Troost, J. P., Cummings, D., Massengill, S. F., Gbadegesin, R. A., Greenbaum, L. A., Shatat, I. F., Cai, Y., Kapur, G., Hebert, D., Somers, M. J., Trachtman, H., Pais, P., Seifert, M. E., Goebel, J., Sethna, C. B., Mahan, J. D., Gross, H. E., Herreshoff, E., ... Gipson, D. S. (2017). Previously cited.

²⁵⁷ Reeve, B. B., Edwards, L. J., Jaeger, B. C., Hinds, P. S., Dampier, C., Gipson, D. S., Selewski, D. T., Troost, J. P., Thissen, D., Barry, V., Gross, H. E., & DeWalt, D. A. (2018). Previously cited.

²⁵⁸ Howell, C. R., Thompson, L. A., Gross, H. E., Reeve, B. B., DeWalt, D. A., & Huang, I.-C. (2016). Previously cited.

²⁵⁹ Reeve, B. B., Edwards, L. J., Jaeger, B. C., Hinds, P. S., Dampier, C., Gipson, D. S., Selewski, D. T., Troost, J. P., Thissen, D., Barry, V., Gross, H. E., & DeWalt, D. A. (2018). Previously cited.

²⁶⁰ The APA encourages additional research and evaluation of these instruments for their diagnostic and treatment monitoring utility.

²⁶¹ PROMIS. (2013, May). *Instrument development and validation scientific standards version 2.0 (revised May 2013)*. http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf