Ment Health Fam Med (2019) 15: 884-892

Review Article

The Use of Behavioral Health Rating Scales for Primary Care Providers

Jessica Jeffrey
University of California, Los Angeles, USA

Nastassia J. Hajal
University of California, Los Angeles, USA

Alexandra Klomhaus
University of California, Los Angeles, USA

Lauren Marlotte
University of California, Los Angeles, USA

Patricia Lester
University of California, Los Angeles, USA

ABSTRACT

In recent years, there has been a shift in U.S. healthcare to increase accountability in care by requiring objective measurement and documentation of disease state improvement for payment. While the implementation of objective measures for physical disease has been relatively straightforward, integration of ongoing, objective measurement of behavioral health outcomes in primary care settings has been slower. This slow uptake may be at least partially due to the limited training that most primary care providers receive in behavioral health assessment. In this manuscript we review existing literature on the effectiveness of behavioral health-related measurement-based care (MBC), discuss the challenges associated with its adoption in primary care settings, review recommended behavioral health rating scales, and present a case example to illustrate how behavioral health MBC may improve care. Finally, we discuss recommendations for incorporating behavioral health MBC into practice in a manner that is beneficial for the patient, provider, and health system.

MeSh Headings/ Keywords: Primary care providers, Behavioral health, Healthcare.

Introduction

With the goals of improving patient outcomes while controlling healthcare costs, U.S. healthcare systems have been shifting their payment systems from volume to value of care provided [1,2]. One way this is being pursued is by requiring objective measures for documentation of disease state improvement for payment. Implementation of objective measures for physical disease has been relatively straightforward (i.e., third-party payer requirement of improvements in HbA1c scores for diabetes patients) but adoption of objective measures of behavioral health has been slow [3,4]. Behavioral health symptoms are the chief complaint in up to 25% of adults [5,6] and 20% of youth in primary care settings [7], and more Americans receive behavioral healthcare from primary care providers than behavioral health specialists [8-10]. Despite this fact, medical professionals generally have limited training in the use and implementation of behavioral health assessment [11,12]. Measurement-based care (MBC) of behavioral health includes screening, regular use of rating scales to assess symptom severity, and a focus on treatment-to-target symptom reduction and/or remission as measured by rating scale scores [13]. Rating scales in MBC are used as a starting point for clinical evaluation, not as a substitute for clinical judgment [4]. The slow adoption of MBC is unfortunate given the positive effects of MBC on treatment outcomes, patient-provider communication, and patient engagement in care [14-17].

This article focuses on the use of rating scales to guide behavioral health-related MBC practice recommendations for primary care physicians. Specifically, we review the evidence of the effectiveness of MBC, examine challenges to adoption, and provide recommendations for incorporating MBC into busy primary care practices. We also review psychometrically validated behavioral health rating scales that are available within the public domain and present a case example to illustrate how MBC may improve health outcomes. Finally, we discuss the benefits of MBC in the context of health systems.

Why Use Behavioral Health Rating Scales?

MBC is associated with positive outcomes in behavioral, primary, and specialty healthcare settings. In diverse areas of behavioral health treatment (including individual, couples, and home-based therapies), MBC is associated with greater improvement in symptoms [18-20] (particularly for complex cases [14]) and lower likelihood of treatment failure [20]. Furthermore, MBC may be associated with faster improvement in symptoms. One study showed that patients whose clinicians reviewed the results of biweekly symptom rating scales showed a treatment response after fewer sessions than patients whose clinicians did not view the reports [21]. These effects are not unique to behavioral health settings; when implemented in primary and specialty medical care settings, MBC is associated with improvement in behavioral health symptoms and quality of life [14,22]. There is also potential for MBC to enhance treatment of physical symptoms, given that behavioral health problems are associated with unexplained somatic symptoms, poorer control of chronic medical conditions, and greater use of acute medical services [9]. In fact, the application of MBC to clinical practice is a key component of integrated care programs, which have been gaining in popularity due improved patient outcome and ability to reduce overall health care costs [9,23].

There are several possible mechanisms for these positive treatment effects. MBC techniques that use standardized measures enable providers to systematically examine their patients' functioning relative to same gender, developmental level, and clinical population peers [24], which can be extremely helpful not only in initial diagnosis and treatment planning, but also in routinely monitoring whether their chosen course of treatment is making clinically significant change [4,24]. The failure to identify patients who are not responding to treatment contributes to patients being maintained on ineffective treatments [25]. Importantly, MBC facilitates identification of non-responders, enabling clinicians to alter the treatment plan in a timely fashion [13]. It also facilitates detection of residual symptoms (e.g., a patient whose depressive symptoms show marked improvement overall, but who still experiences poor appetite, low energy, and hypersomnia). Awareness of lingering subclinical symptoms allows providers to address them to prevent relapse. Relatedly, many providers do not regularly and systematically assess domains that are not the main focus of treatment, but that may have an impact on intervention outcome. For example, without behavioral health MBC, a primary care provider might miss that a cardiology patient's inability to lose the recommended amount of weight is due to not to a physical problem or willful non-adherence, but to behavioral withdrawal symptoms of depression. MBC enables primary care providers to assess multiple areas of behavioral functioning to identify problems, whether expected or not. This allows them to structure the appointment to address the areas of most need (e.g., providing a referral to a cognitive-behavioral therapist to treat depression, assessing for harm and engaging in safety planning) as opposed to spending time providing less relevant education or intervention [26].

MBC also influences several processes over the course of treatment that affects outcome, including improved patient-provider communication [14], treatment experience [27], and treatment persistence [19]. Including a rating scale as part of ongoing assessment of functioning provides multiple modes of communication for patients to share information with providers. Patients may provide more accurate information when answering questions through written measures [26,28] versus face-to-face interview, particularly when first getting to know a provider or when regarding sensitive topics [29]. Additionally, rating scales help patients to more effectively communicate to their providers when treatments do not seem to be working [4]. When primary care providers engage in MBC, patients' confidence in their providers' diagnoses increases and patients share that they feel more validated [29]. In addition to improving patient-provider communication, MBC improves patients' understanding of their condition and course of treatment [29]. Answering the same questions on a regular basis may increase patients' awareness of their thoughts, feelings, and behaviors, which in and of itself is a core component of empirically supported behavioral health interventions. Finally, asking patients to actively participate in treatment monitoring has potential to empower them and increase sense of self-efficacy, promoting physical and behavioral health [30-32] and improve treatment adherence [33,34].

Implementing Rating Scales in Primary Care Practices

Challenges to implementing MBC include those related to clinic workflow within busy primary care settings (e.g. staff required to administer rating scales, score and interpret results), financial constraints (cost of measures), and staffs' comfort introducing behavioral health rating scales to patients [35].

To foster adoption, rating scales must be integrated seamlessly with the clinic workflow. Many systems have found the administration of rating scales in the waiting room, prior to an appointment with a provider, as effective for achieving MBC [36,37]. Although rating scales have traditionally been administered via paper-and-pencil, technological innovations (such as online administration and completion of scales on hand-held devices) have increased the efficiency of collection patient measures [4]. Web-based rating scale delivery systems enable patient responses to be scored and immediately provided to primary care providers so they can deliver customized clinical guidance based on real-time patient results [36]. Importantly, systems that provide feedback to providers in a structured manner during the clinical encounter are most effective for improving patient outcomes [4]. Further improvements to workflow occur when web-based rating scale delivery systems are integrated within the primary care providers' electronic health record. Integration with the electronic health record not only allows for efficient and timely documentation of scored

results, but also coordination of care between providers, and tracking of symptoms overtime to assess treatment response [36]. Web-based rating scale administration appears to be preferable to patients [38,39] with no negative impact on valid data collection [40]. Consideration must also be given to rating scale length to prevent fatiguing patients or interfering in the time allotted for the provider's examination. Importantly, the adoption of behavioral health integrated care models, of which MBC is a key component [13], into primary care settings may allow for the creation of robust workflows to obtain behavioral health rating scales.

Many widely used behavioral health measures are expensive, presenting a logistic barrier for clinics that may otherwise be equipped and motivated to employ MBC. Fortunately, several psychometrically sound measures are available within the public domain and may be downloaded and administered repeatedly for clinical care without charge. Several large health care systems have invested in MBC programs using a web-based program and/or the electronic health record [4,36]. Although smaller practices may face challenges adopting comprehensive electronic systems due to cost constraints [41], there are some free online resources for administration and scoring of certain measures (e.g., Strengths and Difficulties Questionnaire) [42].

To maximize the effectiveness of MBC, administrative staff training should be integrated in the process, as staff may be asked to introduce and provide patients with the rating scales. An important aspect of training includes attentiveness to the cultural shift required for staff to gain confidence discussing the purpose of rating scales with patients [36]. Staff may be instructed that it is often helpful to begin by explaining to the patient that completing behavioral health measures is similar to having their temperature or blood pressure taken at the start of a medical appointment; in other words, these rating scales will help to capture their "behavioral health vital signs" [36]. Staff should also feel comfortable describing the approximate duration for completion, confidentiality and its limits, and security of responses within the health record.

Choosing Behavioral Health Rating Scales for Clinical Practice

To be useful in clinical practice, rating scales must be psychometrically sound. Each measure should be reliable (i.e., items within each scale should assess the same construct and do so consistently) and valid (i.e., the scale should measure the true underlying condition it was developed to assess in the target population) [43]. Of particular importance introducing rating scales into new settings is the concept of face validity, which indicates the degree to which a measure is subjectively viewed as related to the overarching construct of interest. Patients, providers, and administrative staff are more likely to use rating scale assessments if they are viewed as relevant to patient care [44,45].

To maximize utility, rating scales used in clinical practice should be standardized. Standardization involves administering a measure to large sample of a certain population (e.g., a community sample of preschool-aged boys in the U.S.), enabling providers to see how a patient scores in relation to socio-demographic peers [24]. For example, an internist who administers a standardized alcohol use scale to an 18-year-old female patient can gauge how the patient's drinking compares to other young adult females. This process is what enables us to determine whether a patient is scoring within normal limits (typically <84th percentile), or in the elevated (84th-98th percentile) or clinically significant range (>98th percentile) based on where his/her score falls in relation to the larger population. Although these benchmarks have limitations, they are useful in treatment planning and promote continuity of care by facilitating communication between professionals.

To be clinically useful, rating scales must also have adequate levels of sensitivity and specificity. Specificity refers to the proportion of people without a disorder who test or screen negative for the disorder (i.e. a measure's ability to correctly identify those without the disorder), while sensitivity refers to the proportion of people with a disorder who test or screen positive (i.e. a measure's ability to correctly identify those with the disorder). Moreover, use of a rating scale for MBC necessitates sensitivity to change [46], defined as a measure's ability to differentiate between individuals who improved, remained unchanged, or deteriorated in functioning by comparing multiple administrations of the measure over time.

Recommended Measures

The measures described in Tables 1-3 cover a broad range of behavioral health symptoms and related constructs, including symptoms of depression, anxiety, traumatic stress, externalizing behaviors, substance use, and contextual factors. Rating scale measures for both pediatric and adult populations are included. These particular rating scales were chosen based on being psychometrically valid, freely available within the public domain, and clinically indicated for screening and/or monitoring of treatment progress. While all were designed and are used by psychologists and psychiatrists specializing in standardized behavioral health assessment, some were developed specifically for primary care settings.

When choosing scales for use in clinical practice, it is critical to consider that psychometric characteristics may vary depending on demographic and clinical factors. Several of the rating scales listed in Tables 1-3 have multiple reliability, sensitivity, and specificity statistics displayed; e.g., the adult PTSD measures listed (Table 2) were originally developed for military-connected individuals and have since been normed in civilian populations. Because psychometric characteristics depend on the specific norming populations (e.g., Veteran vs. civilian, male vs. female, community vs. clinic-referred), the tables display multiple statistics for some scales (e.g., PC-PTSD, AUDIT).

Table 1: Rating scales designed to assess depression and anxiety symptoms.

Measure	Target Population (Reporter)	Number of Items	Time Period Assessed	Response Options	Reliability	Sensitivity	Specificity	Screening (S) and/ or Monitoring (M)
Depression								
Patient Health Questionnaire (PHQ-9) [49]	Adults & Adolescents (Self-report)	9[1]	Past 2 weeks	Likert, 0 (not at all) to 3 (nearly every day)	α = .86 -.89	88%	88%	S + M
Mood & Feelings Questionnaire (MFQ) [50-52][2]	Child (Self-report for ages 7+ years)	13- & 33-item versions	Past 2 weeks	Likert, 0 (not true) to 2 (true)	α =.94 & α = .84	78% & 66%	78% & 61%	S + M
	Child (Parent-report)	13- & 33-item versions	Past 2 weeks	Likert, 0 (not true) to 2 (true)	α = .92 & α = .84	63% & 66%	61% & 66%	S + M
Anxiety								
Generalized Anxiety Disorder-7 (GAD-7) [53]	Adult (Self-report)	7	Past 2 weeks	Likert, 0 (not at all) to 3 (nearly every day)	α = 0.92	89%	82%	S + M
Screen for Child Anxiety Related Disorders (SCARED) [54,56][3]	Child (Self-report for ages 7+ years)	5-, 38-, and 41-item versions	Past 3 months	Likert, 0 (Not True or Hardly Ever True) to 2 (Very True/ Often True)	α = .74−.93 & α = .78−.90	67-80% & 71%	50-84% & 61-71%[4]	S + M
	Child (Parent-report)							S + M

[1]Plus 1 item on impairment in functioning, answered only if at least 1 of items 1-9 is endorsed.
[2]Multiple values for reliability, sensitivity, and specificity refer to 13- and 33-item forms, respectively.
[3]Multiple values for reliability, sensitivity, and specificity refer to 38- and 41-item forms, respectively.
[4]Ranges depend on comparison group, i.e., no anxiety, disruptive behavior disorder, or depression, on factor/subscale, and the reporter (parent vs. child).
* M indicates that measure was designed to be used as a monitoring tool or has demonstrated effectiveness as a monitoring tool in previous research.

Table 2: Rating scales designed to assess traumatic stress and externalizing behavior symptoms.

Measure	Target Population (Reporter)	Number of Items	Time Period Assessed	Response Options	Reliability	Sensitivity	Specificity	Screening (S) and/or Monitoring (M)
Traumatic Stress								
Primary Care PTSD Scale (PC-PTSD) [57-59][1]	Adult (Self-report)	5	Past month	Binary (Yes/No)	None	95%, 78%, & 85%	85%, 87% & & 82%[2]	S
PTSD Checklist (PCL-5) [60,61][3]	Adult (Self-report)	20	Past month	Likert, 0 (not at all) to 4 (extremely)	α_{vet} = 0.96 α_{civ} = 0.94	88% & 66-77%	69% & 95-97%[4]	S + M
Child PTSD Symptom Scale for DSM-V [62]	Child (Self-report for 8-18 years)	6- and 27-item versions	Past month	Likert, 0 (not at all) to 4 (6 or more times a week/ almost always)	α = .63-.92[5]	.88-93%	82%	S
Child Stress Disorders Checklist [63]	Child (parent-report for children 2-18 years)	36	Past month	Likert, 0 (Not True) to 2 (Very True or Often True)	α = .84	None	None	S + M
Externalizing Behaviors								
Adult ADHD Self-Report Scale (ASRS) [64-66][6]	Adult (Self-report)	6- and 18-item versions	Past 6 months	Likert, 0 (Never) to 4 (Very Often)	α = .63-.72 & α = .88	68.7% & 56.3%	99.5% & 98.3%	6-item: S 18-item: S + M
SNAP-IV [67,68][7]	Child (Teacher- and Parent-report)	18-, 26- and 90-item versions	None	Likert, 0 (Not at All) to 3 (Very Much)	α = .92-.97[8] & α = .79-.94	82.3%	82.4%[9]	S

[1]DSM-IV and DSM-5 versions available. Multiple values for sensitivity and specificity refer to Veteran, active duty military, and civilian populations, respectively.
[2]Military and Veteran values correspond to DSM-5.
[3]Multiple values for reliability, sensitivity, and specificity refer to Veteran and civilian populations, respectively.
[4]Range depends on presumed prevalence.
[5]Range depends on form (short- versus long-) and subscale (for long-form); total, intrusion, avoidance, changes in cognition and mood, and arousal and hyperactivity).
[6]Multiple values for reliability, sensitivity, and specificity refer to 6- and 18-item forms, respectively.
[7]Multiple values for reliability, sensitivity, and specificity refer to teacher- and parent-report forms, respectively.
[8]Range depends on subscale: inattention and hyperactivity/impulsivity.
[9]Based on a study from Spanish-speaking youth.

Table 3: Rating scales designed to assess substance use, general child behaviors, and contextual factors.

Measure	Target Population (Reporter)	Number of Items	Time Period Assessed	Response Options	Reliability	Sensitivity	Specificity	Screening (S) and/or Monitoring (M)
Substance Use								
CRAFFT [69,70]	Adolescent (Self-report)	Part A: 3 Part B: 6	Past 12 months	Binary (Yes/No)	α = .68	79-91%	93-97%[1]	S
DAST-10 [71,72]	Adult (Self-report)	10	Past 12 months	Binary (Yes/No)	α = .86	100%	84%	S
AUDIT [73,74]	Adult (Self-report)	10	Not specified[2]	Likert, 0 (Never) to 4 (Daily/Almost Daily)	α = .80-.90	92%	94%	S
Child Difficulties (Broadband Measure) & Prosocial Behaviors								
Strengths & Difficulties Questionnaire (SDQ [75])	Child (Parent report)	25	Past 6 months[3]	Likert, 0 (Not True) to 2 (Certainly True)	α = .57-.85[4]	25-74%	91-96%	S + M
	Child (Self-report)	25	Past 6 months	Likert, 0 (Not True) to 2 (Certainly True)	α = .41-.81	13-40%	89-97%	S + M
Contextual Factors								
McMaster Family Assessment Device [76,77]	Adult and Adolescent (Self-report)	12-, 53[5]-, and 60-item versions	Not specified	Likert, 1 (Strongly Agree) to 4 (Strongly Disagree)	α = .72-.92	57-83%	60-79%[6]	S + M
Work and Social Adjustment Scale [78,79]	Adult (Self-report)	5	Not specified	Likert, 0 (not at all impaired) to 8 (very severely impaired)	α = .70-.94	88%	78%	S + M

[1]Range depends on what kind of use is being evaluated according to DSM-V: problem use, or SUD.
[2]The measure includes questions that consider frequency of use over days/weeks/months/year, with a set of questions that consider "During the past year…".
[3]Baseline assessment considers the past 6 months; follow-up assessments consider the past month.
[4]Range depends on the subscale: total difficulties, emotional, conduct, hyperactivity, peer problems, prosocial, and total impact.
[5]The original FAD was 53-items, which is what the validity and reliability statistics are based on. There is also a more recent 60-item version. The 12-item general functioning subscale may be used on its own as a brief screening measure.
[6]Range depends on subscale (Problem Solving, Communication, Roles, Affective Responsiveness, Affective Involvement, Behavior Control, General Functioning).

Other factors that may warrant consideration include alignment with diagnostic criteria, grading of symptom severity, and multiple forms. Many measures are structured to align with diagnostic criteria for various psychological disorders and some may have the ability to further provide information about symptom severity (e.g., PHQ). This is helpful for tracking treatment progress over time, i.e., if a patient still meets diagnostic criteria but the severity reduces from severe to mild. For example, PHQ-9 scores can effectively and accurately differentiate between patients with persistent major depression, and partial and full depression remission [47].

Many measures offer a "short form" (e.g., SCARED). Although long forms typically offer richer information and boast stronger psychometric characteristics, short forms may be optimal for patient satisfaction (particularly for child-reported measures) and efficiency of administration and scoring (particularly if scales are completed via paper and pencil as opposed to electronically). Another option is to provide a "gating" measure, where responses to a few initial questions determine whether a longer form will be completed. For example, the 5-item PC-PTSD may be used as a gating measure for longer PTSD scales, such as the 20-item PCL. Specifically, patients obtaining a score of <3 on the PC-PTSD would not be prompted to complete further PTSD measures, but those with a score of >3 would be prompted to complete a more in-depth screening.

Case Example: Depressive Symptoms Impacting Diabetes Treatment

The following case example illustrates the use of self-report behavioral health rating scales by two primary care providers who work together within an office-based practice. The case was created for the purposes of exemplifying the recommendations in this paper.

Dr. J is a primary care provider for Cindy, a 45-year-old mother. Dr. J works in the same practice as Dr. K, a pediatrician who sees Cindy's 11-year-old son Sam. Cindy is followed by Dr. J for type 2 diabetes. Cindy's diabetes had previously been well-controlled by diet, exercise, and medication, but her HbA1c levels have increased over the past year (currently 9%). Cindy reports she is lethargic and is no longer exercising; she previously walked three times per week at lunch with a friend but has recently felt too tired to continue. She admits she has also been forgetting to consistently take her prescribed Metformin to manage her diabetes. Dr. J counsels Cindy about checking her blood sugar levels, reviews the importance of exercising regularly, and encourages medication adherence. Dr. J notes the flagging system used in the office that indicates Cindy completed

a PHQ-9 in the waiting room prior to her appointment. Upon review of the PHQ-9, Dr. J notes elevated depressive symptoms (PHQ-9 score indicating moderate depression, no endorsement of suicidal ideation). Dr. J's clinical interview confirms that over the past month Cindy has been experiencing low mood, less enjoyment of activities, low energy, poor ability to concentrate, and insomnia (she denies thoughts of suicide or a suicide attempt). Dr. J uses the PHQ-9 to educate Cindy about the symptoms of depression. Cindy expresses relief that there is a clear way to understand and conceptualize her symptoms, and motivation to participate in treatment for her depressive symptoms. Further review of Cindy's rating scale assessment results indicates elevations in unhealthy family functioning (as measured by the McMaster Family Assessment Device). Dr. J reviews these findings with Cindy, who describes increased conflict with her 11-year-old son, Sam. Cindy explains that Sam has been angry with his parents because of their recent divorce. Dr. J seeks consultation from a behavioral health provider who suggests that Cindy would benefit from a referral to therapy.

Cindy provides permission for her therapist to review the behavioral health rating scales from Dr. J's office. Based on the elevated PHQ-9 and FAD scores, Cindy's therapist explores how her relationships might be affecting her mood and stress levels. Cindy describes Sam as "irritable" and she reports often having conflict with Sam related to traveling between two homes now that his parents are divorced. The therapist suggests Sam see his physician for an evaluation as well.

As part of MBC, Dr. K's office routinely collects parent-report rating scales on Sam's functioning from his mother, as well as Sam's self-report on several rating scales. Sam's scores revealed elevated depressive symptoms (Mood and Feeling Questionnaire score indicating mild depression), which is consistent with Cindy's mention of Sam's "irritability" during her therapy sessions. Dr. K refers Sam to a therapy group for school-aged children of divorced parents. Cindy also considers her therapist's recommendation to consider incorporating a few parent-child sessions into their work.

Six weeks later, Cindy and Sam return for follow-up appointments. In accordance with MBC practices, they complete behavioral health rating scales again. Cindy shares that she is continuing her personal therapy and her son is participating in group therapy. She believes that her relationship with her son has improved following several sessions they had together with Cindy's therapist. Cindy reports feeling more energetic and she has begun exercising again. She also shares that it has been easier to remember to take her diabetes medication since her mood has improved. Cindy's PHQ-9 follow up score is 9 and Sam's MFQ is 3 (both within normal limits). The FAD follow-up score does not reveal any disturbances in family functioning. Cindy's HbA1c decreased to 7.5%. Cindy and Sam were advised to return in 3 months for follow-up appointments.

Case Discussion

MBC practices optimize both behavioral health and medical outcomes. In the case above, scores from routinely administered rating scales provided Dr. J an opportunity to explore Cindy's mood and track progress over the course of treatment. The introduction of the PHQ-9 allowed Cindy to increase understanding of her symptoms, thus increasing engagement in treatment for both behavioral health and physical health issues. Rather than attributing Cindy's poor diabetes control to willful non-adherence, Dr. J explored behavioral health symptoms that may have prevented her from engaging in activities recommended for diabetes control. The PHQ-9 prompted and assisted Dr. J in conducting a clinical interview that resulted in a depression diagnosis. Furthermore, since the PHQ-9 measures gradations in symptom severity that are sensitive to change, Dr. J was able to see that Cindy had made clinically significant improvement at the 6-week follow-up session. Cindy's diabetes was also improved, which Cindy attributed to being better able to manage her condition due to increased energy and ability to concentrate. She had more energy to exercise and greater ability to organize and thus follow through on her diabetic diet and medications.

Since Dr. J's clinic workflow also included rating scales that measured contextual factors such as family functioning, Dr. J was able to identify that family dysfunction was likely contributing to Cindy's depressive symptoms and that a behavioral health assessment for Sam was also warranted. Behavioral health MBC also helped Dr. J identify that Sam was experiencing mild depression and difficulties related to his parents' divorce. Like his mother, Sam greatly benefited from treatment, as exhibited by his follow-up depression rating scale score, which was in the normal range.

Discussion

MBC is integral to optimizing patients' behavioral and physical health outcomes [9,14] and increasing provider accountability in patient care. MBC may be employed to augment patient evaluation, monitor behavioral health treatment progress, and increase patients' engagement in their care [48]. Patients receive higher quality care when they are educated about their symptoms, partner with their providers through a shared treatment plan, and monitor treatment outcomes along with their providers. It is advantageous for providers in primary care settings to have ready access to a broad range of psychometrically validated, patient-administered, easily scored and interpreted, freely available behavioral health rating scale measures. This paper reviewed the effectiveness evidence for behavioral health MBC, examined challenges associated with adoption, and provided recommendations for implementation into primary care practices. We also reviewed psychometrically validated and publicly available rating scales to measure a broad range of behavioral health symptoms in both adult and pediatric

populations and described some of the considerations when choosing rating scales for practice.

The benefits of MBC for patient care are clear. However, the effective integration of behavioral health rating scales into a clinical setting requires seamless integration of rating scales with clinic workflow and an office culture that supports the use of rating scales. Providers and clinic administrators offer important input on how to make the MBC process both useful and efficient. Maintaining an ongoing dialogue about the use of MBC and the impact on clinical operations may promote ongoing ideas for continual improvement. Many providers may find that the benefits of MBC to patient care outweigh initial challenges related to workflow. The growing emphasis on integrated care and awareness of the importance of a holistic view of health and wellness suggests that primary care providers must move toward including behavioral health outcomes as part of their objective measurement and documentation.

At a healthcare system level, MBC may be used to further increase quality and accountability in care by guiding professional development practices and practice-level quality improvement [4]. Examining population-level data from a primary care practice obtained through MBC provides an opportunity to learn about the prevalence and severity of behavioral health symptoms experienced by a group of individuals. This data may then be used to guide trainings and workshops for providers. Data may also be utilized to plan for new service offerings within a primary care practice. For instance, data from a primary care practice that utilizes the GAD-7 in MBC may reveal that a high percentage of patients have a possible anxiety disorder. Further analysis from the medical record may reveal that many of these patients experience chronic pain symptoms and are higher utilizers of medical services. At a systems level, this information may be used to heighten the identification of patients with comorbid anxiety and pain, and train primary care providers to provide straightforward interventions to target anxiety in the office such as mindfulness practices, progressive muscle relaxation, and visual imagery. Additionally, the health system may also decide to invest in more intensive case management services for these patients. Continuous data analysis examining behavioral health symptoms and service utilization may be conducted to examine the impact of primary care trainings and implementation of new services; this serves as the framework for continuous quality improvement in practice.

At a national level, MBC may be used to demonstrate the value of behavioral health treatment to healthcare purchasers and payers [4]. Behavioral health and medical outcomes data, along with healthcare utilization and cost data, may demonstrate the positive impact of treatment of behavioral health symptoms for primary care patients. Patient outcomes and cost-effectiveness data may be used to support reimbursement for integrated behavioral health and primary care models. MBC may be achieved within primary care settings to improve patient outcomes and promote systems-level changes that support behavioral health treatment in primary care settings.

References

1. Miller HD. From Volume To Value: Better Ways To Pay For Health Care. Health Aff (Millwood). 2009; 28: 1418-1428.
2. Lee VS, Kawamoto K, Hess R, Park C, Young J, et al. Implementation of a Value-Driven Outcomes Program to Identify High Variability in Clinical Costs and Outcomes and Association With Reduced Cost and Improved Quality. *JAMA.* 2016; 316: 1061-1072.
3. Ratzliff A, Phillips KE, Sugarman JR, Unützer J, Wagner EH. Practical Approaches for Achieving Integrated Behavioral Health Care in Primary Care Settings. *Am J Med Qual.* 2017; 32: 117-121.
4. Fortney JC, Unützer J, Wrenn G, Pyne JM, Smith GR, et al. A Tipping Point for Measurement-Based Care. *Psychiatr Serv.* 2017; 68: 179-188.
5. Barrett JE, Barrett JA, Oxman TE, Gerber PD. The prevalence of psychiatric disorders in a primary care practice. *Arch Gen Psychiatry.* 1988; 45: 1100-1106.
6. Korff MV, Shapiro S, Burke JD, Teitlebaum M, Skinner EA, et al. Anxiety and Depression in a Primary Care Clinic: Comparison of Diagnostic Interview Schedule, General Health Questionnaire, and Practitioner Assessments. *Arch Gen Psychiatry.* 1987; 44: 152-156.
7. Tyler ET, Hulkower RL, Kaminski JW. Behavioral Health Integration in Pediatric Primary Care: Considerations and Opportunities for Policymakers, Planners, and Providers [Internet]. New York, NY: Milbank Memorial Fund; 2017.
8. Collins C, Hewson DL, Munger R, Wade T. Evolving Models of Behavioral Health Integration in Primary Care [Internet]. New York, NY: Milbank Memorial Fund; 2010.
9. Unützer J, Park M. Strategies to Improve the Management of Depression in Primary Care. *Prim Care.* 2012; 39: 415-431.
10. Williams J, Klinepeter K, Palmes G, Pulley A, Foy JM. Diagnosis and Treatment of Behavioral Health Disorders in Pediatric Practice. *Pediatrics.* 2004; 114: 601-606.
11. Ani C, Bazargan M, Hindman D, Bell D, Farooq MA, et al. Depression symptomatology and diagnosis: discordance between patients and physicians in primary care settings. *BMC Fam Pract.* 2008; 9: 1.
12. Muse M, McGrath RE. Training comparison among three professions prescribing psychoactive medications: psychiatric nurse practitioners, physicians, and pharmacologically trained psychologists. *J Clin Psychol.* 2010; 66: 96-103.
13. Kroenke K, Unutzer J. Closing the False Divide: Sustainable Approaches to Integrating Mental Health Services into Primary Care. *J Gen Intern Med.* 2017; 32: 404-410.
14. Carlier IV, Meuldijk D, Van Vliet IM, Van Fenema E, Van der Wee NJ, et al. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *J Eval Clin Pract.* 2012; 18: 104-110.
15. Edbrooke-childs J, Wolpert M, Deighton J. Introduction to the Special Section on Implementing Feedback from Outcome Measures in Child and Adult Mental Health Services. *Adm Policy Ment Health Ment Health Serv Res N Y*. 2016; 43: 277-278.
16. Jensen-Doss A, Haimes EMB, Smith AM, Lyon AR, Lewis CC, et al. Monitoring Treatment Progress and Providing Feedback is Viewed Favorably but Rarely Used in Practice. *Adm Policy Ment Health Ment Health Serv Res*. 2018; 45: 48-61.
17. Moltu C, Veseth M, Stefansen J, Nøtnes JC, Skjølberg Å, et al. This is what I need a clinical feedback system to do for me: A qualitative inquiry into therapists' and patients' perspectives. *Psychother Res*. 2018; 28: 250-263.
18. Anker MG, Duncan BL, Sparks JA. Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *J Consult Clin Psychol*. 2009; 77: 693.

19. Chang TE, Jing Y, Yeung AS, Brenneman SK, Kalsekar ID, et al. Depression Monitoring and Patient Behavior in the Clinical Outcomes in MEasurement-Based Treatment (COMET) Trial. *Psychiatr Serv*. 2014; 65: 1058-1061.

20. Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. [Internet]. American Psychological Association; 2010.

21. Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. *Psychiatr Serv*. 2011; 62: 1423-1429.

22. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, et al. Measuring Quality of Life in Routine Oncology Practice Improves Communication and Patient Well-Being: A Randomized Controlled Trial. *J Clin Oncol.* 2004; 22: 714-724.

23. Clarke RMA, Jeffrey J, Grossman M, Strouse T, Gitlin M, et al. Delivering On Accountable Care: Lessons From A Behavioral Health Program To Improve Access And Outcomes. *Health Aff (Millwood)*. 2016; 35: 1487-1493.

24. Sattler JM. Assessment of Children: Cognitive Foundations, 5th Edition. 5th edition. San Diego: Jerome M. 2008. 796.

25. Henke RM, Zaslavsky AM, McGuire TG, Ayanian JZ, Rubenstein LV. Clinical Inertia in Depression Treatment. *Med Care.* 2009; 47: 959-967.

26. Youn SJ, Kraus DR, Castonguay LG. The treatment outcome package: Facilitating practice and clinically relevant research. *Psychotherapy.* 2012; 49: 115-122.

27. Priebe S, McC R, Bullenkamp J, Hansson L, Lauber C, et al. Structured patient–clinician communication and 1-year outcome in community mental healthcare. *Br J Psychiatry*. 2007; 191: 420-426.

28. Groves RM. Probing the Causes of Nonresponse and Efforts to Reduce Nonresponse. In: Survey Errors and Survey Costs [Internet]. 1989; 185-238.

29. Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, et al. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ.* 2009; 338: b663.

30. Keshen A, Helson T, Town J, Warren K. Self-efficacy as a predictor of treatment outcome in an outpatient eating disorder program. *Eat Disord.* 2017; 25: 406-419.

31. Kuo C-C, Lin C-C, Tsai F-M. Effectiveness of Empowerment-Based Self-Management Interventions on Patients with Chronic Metabolic Diseases: A Systematic Review and Meta-Analysis. *Worldviews Evid Based Nurs.* 2014; 11: 301-315.

32. Strecher VJ, McEvoy DeVellis B, Becker MH, Rosenstock IM. The Role of Self-Efficacy in Achieving Health Behavior Change. *Health Educ Q.* 1986; 13: 73-92.

33. Zullig LL, Peterson ED, Bosworth HB. Ingredients of Successful Interventions to Improve Medication Adherence. *JAMA.* 2013; 310: 2611-2612.

34. Germann JN, Kirschenbaum DS, Rich BH. Child and Parental Self-Monitoring as Determinants of Success in the Treatment of Morbid Obesity in Low-Income Minority Children. *J Pediatr Psychol*. 2007; 32: 111-121.

35. Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychother Res.* 2015; 25: 6-19.

36. Jeffrey J, Sinclair M, Linonis R, Semaan A, Hsiao T, et al. Integration of a Web-based Behavioral Health Assessment within a Collaborative Care Setting. *J Mob Technol Med.* 2018; 7: 9-15.

37. Goldstein LA, Connolly Gibbons MB, Thompson SM, Scott K, Heintz L, et al. Outcome Assessment via Handheld Computer in Community Mental Health: Consumer Satisfaction and Reliability. *J Behav Health Serv Res.* 2011; 38: 414-423.

38. Ogles BM, France CR, Lunnen KM, Bell T, Goldfarb M. Computerized Depression Screening and Awareness. *Community Ment Health J.* 1998; 34: 27-38.

39. Zimmerman M, Martinez JH. Web-based assessment of depression in patients treated in clinical practice: reliability, validity, and patient acceptance. *J Clin Psychiatry.* 2012; 73: 333-338.

40. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): Conceptualizing and Measuring Activation in Patients and Consumers. *Health Serv Res.* 2004; 39: 1005-1026.

41. Ajami S, Bagheri-Tadi T. Barriers for Adopting Electronic Health Records (EHRs) by Physicians. *Acta Inform Medica.* 2013; 21: 129-134.

42. Youth in Mind. SDQ Scoring.

43. Meyers K, Winters NC. Ten-Year Review of Rating Scales. I: Overview of Scale Functioning, Psychometric Properties, and Selection. *J Am Acad Child Adolesc Psychiatry.* 2002; 41: 114-122.

44. Priebe S, McCabe R, Bullenkamp J, Hansson L, Rössler W, et al. The impact of routine outcome measurement on treatment processes in community mental health care: approach and methods of the MECCA study. *Epidemiol Psychiatr Sci.* 2002; 11: 198-205.

45. Connell J, Carlton J, Grundy A, Taylor Buck E, Keetharuth AD, et al. The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). Qual Life Res. 2018; 27: 1893-1902.

46. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *J Affect Disord.* 2004; 81: 61-66.

47. Kroenke K, Spitzer RL. The PHQ-9: A New Depression Diagnostic and Severity Measure. *Psychiatr Ann.* 2002; 32: 509-515.

48. Lewis CC, Boyd M, Puspitasari A, Navarro E, Howard J, et al. Implementing Measurement-Based Care in Behavioral Health: A Review. *JAMA Psychiatry.* 2018.

49. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. *J Gen Intern Med.* 2001; 16: 606-613.

50. Wood A, Kroll L, Moore A, Harrington R. Properties of the Mood and Feelings Questionnaire in Adolescent Psychiatric Outpatients: A Research Note. *J Child Psychol Psychiatry.* 1995; 36: 327-334.

51. Rhew IC, Simpson K, Tracy M, Lymp J, McCauley E, et al. Criterion validity of the Short Mood and Feelings Questionnaire and one- and two-item depression screens in young adolescents. *Child Adolesc Psychiatry Ment Health*. 2010; 4: 8.

52. Thabrew H, Stasiak K, Bavin L-M, Frampton C, Merry S. Validation of the Mood and Feelings Questionnaire (MFQ) and Short Mood and Feelings Questionnaire (SMFQ) in New Zealand help-seeking adolescents. *Int J Methods Psychiatr Res*. 2018; 27: e1610.

53. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: The GAD-7. *Arch Intern Med.* 2006; 166: 1092-1097.

54. Birmaher B, Brent DA, Chiappetta L, Bridge J, Monga S, et al. Psychometric Properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): A Replication Study. *J Am Acad Child Adolesc Psychiatry.* 1999; 38: 1230-1236.

55. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, et al. The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale Construction and Psychometric Characteristics. *J Am Acad Child Adolesc Psychiatry.* 1997; 36: 545-553.

56. Caporino NE, Sakolsky D, Brodman DM, McGuire JF, Piacentini J, et al. Establishing Clinical Cutoffs for Response and Remission on the Screen for Child Anxiety Related Emotional Disorders (SCARED). *J Am Acad Child Adolesc Psychiatry*. 2017; 56: 696-702.

57. Prins A, Ouimette P, Kimerling R, Camerond RP, Hugelshofer DS, et al. The primary care PTSD screen (PC–PTSD): development and operating characteristics. *Prim Care Psychiatry.* 2004; 9: 9-14.

58. Freedy JR, Steenkamp MM, Magruder KM, Yeager DE, Zoller JS, et al. Post-traumatic stress disorder screening test performance in civilian primary care. Fam Pract. 2010; 27: 615-624.

59. Prins A, Bovin MJ, Smolenski DJ, Marx BP, Kimerling R, et al. The Primary Care PTSD Screen for DSM-5 (PC-PTSD-5): Development and Evaluation Within a Veteran Primary Care Sample. *J Gen Intern Med.* 2016; 31: 1206-1211.

60. Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and Initial Psychometric Evaluation. *J Trauma Stress.* 2015; 28: 489-498.

61. Bovin MJ, Marx BP, Weathers FW, Gallagher MW, Rodriguez P, et al. Psychometric properties of the PTSD checklist for diagnostic and statistical manual of mental disorders–fifth edition (PCL-5) in veterans. *Psychol Assess.* 2016; 28: 1379.

62. Foa EB, Asnaani A, Zang Y, Capaldi S, Yeh R. Psychometrics of the Child PTSD Symptom Scale for DSM-5 for Trauma-Exposed Children and Adolescents. *J Clin Child Adolesc Psychol.* 2018; 47: 38-46.

63. Saxe G, Chawla N, Stoddard F, Kassam-adams N, Courtney D, et al. Child Stress Disorders Checklist: A Measure of ASD and PTSD in Children. *J Am Acad Child Adolesc Psychiatry.* 2003; 42: 972-978.

64. Kessler RC, Adler L, Ames M, Demler O, Faraone S, et al. The World Health Organization adult ADHD self-report scale (ASRS): a short screening scale for use in the general population. *Psychol Med.* 2005; 35: 245-256.

65. Kessler RC, Adler LA, Gruber MJ, Sarawate CA, Spencer T, et al. Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) Screener in a representative sample of health plan members. *Int J Methods Psychiatr Res.* 2007; 16: 52-65.

66. Adler LA, Spencer T, Faraone SV, Kessler RC, Howes MJ, et al. Validity of Pilot Adult ADHD Self- Report Scale (ASRS) to Rate Adult ADHD Symptoms. *Ann Clin Psychiatry.* 2006; 18: 145-148.

67. Bussing R, Fernandez M, Harwood M, Wei Hou null, Garvan CW, et al. Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: psychometric properties and normative ratings from a school district sample. *Assessment.* 2008; 15: 317-328.

68. Alda JA, Serrano-Troncoso E. Attention-deficit hyperactivity disorder: agreement between clinical impression and the SNAP-IV screening tool. *Actas Esp Psiquiatr.* 2013; 41: 76-83.

69. Knight JR, Sherritt L, Shrier LA, Harris SK, Chang G. Validity of the CRAFFT Substance Abuse Screening Test Among Adolescent Clinic Patients. *Arch Pediatr Adolesc Med.* 2002; 156: 607-614.

70. Mitchell SG, Kelly SM, Gryczynski J, Myers CP, O'Grady KE, et al. The CRAFFT Cut-Points and DSM-5 Criteria for Alcohol and Other Drugs: A Reevaluation and Reexamination. *Subst Abuse.* 2014; 35: 376-380.

71. Cocco KM, Carey KB. Psychometric properties of the Drug Abuse Screening Test in psychiatric outpatients. *Psychol Assess.* 1998; 10: 408-414.

72. Saitz R, Cheng DM, Allensworth-Davies D, Winter MR, Smith PC. The Ability of Single Screening Questions for Unhealthy Alcohol and Other Drug Use to Identify Substance Dependence in Primary Care. *J Stud Alcohol Drugs.* 2014; 75: 153-157.

73. Reinert DF, Allen JP. The Alcohol Use Disorders Identification Test (AUDIT): A Review of Recent Research. *Alcohol Clin Exp Res.* 2002; 26: 272-279.

74. Saunders JB, Aasland OG, Babor TF, Fuente JRDL, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction.* 1993; 88: 791-804.

75. Goodman R. Psychometric Properties of the Strengths and Difficulties Questionnaire. *J Am Acad Child Adolesc Psychiatry.* 2001; 40: 1337-1345.

76. Miller IW, Epstein NB, Bishop DS, Keitner GI. The McMaster Family Assessment Device: Reliability and Validity*. *J Marital Fam Ther.* 1985; 11: 345-356.

77. Miller IW, Ryan CE, Keitner GI, Bishop DS, Epstein NB. The McMaster Approach to Families: theory, assessment, treatment and research. *J Fam Ther.* 2000; 22: 168-189.

78. Mundt JC, Marks IM, Shear MK, Greist JH. The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *Br J Psychiatry J Ment Sci.* 2002; 180: 461-464.

79. Jansson-Fröjmark M. The Work and Social Adjustment Scale as a Measure of Dysfunction in Chronic Insomnia: Reliability and Validity. *Behav Cogn Psychother.* 2014; 42: 186-198.

ADDRESS FOR CORRESPONDENCE:

Jessica Jeffrey, University of California, Los Angeles, USA, E-mail: JJeffrey@mednet.ucla.edu

Submission: 16 September 2019

Accepted: 17 October 2019