ADVANCING MEASUREMENT-INFORMED CARE IN

Community Behavioral Health

NATIONAL COUNCIL *for* Mental Wellbeing

CENTER OF EXCELLENCE
for Integrated Health Solutions

Funded by Substance Abuse and Mental Health Services Administration and operated by the National Council for Mental Wellbeing

AUTHORS:

Henry Chung, MD
Professor of Psychiatry,
Albert Einstein College of Medicine

Deborah Scharf, PhD, CPsych
Associate Professor,
Department of Psychology, Lakehead University

Joe Parks, MD
Medical Director,
National Council for Mental Wellbeing

Jeff Capobianco, PhD
Senior Consultant,
National Council for Mental Wellbeing

Vamika Mann, MA
Doctoral Candidate,
Department of Psychology, Lakehead University

Alexandra Plante, MA
Senior Advisor, Substance Use Disorder,
National Council for Mental Wellbeing

MEASUREMENT INFORMED CARE ADVISORY GROUP

Joe Parks, MD
Medical Director,
National Council for Mental Wellbeing

Deborah Scharf, PhD, CPsych
Associate Professor,
Department of Psychology, Lakehead University

Henry Chung, MD
Professor of Psychiatry,
Albert Einstein College of Medicine

Jack Todd Wahrenberger MD MPH
Chief Medical Officer,
Pittsburgh Mercy Health System

Jeff Capobianco, PhD
Senior Consultant,
National Council for Mental Wellbeing

John Bischof, MD
Medical Director Behavioral Health,
CareOregon

Jorge R. Petit, MD
Founder/CEO,
Quality Healthcare Solutions

Karen L. Fortuna, PhD, LICSW
Assistant Professor of Psychiatry,
Geisel School of Medicine at Dartmouth

Lori Raney, MD
Owner, Collaborative Care Consulting

Luming Li, MD, MHS
Chief Medical Officer,
The Harris Center for Mental Health & IDD

Martin Rosenzweig, MD
Chief Medical Advisor, GreyMatter

Patrick Runnels, MD, MBA
Chief Medical Officer, Population Health University Hospitals
Chief Medical Officer, Veale Initiative for Health Care Innovation
Professor, Department of Psychiatry
Case Western Reserve School of Medicine

Rochelle Head-Dunham, MD, DFAPA, FASAM
Executive Director and Medical Director
Metropolitan Human Services District;
LSU and Tulane University Schools of Medicine

Sosunmolu Shoyinka, M.D, MBA
Founder, Centia Health LLC,
Clinical Associate Professor of Psychiatry,
University of Pennsylvania Perelman School of Medicine

Virna Little, Psy. D, LCSW-r, MBA, CCM, SAP
Co-founder & COO, Concert Health

H. Westley Clark, M.D., J.D., M.P.H.
Dean's Executive Professor of Public Health,
Santa Clara University

CORRESPONDING AUTHOR:

Henry Chung, HChung@Montefiore.org

FINANCIAL SUPPORT:

This report is supported by the National Council for Mental Wellbeing's Center of Excellence for Integrated Health Solutions, funded by a grant award from the Substance Abuse and Mental Health Services Administration (SAMHSA) and managed by the National Council for Mental Wellbeing. The views, opinions and content expressed in this presentation do not necessarily reflect the views, opinions or policies of the Center for Mental Health Services, SAMHSA or the U.S. Department of Health and Human Services.

ACKNOWLEDGMENTS:

The team would like to extend our sincerest gratitude to Sr. Project Manager Tiffany Francis and other National Council collaborators, including members of our expert panel, and session participants at the NatCon23 meeting in Los Angeles on May 3 and the SAMHSA lunch-and-learn on May 30, 2023, who provided input on the project process and results to produce this position paper.

Overview

Measurement-based care (MBC) is the practice of systematically repeating clinical assessments of symptoms and functional outcomes for the purpose of directing clinical care, including changes in the type or intensity of treatment (Alter et al., 2021). MBC must be adapted when applied to behavioral health care, because behavioral health clinicians disproportionately consider difficult-to-quantify psychosocial factors when determining treatment plans and otherwise directing care. Yet research shows that behavioral health care quality improves when clinicians repeatedly assess and provide feedback on client progress toward quantifiable outcomes, so long as their measures are meaningful to both clinicians and service users.

Measurement-*informed* care (MIC), in which measurement is one of many factors that informs (rather than determines) behavioral health treatment, has great potential to advance the quality and impact of the nation's behavioral health care. This includes improving organizational quality monitoring, accountability and progress toward financial sustainability goals. MIC uptake and sustainability depend on the broad availability of, and agreement upon, a set of reliable and valid quality measures that behavioral health teams can reasonably use to inform care. At the time of this report, national quality programs have adopted few behavioral health measures aligned with MIC. Among those few that are aligned, such as measures for depression response and remission, there are significant implementation issues, including low adoption by clinicians and payers, high clinician or service-user burden, and/or complex data specifications that limit clinicians' use of the measures as intended. Measurement complexity also contributes to high costs.

This report describes our review of existing behavioral health measures against an expert panel-approved set of criteria to identify measures potentially available to support MIC. We then put forward a set of concepts and candidate measures to advance MIC through transdiagnostic measurement. Transdiagnostic MIC has the potential to reduce organizations' reporting burden, increase clinician and service-user uptake in treatment decision-making, and shift attention to the needs of the whole person to better align with the ideology of the field.

State of The Field

Repeated outcome measurement is a well-studied practice in behavioral health (Lambert et al., 2018; Tartakovsky, 2016). Research shows that routine clinical measurement of behavioral health symptoms and other indicators (e.g., therapeutic alliance), when fed back to clients, yields a range of benefits, such as improved client retention, reduced no-show rates and improved clinical outcomes (Lambert et al., 2018; Tartakovsky, 2016).

Within the past decade in the behavioral health field, repeated clinical measurement has been promoted significantly, not only to improve outcomes but also for systemic purposes such as justifying, apportioning and financing care. This promotion aligns with care practices for some chronic physical health conditions that include repeated measurements of single biometric indicators (e.g., HbA1C for diabetes or blood pressure for hypertension) (American Diabetes Association, 2019; Carey et al., 2021). Research on the Collaborative Care Model (CoCM) supports this approach in behavioral health. CoCM studies show that a single measure, the PHQ-9 (Kroenke et al., 2001), can inform primary care team actions that enhance effectiveness, efficiency and accountability in depression care (Fortney et al., 2017). Given the demonstrated and anticipated benefits of a single indicator for some conditions and populations (e.g., MBC; **see Box 1**), payers and regulators are increasingly encouraging data collection for similar behavioral health outcomes reporting (e.g., PHQ-9 for depression care). Such rapid shifts in the role and importance of measurement in behavioral health challenge the field, leadership and policy organizations to select and advocate for measures that reflect behavioral health science, ideology and customs (**see Box 2**).

The National Council supports advances in behavioral health measurement. It also recognize that measures and practices must reflect whole-person (integrated), contextual and biopsychosocial approaches.

BOX 1. Anticipated Benefits of Enhanced Use of Measurement in Behavioral Health

Improved care quality

- Increased service-user education, health literacy, treatment engagement and adherence through shared decision-making.
- Increased clinician attention and responsivity to symptom changes (e.g., treatment to target).
- Increased aggregate data availability to support population-based care.

Increased care value

- Improved efficiency through routinized data collection and review.
- Improved integration with general medicine providers.

Improved clinical outcomes

- Improved response and remission rates.

Improved sustainability of care

- Increased reimbursement opportunities for direct billing and value-based programs.
- Improved behavioral health value proposition.
- Richer data collection and outcomes reporting to support accreditation requirements.

Illustrated impact of care

- Community behavioral health service impacts showcased to shift the policy and payer mandate toward indicators that are most meaningful to service users, families, communities and providers.

(Alter et al., 2021; Finn & Tonsager, 1997; Fortney et al., 2017; Hallgren et al., 2022; Kendrick et al., 2023; Lambert et al., 2018; Ridout et al., 2023; Van Tiem et al., 2022)

Throughout this paper, we have aimed to articulate, for the first time, general principles for using MIC in behavioral health. This includes specifying the meaning, use and value of measurement in behavioral health care, while taking into account the customs and beliefs of service users, clinicians, payers and regulators. We propose the following general principles for MIC:

1. Measurement is essential to treatment decision-making.
2. Measurement alone is not sufficient as a sole process for making treatment decisions.
3. Measurement errors can occur due to the subjective nature of measurement tools.
4. Measurement must be done in the context of health literacy (Shahid et al., 2022) and fairness considerations to ensure service users understand the meaning, purpose and utility of treatments.

WHICH MEASURES?

Rapid shifts in the use and importance of measurement in behavioral health have led to a proliferation of reporting requirements and measures, causing confusion, misalignment of approaches and reporting burden (Jacobs et al., 2023). In response, behavioral health policy and leadership organizations have produced reports to direct clinicians and organizations toward a curated set of measures that can help them meet their emerging measurement needs, including improving care, integrating with general medical systems, demonstrating quality of care and value propositions, and participating in required accountability, accreditation and payment processes.

For example, the Meadows Mental Health Policy Institute (Alter et al., 2021) produced a follow-up to the Kennedy Forum's report (Wrenn & Fortney, 2015), in which they used literature review and surveys of community interested parties to identify 36 service-user self-report scales — 22 of which were for adult populations — that assessed specific behavioral health conditions or overall functioning, could screen for and facilitate diagnosis, and were sensitive to clinical change. The set included multiple scales for some conditions (e.g., two scales for depression: PHQ-9 [Kroenke et al., 2001] and PROMIS depression [Schalet et al., 2016]) and singular scales for others. The authors concluded that some scales can be used by community behavioral health clinicians and organizations (proprietary and nonproprietary), and identified a need for additional quality measures focused on clinical outcomes for some diagnoses and populations. They also identified a range of implementation issues related to reimbursement mechanisms, mechanisms for electronic service-user report, electronic health record (EHR) functionality for scoring and results sharing, and more (Alter et al., 2021).

The American Psychiatric Association (APA) further described how solo and small behavioral health practices can align clinical measurement with other accountability and finance requirements (Ridout et al., 2023). Its report includes a framework to aid in the selection of candidate scales for specific behavioral health diagnoses, with considerations related to cost, ease of use and administration, and the reliability and validity of scales across clinical settings.

Many scales identified by the APA, Kennedy Forum and Meadows groups overlap. Although this could be due to overlap in contributors, it may also suggest some consensus in the field about the best available scales for specific diagnoses and where measurement opportunities exist.

BOX 2. Environmental and Ideological Challenges to Measurement in Behavioral Health

Feasibility

- Clinician time for administration, scoring, interpretation, feedback and provider data exchange.
- Electronic health record (EHR) or other automated infrastructure for data capture, synthesis, presentation and interpretation.
- Expense associated with measures themselves and/or changes to EHR or practice to satisfy changing reporting requirements.
- Proliferation of measures for multimorbid conditions creates uncertainty about what to measure and how to measure them.
- Variability in reporting structures and reporting requirements between states.

Clinician issues

- Insufficient clinician knowledge/training to effectively collect, score, interpret and share repeated measures data.
- Devaluation of any directives ("change fatigue") that providers perceive as detracting from their ability to address service users' immediate needs.

Service-user issues

- Perceived service-user resistance to completing questionnaires due to burden, perceived low importance, or non-alignment with quality of life and functional goals.
- Validity issues with self-report data due to literacy challenges.

Context and beliefs

- Oversimplification of behavioral health practice (e.g., reductive approach to service-user engagement, goals).
- Disconnection from the human experience.
- Perceived failure to capture important yet intangible aspects of the work.
- Perception that adequate assessment of behavioral health practice is necessarily burdensome because it is multifaceted.
- Overextension of privacy practices creates resistance to sharing measured outcomes transparently with all interested parties.

(Ko et al., 2023; Lewis et al., 2019; Ridout et al., 2023).

Determining how to prioritize the development of potential measures aligned with the respective diagnostic symptom scales identified in the APA and Meadows reports is a challenge that needs to be addressed. Policymakers and regulators must work with quality measurement experts to re-consider scales' criteria for what to include/exclude, population numerators and denominators, and assessment frequency or other specifications. The extent to which scales improve alignment between clinical care and reporting requirements, reduce service-user and clinician burden, and reflect the quality, breadth and value of behavioral health work must also be considered.

Defining Measurement-informed Care: A Balanced, Sustainable Approach

We propose that the behavioral health field must reconcile the wide gap between how it typically uses available scales for treatment decision-making — such as those that screen for and/or help diagnose conditions such as depression, anxiety and post-traumatic stress disorder (PTSD) — with measurement for organizational and systemic needs. We call this a measurement-*informed* care (MIC) approach. We propose that MIC consists of 1) using reliable and valid tools, 2) for repeated measurements at clinically meaningful, regular intervals 3) of service-user-reported outcomes and/or biometric indicators, 4) as one among many sources of information (service-user preferences, social and environmental factors, customs and beliefs, quality of life and functional needs/goals, family support, health literacy, etc.), 5) to make treatment decisions to support clinical progress, such as the need for changes to the treatment plan, and 6) as indications of accountability and efficiency of care.

To be sustainable, MIC requires strong consensus among all interested parties about the appropriate scales and their specified use across populations and settings. Payers, grantors, administrators, clinicians, service users, families, caregivers and other parties must also agree that selected measures minimize burden, can be used to establish quality benchmarks, demonstrate service-user recovery and can be used to promote quality behavioral health care overall. This includes maintaining consistency with population health priorities and program/organizational accountability requirements for behavioral health payment models (such as relevant value-based payment [VBP] structures), integration and parity.

To truly be aligned with whole-person care, measures must also be implemented as information to be considered alongside other sources, such as service-user and family reports, clinical observation and history, social and environmental factors that affect health, level of health literacy, disability status and community context. Since many of these critical sources of information are qualitative by design, teams implementing MIC in behavioral health care must have the latitude to integrate multiple sources and types of information into clinical decision-making and assessments of care quality.

Purpose

The purpose of this report is to foster and promote an effective and sustainable approach to behavioral health MIC. To do this, we:

1. Articulate criteria that can be applied to existing (or forthcoming) behavioral health measures to determine their suitability for MIC.
2. Apply these criteria to existing health care quality organization-endorsed (e.g., NCQA) behavioral health quality measures, to identify the best available MIC-supporting measures.
3. Propose next steps toward achieving sustainable MIC in behavioral health, including shifting from diagnosis-specific to transdiagnostic, functional outcomes suitable for varied populations and settings, and putting forth candidate measures for further consideration by the field.

In the next section, we describe a research project articulating a vision and process for organizations to achieve sustainable MIC. Our findings include guidance on immediate actions and next steps for implementation, advocacy and research of MIC within the field of behavioral health.

Research Project

Using the available quality measures endorsed by national health care quality organizations, this three-phase project was designed to summarize them by content area (e.g., diagnosis) and function (e.g., process versus outcome), articulate criteria to identify existing MIC-compatible measures, and prioritize the measures in terms of their suitability for MIC. This process included:

- Highlighting existing measures that fit our priority criteria.
- Identifying measures that could fit the criteria with some modification.
- Identifying weaknesses in available measures, in terms of content and function.
- Recommending candidate measures, where available, for further research and development.

We have limited our discussion to MIC for adult (aged 18-65 years) outpatient populations served in community behavioral health settings. While we expect that this methodological and conceptual work could generalize to MIC in other behavioral health care settings (e.g., inpatient care) and populations (e.g., children, adolescents, older adults), we recognize that researchers will need to ensure that essential differences in other areas are not missed. Ideally, our method provides a framework that others with complementary expertise (e.g., children, adolescents, older adults) can use to advance MIC in the behavioral health systems and populations they serve.

Phase I: Literature and Measures Review

The first phase of the research project was to catalog behavioral health measures endorsed by national quality organizations (**see Box 3**) and categorize them according to a range of features useful for determining their suitability for MIC.

METHODS

Two team members (VM, DS) reviewed and extracted all adult, outpatient behavioral health measures currently (fall 2022 and winter 2023) endorsed by national quality organizations active within the United States.[1] The list of quality organizations contributing measures is in **Box 3**. We then extracted these measures into a central repository and coded them along dimensions that we anticipated might be relevant to MIC. These dimensions included measure name, source, type, text summary, disorder, disorder type (e.g., serious mental illness, substance use disorder), age (e.g., adolescent, older adult), population, additional criteria for what to include/exclude specified in the measure, numerator/denominator information, reliability and validity information, endorsement status and endorsing organization(s), whether the measure was process versus outcome oriented, if it specified a standardized assessment tool and which one(s) were included (e.g., PHQ-9), and any references to peer-reviewed, academic publications. We used descriptive statistics to summarize the contents of the repository.

BOX 3. Measures Review Sources (December 2022 to January 2023)

1. American Psychiatric Association – Mental and Behavioral Health Registry (APA-MBHR)
2. Certified Community Behavioral Health Clinics (CCBHCs)
3. Consumer Assessment of Healthcare Providers and Systems (CAHPS)
4. Medicaid Innovation Accelerator Program (IAP)
5. National Committee for Quality Assurance (NCQA)
6. National Outcome Measures (NOMs)
7. National Quality Form (NQF)
8. Patient-reported Outcomes Measurement Information System (PROMIS)
9. Personal Outcome Measures (POMs)
10. Quality Payment Program (QPP)
11. Substance Abuse and Mental Health Services Administration (SAMHSA)
12. World Health Organization (WHO)

RESULTS

Our review of endorsed measure sets resulted in a total of 215 unique measures. The measures were nearly evenly divided between outcome (n=105; 49%) and process (n=110; 51%). Among the outcome measures, approximately one-third (n=52; 31%) were patient-reported outcome measures (PROMs). Note that this deviates from the broader set of general medical measures in the Meadows report, which showed that, of the 1,000 different quality measures utilized across all CMS programs, 95% were process focused and not outcome focused (Alter et al., 2021).

We coded and categorized the content focus of each measure. This was challenging due to variability in how core concepts were specified. For example, simply identifying "adult" measures was problematic, as "adult" was defined in ways that sometimes included or excluded younger and/or older adult groups (e.g., 18-64, 18-75, 18-85, 19+, 25-64). We retained measures that included most of our target age range of 18-65 years.

[1] *For comparison, 16 measures from an international organization, the Canadian Institute for Health Information (CIHI), were included in the original data set but have been removed from the data presented in this report for simplicity and applicability to National Council membership.*

Clinical terms were also defined differently, which complicated our attempts to categorize measures by diagnostic group. For instance, while many measures addressed symptoms within a single diagnostic category (e.g., GAD-7 for generalized anxiety disorder; Spitzer et al., 2006), others were designed for broader clinical groups (e.g., "serious mental illness") that were either not defined or defined differently across measures (e.g., psychotic disorders versus serious mental illness versus schizophrenia including schizoaffective disorder). Substance use was defined with similar variability (e.g., alcohol, alcohol and drugs, opioid use disorder, tobacco), as were classifications of behaviors (e.g., self-harm) relevant to multiple diagnoses.

Only a small subset of identified measures (n=32; 15%) was endorsed by Medicaid (including section 2703 Health Homes for Chronic Conditions; Medicaid, 2023) and/or was required reporting for Certified Community Behavioral Health Clinics (CCBHCs).

Phase II: Expert Panel Initial Review

A data summary and a copy of the endorsed, coded behavioral health data set was provided to the project's expert panel. The expert panel reviewed the database of endorsed behavioral health measures, advised on strategies for organizing and evaluating the available measures for MIC, and then provided feedback on the suitability of those measures. They also advised on next steps to advance the field.

METHODS

Our expert panel was composed of 12 people with behavioral health expertise in addiction care, administration, clinical care, integrated care, primary care, psychology, psychiatry, quality measurement, research and social work. Experts were identified through the National Council's network and through project consultants' social networks. The panel met in January 2023 to review the data summary and database of measures. At the first meeting, experts were asked:

- What measures do you use in your program/organization? How did you choose them?
- Where do you see the potential impacts of these measures in your work and your field?
- What are strengths, challenges and opportunities for improvement?
- How can a white paper on recommended MIC measures for behavioral health be most impactful to clinicians and interested parties? What is on your wish list for future measures?
- What specific concerns or considerations should we account for while undertaking this research?

After reviewing the measure set, they were asked:

- What additional sets of endorsed measures might we have missed in our review?
- How should we prioritize measures for MIC?

- If measures are not currently endorsed, should we propose a limited set of new measures?
- If measures are endorsed, should we propose modifications that would improve their feasibility and relevance? If so, what strategies should we use to do this?

RESULTS

The primary theme expert panel members raised was the disconnect between regulators', payers' and general medical partners' measurement expectations and the availability of suitable measures for behavioral health. Specifically, **while regulators and payers were expecting behavioral health systems and clinicians to engage in MIC, after their review experts asserted that there were not widely accepted and agreed-upon measures to use.**

Experts further discussed issues within the following three themes:

- Robust MIC remains challenging because the quality organization–endorsed measures that payers and regulators mandate or incentivize (e.g., Healthcare Effectiveness Data and Information Set [HEDIS], NCQA) differ from those that indicate clinical change or are important to service users, families and caregivers.
  - Interested parties may have different perceptions of measures' utility due to content (e.g., follow-up after hospitalization within 30 days is a useful systemic measure of care quality but says little about how clinicians should direct an individuals' care) and timing (e.g., PHQ-9 is a useful indicator of clinical progress, but annual PHQ-9 reassessments, as required for CCBHC, are too infrequent to direct care).
- Behavioral health leadership and policy organizations can and should help direct the concepts and content that will improve measures for behavioral health MIC and align them with those put forth by national quality organizations, thoughtfully and in an organized fashion. A future standardized suite of diagnostic, symptom and functional-impairment severity scales (used as numeric indicators of clinical change in people receiving care) and measures (suitable scales that specify adequate frequency of measurement, numerators/denominators, criteria for what to include/exclude) should include:
  - Outcome measures to address issues/conditions considered significant enough to warrant regular screening, such as those identified by the United States Preventive Services Task Force (USPSTF) or other national advisory/scientific body.
  - Transdiagnostic and functional measures.
  - Measures of quality that include clinical, functional and utilization outcomes that can be used post hoc to calculate value of care.

Following this meeting, and using the feedback we received, we returned to the expert panel four months later to solicit input on a framework we developed for the National Council to organize behavioral health measures for MIC. The framework consists of two groups, or tiers, of measures.

TIER 1 MEASURES

Conceptually, Tier 1 measures are those that represent the state of the field in behavioral health MIC. This includes measures that

are well validated and accepted, such as at the level of being named in clinical consensus statements, widely used and valued by clinicians and service users, families and caregivers, and those that can inform organizational needs (e.g., accountability, billing). Tier 1 measures are for conditions that impact the general population to the extent that the USPSTF has identified them for routine screening/follow-up and/or treatment if the disorder is present. Tier 1 measures also are aligned with reporting requirements for Medicaid (e.g., through programs such as Health Homes [Medicaid, 2023]) or a reporting requirement for CCBHCs, used to indicate quality and/or for certification/accreditation and sustainability purposes. Tier 1 measures have quantitative cut points to inform aggregate population health analysis and are one among many pieces of information that clinicians can use to indicate clinical progress or a need for changes in the treatment plan within individuals' context of care.

Upon review of these criteria and measures, the expert panel advised that we add additional criteria for Tier 1 measures:

- They should include limited biometric control measures that are important for the care of individuals receiving antipsychotic medications and for advancing integrated care (e.g., HgBA1c, blood pressure).
- There should be the potential for possible expansion to the entire CCBHC population (whole-person care in adult community behavioral health).
- They can be improved in terms of timing and frequency of measurement as needed to incentivize better data collection and reimbursement.
- They are likely to be supported by payers.

The expert panel also noted unresolved issues at this phase of the process:

- The need to modify and improve specifications of some Tier 1 measures to meaningfully inform clinical care as well as improve organizational monitoring efforts for these measures (e.g., the use of episode-based time periods for depression response and remission). This could also include specifications for how data are shared with people receiving care (Shahid et al., 2022).
- The inconsistency of using percent severity reductions to define response whereas categorical cut points are used for remission. This increases the organizational burden of calculating and monitoring individuals' progress on these measures during the measurement year.

Complete criteria for Tier 1 measures to include/exclude are in **Box 4**. After solidifying Tier 1 criteria, we applied them to the measures in our database and identified a Tier 1 measure set displayed in Table 1.

BOX 4. Tier 1 Measures Criteria for What to Include/ Exclude

Include

- **Required reporting for Medicaid or CCBHC (or future major federal behavioral health programs)**
- **Outcome focused**
- **Service-user self-report scales or biometric indicator**
- **Low burden (≤15 items)**
- **Sensitive to clinical change**
- **Psychometrically sound (reliable, valid)**
- **Scales with established norms and clinical severity thresholds**
- **Adult**
- **Outpatient**
- **Suitable for community behavioral health settings and populations**
- **Free and in the public domain**
- **Eligible for reimbursement**

Exclude

- **Process focused**
- **Epidemiological (counts only)**

Table 1. Tier 1 measures

Name	Metric Specifications	Source	Medicaid/ CCBHC	Self-Report	# Items	Implementation and Proposed Modification
Depression Response/ Remission at Six Months	The percentage of adolescent patients (12–17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reached response (PHQ-9 50% reduction) or remission (PHQ-9<5) in six months (+/- 60 days) after an index event date.	APA-MBHR, NCQA	Yes	Yes	10	• Recommended frequency of assessment is monthly. • Consider creating a categorical cut point for response instead of percent reduction. • Episode-based time interval should be revised to last score in calendar year.
Anxiety Response at Six Months	The percentage of adult patients (18 years of age or older) with an anxiety disorder (e.g., generalized anxiety disorder, social anxiety disorder or panic disorder) who demonstrated a response to treatment (GAD-7 score at least 25% less than score at index event) at six months (+/- 60 days) after an index visit.	APA-MBHR	Yes	Yes	8	• Recommended frequency of assessment is monthly. • Consider creating a categorical cut point for response instead of percent reduction. • Episode-based time interval should be revised to last score in calendar year.
Alcohol Use Disorder Outcome Response	The percentage of adult patients (18 years of age or older) who report challenges with drinking alcohol (e.g., can be noted through a screening measure such as the AUDIT-C as described in MIPS Clinical Quality Measure Quality ID #431, aka NQF 2152, or other drug/ alcohol screeners such as the DAST and TAPS) *and* demonstrated a response to treatment at three months (+/- 60 days) after the index visit.	APA-MBHR	Yes	Yes	3	• Create a categorical cut point to indicate alcohol treatment response indicating drinking within NIAAA (or other) safe limits.

Comprehensive Diabetes Care for People With Serious Mental Illness: Hemoglobin A1c (HbA1c) Poor Control (>9.0%)	Members 18-75 years of age with at least one acute inpatient visit or two outpatient visits for schizophrenia or bipolar I disorder, or at least one inpatient visit for major depression during the measurement year *and* diabetes (type 1 and type 2) and whose HbA1c level was greater than 9.0% or was missing a result, or for whom an HbA1c test was not done.	Medicaid (Adult Core Set 2022); ASPE	Yes	No	1	• HbA1c outcome measure is aligned with NCQA diabetes screening of bipolar and schizophrenia patients receiving atypical antipsychotic medications. • Frequency of assessment is at least twice a year when stable at target; frequency is greater at two to three months when HbA1c not at target.
Comprehensive Diabetes Care for People With Serious Mental Illness: Blood Pressure Control (<140/90 mm Hg)	Members 18-75 years of age with at least one acute inpatient visit or two outpatient visits for schizophrenia or bipolar I disorder, or at least one inpatient visit for major depression during the measurement year *and* diabetes (type 1 and type 2) and whose most recent blood pressure screening result was <140/90mm Hg.	NQF	Yes	No	1	• Cut point could be updated with new guidance (130/65 mm Hg).
Diabetes Monitoring for People With Diabetes and Schizophrenia	Assesses adults 18-64 years of age with schizophrenia and diabetes who had both an LDL-C test and an HbA1c test during the measurement year.	NCQA, APA-ADA	Yes	No	2	• Modify the screening and monitoring measure to create an outcome measure consistent with the recommendations of the joint consensus APA/ADA statement on antipsychotic medication. • Create a cut point to indicate LDL-C treatment response.

Notes: ADA = American Diabetes Association; APA-MBHR = American Psychiatric Association – Mental and Behavioral Health Registry; ASPE = Office for the Assistant Secretary for Planning and Evaluation; AUDIT-C = Alcohol Use Disorders Identification Test - Consumption; CCBHC = Certified Community Behavioral Health Clinic; DAST = Drug Abuse Screening Test; LDL-C = Low Density Lipoprotein – Cholesterol; NCQA = National Committee for Quality Assurance; TAPS = Tobacco, Alcohol, Prescription medication and other Substance use.

TIER 2 MEASURES

To create our Tier 2 measure set, we selected transdiagnostic indicators of clinical change. Transdiagnostic measures have great potential for advancing MIC in behavioral health because they are relevant across multiple diagnostic categories and populations served. Such global utility is particularly attractive in behavioral health due to the unreliability of psychiatric diagnoses (Chmielewski et al., 2015) and reliability and validity issues with the widely promoted Global Assessment of Functioning (GAF) score in earlier versions of the Diagnostic and Statistical Manual (Gold, 2014).

Unfortunately, the constructs and measures that we selected for Tier 2 are at a preliminary stage of development. Unresolved issues for future development of transdiagnostic measures include scale selection, standards for use including frequency of assessment, and cutoffs to indicate clinical change. Given these developmental limitations, **we are not endorsing one Tier 2 measure over another at this stage,** but rather proposing candidates and drafts of transdiagnostic measures that other bodies (e.g., NCQA think tank, state Medicaid, CMMI) can use as a foundation for the hard work of developing measure specifications that are clinically meaningful and psychometrically sound. This approach is consistent with recommendations to CMS to "differentiate measures that are proven and appropriate for universal implementation now, as distinct from candidate [measures] that are not yet proven for widespread implementation but represent priority domains for research and development" (DuBard et al., 2023). Criteria for Tier 2 measures to include/exclude are in **Box 5**.

BOX 5. Tier 2 Criteria for What to Include/Exclude

Include

- **Transdiagnostic (i.e., informative for clinical care across diagnoses)**
- **Outcome focused**
- **Patient self-report**
- **Low burden (≤20 items)**
- **Sensitive to change**
- **Suitable for adult community behavioral health**

Exclude

- **Subcomponent of larger issue**

Upon review of these Tier 2 criteria and measures, members of the expert panel advised that we:

- Prioritize clinically relevant concepts reflective of "whole-person" care, including service-user experience and goal attainment.
- Prioritize concepts that showcase the value of behavioral health care (e.g., functional measures, quality of life, recovery).
- Exclude measures of behaviors and/or disposition outcomes with very low base rates (e.g., suicide).
- Reflect the low appetite for more data collection in the public sector by making Tier 2 measures highly selective. For example:
  - Advocate research that produces simpler, valid measures of concepts that minimize burden if no such scales and measures are available.
  - Exclude measures that are difficult to score.
  - Exclude measures that are proprietary, to minimize expense and burden.

Tier 2 unresolved issues at this stage were whether to include constructs related to social and environmental factors that affect health, such as housing and criminal justice involvement. Interested parties also remained concerned about the absence of measures of service-user characteristics (e.g., trauma history) and care processes (e.g., therapeutic alliance [see FIT; Tartakovsky, 2016]), although these are not clinical outcomes of care per se.

Tier 2 constructs and candidate measures for further development are in Table 2. Unlike Tier 1 measures which come from a finite set (CCBHC and Medicaid), our list of Tier 2 measures is long. This is because we are uncertain which Tier 2 measures will emerge as those most acceptable to the majority of interested parties. We also acknowledge that additional Tier 2 candidate measures and scales we have not identified may be suitable for consideration. We encourage organizations that are successfully using Tier 2 measures (those listed here or additional viable candidates) for MIC to publish their work. This will help build the evidence base for their utility and help narrow the focus of the field.

Phase III: Stakeholder Consultation

We presented draft Tier 1 and 2 measure sets at meetings held by the National Council, SAMHSA and National Association of Medicaid Directors (NAMD) to solicit feedback on the approach and findings from a broad group of interested parties. More than 500 interested parties attended these spring 2023 events. Many interested parties reported they are actively working on streamlining and improving their current set of measures and that the approach and measures presented were helpful.

Regarding Tier 1, individual interested parties attending these events recommended that we endorse only billable measures, free and publicly accessible measures, and those that align with CCBHC reporting to maximize the financial sustainability of behavioral health. Interested parties also proposed adding behavioral health measures from additional and forthcoming measure sets, such as California and Mental Health Corporations of America, the latter of which includes benchmarks. Interested parties also confirmed the value of including a subset of medical measures for individuals who take antipsychotic medication.

Table 2. Tier 2 Concepts and Candidate Scales for Consideration and Further Development

Concept	Measure Preliminary Draft	Candidate Scale(s)	Candidate Scale Specification	Free of Charge	Self-Report	# Items	Proposed Implementation or Modification
Disease Self-Management	Person reports adequate knowledge, skill, means and confidence to manage their health care needs.	Gains in Patient Activation (PAM) Scores at 12 Months (Hibbard et al., 2004)	The PAM is a 10- or 13-item questionnaire that assesses an individual´s knowledge, skill and confidence for managing their health and health care. The measure assesses individuals on a 0-100 scale. There are four levels of activation, from low (1) to high (4). The performance score would be the change in score from the baseline measurement to follow-up measurement, or the change in activation score over time for the eligible patients associated with the accountable unit.	Free for research but not for other uses	Yes	10 or 13	• Increase reassessment frequency to every three months. • Specify target change in score (e.g., move up one level).
Functioning	Person reports ability to complete activities of daily living to their satisfaction.	Daily Living Activities (DLA)-20 (Scott & Presmanes, 2001)	The DLA is a brief functional assessment tool for individuals ages 6+ regardless of diagnosis, disability or social background. It provides a 30-day snapshot of 20 domains and a summary of strengths and needs at a specific point related to whole-person health.	Fee for required training before use	Yes	20	• Create categorical cut point to indicate treatment response. • Reassess every three months. • Shorten scale or prioritize among multiple domains.

Concept	Measure Preliminary Draft	Candidate Scale(s)	Candidate Scale Specification	Free of Charge	Self-Report	# Items	Proposed Implementation or Modification
		Patient Reported Outcome Measurement Information System (PROMIS) v1.2 – Global Health Physical 2a and PROMIS Scale v1.2 – Global Health Mental 2a (Hays et al., 2017)	There are two self-report, two-item versions of PROMIS scales to assess general physical and mental health that use the highest discriminating items from longer versions of the scales. Items are rated on five-point Likert scales.	Yes	Yes	4	• Create categorical cut point to indicate treatment response. • Reassess every three months.
		Sheehan Disability Scale (SDS)(Sheehan et al., 1996)	The SDS evaluates symptom impact and impairment in the domains of work/school, social life/leisure and family life/home responsibility within the past week. These three items are responded to on a visual analogue scale ranging through 0 (no impairment), 1-3 (mild), 4-6 (moderate), 7-9 (marked) and 10 (extreme) disability.	Yes	Yes	3 or 5	• Create categorical cut point to indicate treatment response. • Reassess every three months.
		WHODAS 2.0 (Ustün et al., 2010)	The adult self-administered version of the WHODAS is a 12-item measure that assesses disability in adults ages 18 years and older. It assesses disability across six domains, including understanding and communicating, getting around, self-care, getting along with people, life activities (e.g., household, work and/or school activities) and participation in society. Individuals rate how much difficulty they have had in specific areas of functioning during the past 30 days.	Yes	Yes	12	• Create categorical cut point to indicate treatment response. • Reassess every three months. • Limit or prioritize reporting among multiple domains.

Concept	Measure Preliminary Draft	Candidate Scale(s)	Candidate Scale Specification	Free of Charge	Self-Report	# Items	Proposed Implementation or Modification
Goal Attainment	Person reports making progress toward their top priority goals.	Goal Attainment Scaling (GAS)	To use the GAS, a patient identifies two goals among various areas of interest (see list in "What Matters Most Goal Inventory") and rates progress along standardized, five-point Likert scale ("-2 = Much less than expected" through "+2 Much more than expected"), reassessed every three months. Anything at or above is coded as goal achieved (NCQA, 2023).	Yes	Yes	2	• Format for unsupported patient self-report with prompts and patient training/ education.
Patient Experience of Care	Person reports that their experience of receiving care was positive.	CAHPS Experience of Care and Health Outcomes (ECHO) (AHRQ, 2004)	Five CAHPS summary measures (composites) were created and tested for their psychometric properties: getting treatment quickly, how well clinicians communicate, perceived improvement, getting treatment information from the plan and information about treatment options. Item-total correlations were also assessed for the 31 single items with their respective composites. All but four of these single items had item-total correlations above 0.50 across sites.	Yes	Yes	31, but users can select individual subscales	• Shorten scale or limit to particular subscales/ items.
Quality of Life	Person reports that they have a life that is worth living.	Medical Outcomes Study Item Short Form (SF)-12 (Gandek et al., 1998)	The SF-12 is a widely used, generic, patient-report measure created to assess health-related quality of life in the general population. Items describe difficulties with a range of activities and functions and are rated on a five-point scale from "None" to "Extreme or cannot do."	Yes	Yes	12	• Create categorical cut point to indicate treatment response. • Reassess every three months.

Concept	Measure Preliminary Draft	Candidate Scale(s)	Candidate Scale Specification	Free of Charge	Self-Report	# Items	Proposed Implementation or Modification
Recovery	Person reports that their health and wellness enable them to live a self-directed life and strive to reach their full potential.	Herth Hope Index (HHI) (Nayeri et al., 2020)	The HHI, originally published in 1991, was designed to evaluate hope in young people with chronic illnesses. It consists of 12 items. Responses are on a four-point Likert-type scale, from 1 (completely disagree) to 4 (completely agree). Although it is widely used, the factor solutions are inconsistent across studies.	Yes	Yes	12	• Resolve psychometric inconsistencies toward a more stable factor solution.

Regarding Tier 2, individual interested parties recommended considering additional constructs such as adverse childhood experiences (ACEs), integrated care, service-user engagement, quality of care and "other"-reported outcomes such as family and caregiver assessments of service-user wellbeing.

Interested parties also provided feedback on National Council processes for MIC. They suggested that measure set development include general feedback from users (e.g., service users, clinicians) about acceptability, from technology partners (e.g., EHR vendor) about feasibility, and from managed care organizations (MCOs) about suitability for reimbursement purposes. They further requested broad training and resources to support MIC (see Box 6).

BOX 6. Interested Parties' Self-identified Training Needs to Implement Our Tiered Approach to MIC

- **Provide lists of measures including those green-lighted for immediate use and those under consideration for future use.**
- **Create worksheets or handouts to further explain some of the concepts and guide implementation.**
- **Provide training on how to engage clients, share measures information and apply a health literacy-informed approach (NEBGH, 2023).**
- **Provide supports for small providers with limited data collection and analytic resources, including partnership strategies.**
- **Train stakeholder organizations to use MIC data for routine continuous quality improvement.**
- **Add MIC to graduate school curriculum and professional training programs.**

Summary and Recommendations

Defining and supporting MIC can improve behavioral health care quality, outcomes and sustainability. The right set of measures could address clinicians' and organizations' concerns about prioritizing resources and infrastructure to promote MIC in community behavioral health care and beyond. Among service users and clinicians, MIC has potential to efficiently inform care and improve clinical outcomes through the effective use of data (Fortney et al., 2017; Lambert et al., 2018). Within organizations, a tightly curated and payer-aligned measure set can help reduce burden from competing clinical, quality, accountability and fiscal reporting practices (Alter et al., 2021; Jacobs et al., 2023; Ridout et al., 2023). Nationwide, aligned and synthesizable data sets can robustly illustrate the impact of behavioral health services, which can help attract and sustain investment (Alter et al., 2021; Ridout et al., 2023).

This two-tier approach to prioritizing and identifying measures for MIC serves as a template for immediate actions (Tier 1) and near-term innovations (Tier 2) to realize MIC. Tier 1 is feasible and meaningful because it makes use of measures that many behavioral health organizations already use, while also directing quality measurement bodies to consider modifying specifications to lower clinician and system burden while fostering MIC. It also provides a concise set of measures that behavioral health organizations and payers can include when exploring tailored sustainability strategies such as VBP. Our approach to Tier 2, in which we propose concepts and candidate measures instead of coalescing around widely used scales, reflects the state of the field and will require significant stakeholder input and field testing before Tier 2 measures are endorsed by national quality organizations.

RECOMMENDATIONS TO PROMOTE MIC UPTAKE IN THE FIELD AT LARGE

1. NCQA, APA and other national quality organizations should review Tier 1 measures and make efforts to improve the specifications of these measures, to assist in the uptake and scaling of MIC.
2. Federal and state innovation models that support increased access, quality care and VBP in behavioral health should use Tier 1 measures as a core set for incentivization.
3. Organizations should adopt at least one Tier 2 measure in addition to Tier 1 measures to create a shift in measurement toward transdiagnostic, functional and recovery-oriented constructs in place of diagnosis-specific measures. This shift, as previously noted by others (Kilbourne et al., 2018), may result in several benefits to the field of behavioral health, such as a reduction in the number of measures that clinicians and organizations collect and report. These benefits could be particularly apparent for organizations serving varied populations and for individuals and clinicians managing comorbidity, which is increasingly the norm (e.g., service users with bipolar disorder, substance use disorder and PTSD) (Suls et al., 2019). A transdiagnostic, recovery-oriented measurement approach also has promise for shifting payer and regulator focus to the entirety of what behavioral health services help improve, including quality of life and social and environmental factors that affect health, thus more robustly illustrating behavioral health care value.

BOX 7. Recommendations for MIC Implementation by National Council Member Organizations

- **Behavioral health provider organizations should use both Tier 1 and Tier 2 measures in discussions with payers, prioritizing their quality efforts around Tier 1 and pilot testing Tier 2 measures.**
- **Phased implementation:**
  - Low-resource organizations can start with Tier 1 measures (many may already be in place).
  - Better-resourced organizations can choose a wider range of Tier 1 and 2 measures to help lead the field.
  - Pilot studies, continuous quality improvement processes and feedback to interested parties should be used to build buy-in and should be disseminated when possible.
  - Solicit stakeholder input. Ask clinicians and service users about the Tier 1 and 2 measures that you're considering, and use their ideas to optimize implementation.
- **Utilize technical assistance.**
  - Contact the National Council for a list of available resources.

A shift away from diagnostic-specific measures makes sense, given that the purpose of diagnosis is primarily to facilitate treatment matching, not to apportion care. It also circumvents problems with treatment decision-making stemming from errors related to the poor inter-rater reliability of psychiatric diagnoses and the reality that co-occurring conditions widespread. Similarly, although most DSM-5 and International Classification of Diseases (ICD) diagnoses are only made in the presence of significant distress or functional impairment, and although functional impairments are common motivators for individuals to seek treatment (Doll et al., 2021), diagnoses themselves do not indicate nor quantify impairment or suffering. As such, a shift toward explicit functional and recovery-oriented measures has the potential to vastly consolidate behavioral health measurement and improve resource allocation in a way that is consistent with the spirit of the field.

Of course, the complete realization of MIC will depend on more than creating consensus about an optimized measure set. This includes research demonstrating linkages to population outcomes and strategic, systemic investments in health information technology (HIT) for automating administration and integration into the clinician and team workflows, scoring, clinical sharing and reporting of measures in ways that are acceptable to the workforce (i.e., human-centered design). The National Council and other behavioral health leadership bodies can work with federal and state partners and HIT vendors to help support this process by advising on functionality, EHR interoperability, privacy and cost issues that have impeded robust and scalable MIC to date, and by partnering on research that demonstrates the utility of the approach.

BOX 8. Sustainability of MIC

Whether MIC will be sustainable depends on the broad availability of:

- **A curated set of measures that are:**
  - Low burden to patients and providers.
  - Implemented frequently enough to track changes and guide treatment.
  - Widely agreed upon by clinicians, regulators and payers.
- **Comprehensive staff training in MIC principles and practices.**
- **Organizational support (e.g., funding, infrastructure and consultation) to establish and maintain MIC practices.**
- **Use of billing codes already in place for MBC in primary care settings.**
- **Objective, scientific research to demonstrate the societal return on investment of an MIC approach.**

4. Promote sustainability of MIC implementation by limiting the number of standardized measures used across a wide range of regulators and payers. This will require substantial effort toward regulatory actions that must be supported by behavioral health leadership and advocacy organizations. Sustainable MIC implementation will require broad and accessible technical assistance to the behavioral health workforce at all levels, from leadership (e.g., directional guidance and investment), to clinicians (e.g., education on measure administration, scoring and service-user feedback), to administrative and other support staff who will maintain the integrity of the data. Technical assistance must consider differences in organizational culture (e.g., language) and resources (e.g., data analytics) so that all eligible organizations can take part. Technical assistance must be ongoing to account for high staff turnover rates in many behavioral health settings. Additional suggestions for technical assistance are in **Box 6**.

MIC has some unavoidable additional costs, including collecting the information, aggregating it and, in some cases, using proprietary measures or training. Critically, the National Council and other national- and state-level organizations will need to undertake empirical research and substantial advocacy efforts to educate payers about the value of MIC and a consolidated measure set, and to coordinate appropriate remuneration for services informed by MIC in pay for performance (P4P) and VBP models. In the future, this may also include strategies for reporting and considering alternative data sources, including qualitative data, that also inform clinical care. A list of MIC sustainability strategies is in **Box 8**.

Ultimately, MIC can help foster communication between clinical organizations and policymakers about the value of sustained investment in behavioral health. Without buy-in from payers and policymakers at all levels about this approach, however, issues of burden and proliferation of measures will persist (DuBard et al., 2023).

LIMITATIONS

Through our consultations, we learned of other measure sets under development that could be useful for National Council members' implementation of MIC (e.g., nine new behavioral health measures to be put forward by the state of California). The discovery of these sets is important. At the same time, it does not change our overall approach for how organizations should move into their next phase of MIC. By limiting Tier 1 measures to those required by Medicaid and/or CCBHC and endorsed by national quality organizations, we accept that we might exclude potentially viable and worthwhile measures in favor of providing interested parties with a finite (or nearly finite) set of measures ready for use now. This same tension occurs in our selection of candidate measures for Tier 2. While we propose candidate measures to begin the conversation about the kinds of measures that our field needs, researchers likely will need to undertake head-to-head comparisons, or new measure development work, to address our identified measurement gaps. For now, Tier 2 provides a set of criteria and reasonable examples to help guide organizations toward an efficient and aligned set of transdiagnostic measures that can be used preliminarily now, presented for scholarship and demonstration of value, which will continue to be refined over time.

Finally, the MIC measures presented in this paper are unlikely to include everything that individual behavioral health clinicians and organizations will want to measure. For instance, our stakeholder consultations reinforced that some groups cannot envision measure sets without information to enable analyses of quality by subgroup and demonstrate individuals' and populations' responses to care. Given this, we encourage National Council membership to progressively align their data collection with the measures put forward for national initiatives toward MIC, and then complement this set of measures with those that balance their organization's unique informational needs with their infrastructure capacity and tolerance for data collection burden.

Areas for Future Research

Our experts identified dissatisfaction with the reassessment schedules included in Tier 1 measures. While useful for organizational reporting, our experts explained that the reassessment schedules in most measures were too infrequent to illustrate clinical responses to treatment or to inform next steps in care. Thus, we propose that research is needed to determine optimal reassessment schedules that meet clinical objectives and balance service-user, clinician and administrative data burden. Additionally, research is needed to simplify measure scoring and establish clearer cutoffs to ensure compatibility with current EHR capabilities and to facilitate clinician interpretation of clinical change at the point of care. Strategies that cover the unavoidable costs of MIC, including gathering information, aggregating information and in some cases covering fees related to propriety measures, must also be developed.

Research is also needed to solicit broader feedback from interested parties. This includes input from technology partners (e.g., EHR vendors) to improve MIC functionalities, MCOs about what is needed to ensure payment for collecting and reporting measures, and organizational interested parties about specific issues that should be the focus of training and resources (**see Box 6**). Service-user, family and caregiver perspectives also are critically needed to ensure that Tier 2 measures are comprehensive, meaningful, yet acceptable in terms of burden, language and content. MIC cannot improve care if the data it produces have no meaning to service users and the clinicians who support them.

Finally, a major challenge to MIC research that requires further investigation is how to acculturate payers and policymakers to the broader construct of MIC, in which symptom data informs but does not exclusively dictate clinical care, and, in contrast, how to increase population health literacy and clinician comfort with quantitative data to shape behavioral health care. We challenge behavioral health researchers to explore how additional clinical information such as social and environmental factors that affect health, customs and beliefs, and service-user preferences can be aggregated into digestible forms to further describe and provide value to behavioral health care.

Conclusions

Behavioral health organizations are under immense pressure to be accountable to a range of interested parties via data collected through clinical care. The future financial sustainability of behavioral health care depends on it. Organizations need national, state and local leadership to advocate for and provide guidance around meaningful, aligned, compensable and manageable sets of measures that can meet these requirements while staying true to the spirit of behavioral health, in which symptom scales, laboratory and other biometric data inform but do not solely dictate clinical care, as implied in the term "measurement-based care." In this report, we introduced the concept of measurement-informed care, in which a single set of measures is selected to meet organizations' clinical care and reporting needs, while acknowledging that measures and scales are an important part of the comprehensive biopsychosocial approach to quality behavioral health care. To do this, we provided measures that organizations can use to implement MIC in the current policy and fiscal environment, with separate tiers of measures for those that are (almost) ready for off-the-shelf use and those that have tremendous potential to advance the field but for which more research and testing is needed. While organizations of different sizes, resources and workplace cultures likely will implement it differently, this report includes recommendations that individual organizations can use to move the national landscape toward sustainable MIC. It also specifies required advocacy actions the National Council and its partners should take to support its membership in this important undertaking.

References

Agency for Healthcare Research and Quality (AHRQ). (2004). ECHO survey and reporting kit. https://www.ahrq.gov/sites/default/files/wysiwyg/cahps/surveys-guidance/echo/about/209_echo-reporting-measures.pdf

Alter, C. L., Mathias, A., Zahniser, J., Shah, S., Schoenbaum, M., Harbin, H. T., McLaughlin, & Sieger-Walls, J. (2021). Measurement-based care in the treatment of mental health and substance use disorders. Meadows Mental Health Policy Institute. https://mmhpi.org/project/measurement-based-care-in-the-treatment-of-mental-health-substance-use-disorders/

American Diabetes Association. (2019). Glycemic targets: Standards of medical care in diabetes–2020. Diabetes Care, 43(Supplement_1), 66-76. https://doi.org/10.2337/dc20-S006

Brown, J., Hudson Scholle, S., Croake, S., Rosenau, M., Liu, J., Darter, N., Farson Gray, K., Hart, M., Harvey, B., Lewis, R., Miller, D., Morton, S., Patel, M., & Ranshous, C. (2018, April 8). Development and testing of behavioral health quality measures for health plans: Final report. Office for the Assistant Secretary for Planning and Evaluation; U.S. Department of Health and Human Services. https://aspe.hhs.gov/reports/development-testing-behavioral-health-quality-measures-health-plans-final-report-0

Carey, R. M., Wright, J. T. J., Taler, S. J., & Whelton, P. K. (2021). Guideline-driven management of hypertension: An evidence-based update. Circulation Research, 128(7), 827-846. https://doi.org/10.1161/CIRCRESAHA.121.318083

Chmielewski, M., Clark, L. A., Bagby, R. M., & Watson, D. (2015). Method matters: Understanding diagnostic reliability in DSM-IV and DSM-5. Journal of Abnormal Psychology, 124(3), 764-769. https://doi.org/10.1037/abn0000069

Doll, C. M., Michel, C., Rosen, M., Osman, N., Schimmelmann, B. G., & Schultze-Lutter, F. (2021). Predictors of help-seeking behaviour in people with mental health problems: A 3-year prospective community study. BMC Psychiatry, 21(1), 432. https://doi.org/10.1186/s12888-021-03435-4

DuBard, A., Israel, J., & Mostashari, F. (2023). Less is more: Quality measurement in primary care. Health Affairs Forefront. Retrieved [January 4, 2024], from https://doi.org/10.1377/forefront.20230614.546595

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. Psychological Assessment, 9(4), 374-385. https://doi.org/10.1037/1040-3590.9.4.374

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, 68(2), 179-188. https://doi.org/10.1176/appi.ps.201500439

Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., Bullinger, M., Kaasa, S., Leplege, 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. Journal of Clinical Epidemiology, 51(11), 1171-1178. https://doi.org/10.1016/s0895-4356(98)00109-7

Gold, L. H. (2014). DSM-5 and the assessment of functioning: The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0). Journal of the American Academy of Psychiatry and the Law Online, 42(2), 173-181. https://pubmed.ncbi.nlm.nih.gov/24986344/

Hallgren, K. A., Cohn, E. B., Ries, R. K., & Atkins, D. C. (2022). Delivering remote measurement-based care in community addiction treatment: Engagement and usability over a 6-month clinical pilot. Frontiers in Psychiatry, 13. https://doi.org/10.3389/fpsyt.2022.840409

Hays, R. D., Schalet, B. D., Spritzer, K. L., & Cella, D. (2017). Two-item PROMIS® global physical and mental health scales. Journal of Patient-Reported Outcomes, 1(1). https://doi.org/10.1186/s41687-017-0003-8

Hibbard, J. H., Stockard, J., Mahoney, E. R., & Tusler, M. (2004). Development of the Patient Activation Measure (PAM): Conceptualizing and measuring activation in patients and consumers. Health Services Research, 39(4), 1005-1026. https://doi.org/10.1111/j.1475-6773.2004.00269.x

Horner-Johnson, W., Krahn, G., Andresen, E., & Hall, T. (2009). Developing summary scores of health-related quality of life for a population-based survey. Public Health Reports, 124(1), 103-110. https://doi.org/10.1177/003335490912400113

Jacobs, D. B., Schreiber, M., Seshamani, M., Tsai, D., Fowler, E., & Fleisher, L. A. (2023). Aligning quality measures across CMS — The universal foundation. The New England Journal of Medicine, 388(9), 776-779. https://doi.org/10.1056/NEJMp2215539

Kendrick, T., El-Gohary, M., Stuart, B., Gilbody, S., Churchill, R., Aiken, L., Bhattacharya, A., Gimson, A., Brütt, A. L., de Jong, K., & Moore, M. (2016). Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. The Cochrane Database of Systematic Reviews, 7(7). https://doi.org/10.1002/14651858.CD011119.pub2

Kilbourne, A. M., Beck, K., Spaeth-Rublee, B., Ramanuj, P., O'Brien, R. W., Tomoyasu, N., & Pincus, H. A. (2018). Measuring and improving the quality of mental health care: A global perspective. World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 17(1), 30-38. https://doi.org/10.1002/wps.20482

Ko, H., Gatto, A. J., Jones, S. B., O'Brien, V. C., McNamara, R. S., Tenzer, M. M., Sharp, H. D., Kablinger, A. S., & Cooper, L. D. (2023). Improving measurement-based care implementation in adult ambulatory psychiatry: A virtual focus group interview with multidisciplinary healthcare professionals. BMC Health Services Research, 23(1), 408. https://doi.org/10.1186/s12913-023-09202-3

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. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. Psychotherapy, 55(4), 520-537. https://doi.org/10.1037/pst0000167

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. JAMA Psychiatry, 76(3), 324-335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Medicaid. (2023). Core set of Health Home quality measures for 1945 Medicaid Health Home programs (1945 Health Home core set), 2013-2024. https://www.medicaid.gov/sites/default/files/2023-03/2023-health-home-core-set-history-table.pdf

National Committee for Quality Assurance. (2023). Advancing best practices for goal attainment scaling. https://www.ncqa.org/white-papers/advancing-best-practices-for-goal-attainment-scaling/

Nayeri, N. D., Goudarzian, A. H., Herth, K., Naghavi, N., Nia, H. S., Yaghoobzadeh, A., Sharif, S. P., & Allen, K.-A. (2020). Construct validity of the Herth Hope Index: A systematic review. International Journal of Health Sciences, 14(5), 50-57. https://pubmed.ncbi.nlm.nih.gov/32952505/

Northeast Business Group on Health. (2023). Measurement-based care fact sheet. https://online.flippingbook.com/view/722089316/

Ridout, K. K., Vanderlip, E. R., Ridout, S. J., & Alter, C.L. (2023). Measurement-based care implementation challenges and opportunities in real-world practice. Psychiatric Annals, 53(9), 405-409. https://doi.org/10.3928/00485713-20230818-02

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. https://doi.org/10.1016/j.jclinepi.2015.08.036

Scott, R. L., & Presmanes, W. S. (2001). Reliability and validity of the Daily Living Activities scale: A functional assessment measure for severe mental disorders. Research on Social Work Practice, 11(3), 373-389. https://doi.org/10.1177/104973150101100306

Shahid, R., Shoker, M., Chu, L. M., Frehlick, R., Ward, H., & Pahwa, P. (2022). Impact of low health literacy on patients' health outcomes: A multicenter cohort study. BMC Health Services Research, 22(1), 1148. https://doi.org/10.1186/s12913-022-08527-9

Sheehan, D. V., Harnett-Sheehan, K., & Raj, B. A. (1996). The measurement of disability. International Clinical Psychopharmacology, 11, 89-95. https://journals.lww.com/intclinpsychopharm/fulltext/1996/06003/the_measurement_of_disability.15.aspx

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. Archives of Internal Medicine, 166(10), 1092-1097. https://doi.org/10.1001/archinte.166.10.1092

Suls, J., Green, P. A., & Boyd, C. M. (2019). Multimorbidity: Implications and directions for health psychology and behavioral medicine. Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association, 38(9), 772-782. https://doi.org/10.1037/hea0000762

Tartakovsky, M. (2016). Feedback-informed treatment: Empowering clients to use their voices. PsychCentral. https://psychcentral.com/lib/feedback-informed-treatment-empowering-clients-to-use-their-voices#1

Ustün, T. B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korff, M., & Pull, C. (2010). Developing the World Health Organization Disability Assessment Schedule 2.0. Bulletin of the World Health Organization, 88(11), 815-823. https://doi.org/10.2471/BLT.09.067231

Van Tiem, J., Wirtz, E., Suiter, N., Heeren, A., Fuhrmeister, L., Fortney, J., Reisinger, H., & Turvey, C. (2022). The implementation of measurement-based care in the context of telemedicine: Qualitative study. JMIR Mental Health, 9(11). https://doi.org/10.2196/41601

World Health Organization. (n.d.). WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0. https://www.cms.gov/files/document/whodas-20-instrument.pdf

Wrenn, G., & Fortney, J. (2015). A core set of outcome measures for behavioral health across service settings. The Kennedy Forum. https://pjk-wp-uploads.s3.amazonaws.com/www.thekennedyforum.org/uploads/2017/06/MBC_supplement.pdf