NATIONAL QUALITY FORUM
Driving measurable health improvements together

Behavioral Health and Substance Use, Spring 2022 Cycle: CDP Report

TECHNICAL REPORT
JANUARY 30, 2023

This report is funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I Task Order HHSM-500-T0001.

https://www.qualityforum.org

Contents

Executive Summary

Behavioral health looks at how human behaviors and choices impact mental and physical health. Behavioral health comprises broader concepts, including mental wellness and substance use disorders (SUDs). While many disorders that fall under the behavioral health umbrella are often chronic, people can recover with timely, high quality, coordinated, and evidence-based care. Quality measurement and quality improvement tools remain an important aspect of assessing and improving the treatment of behavioral health conditions.

At present, there are 28 NQF-endorsed behavioral health measures, including topics such as alcohol and drug use, care coordination, depression, medication use, experience of care, tobacco, and physical health.

For this cycle, the Behavioral Health and Substance Use (BHSU) Standing Committee evaluated seven measures undergoing maintenance review against NQF's standard evaluation criteria. The Standing Committee recommended four measures for endorsement but did not recommend the three remaining measures. The Consensus Standards Approval Committee (CSAC) upheld the Standing Committee's recommendations.

The following measures were endorsed:

- NQF #3312 Continuity of Care After Medically Managed Withdrawal From Alcohol and/or Drugs (Centers for Medicare & Medicaid Services [CMS]/Lewin Group)
- NQF #0710e Depression Remission at 12 Months (MN Community Measurement)
- NQF #0711 Depression Remission at Six Months (MN Community Measurement)
- NQF #1884 Depression Response at Six Months – Progress Towards Remission (MN Community Measurement)

The following measures were not endorsed:

- NQF #3313 Follow-Up Care for Adult Medicaid Beneficiaries Who Are Newly Prescribed an Antipsychotic Medication (CMS/Mathematica Policy Research)
- NQF #1885 Depression Response at 12 Months – Progress Towards Remission (MN Community Measurement)
- NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M (MN Community Measurement)

Brief summaries of the measures and their evaluations are included in the body of the report; detailed summaries of the Standing Committee's discussion and ratings of the criteria for each measure are in Appendix A.

Introduction

Behavioral health challenges affect millions of Americans, their families, and communities. One in five U.S. adults, and 1 in 6 children, experience mental illness each year.[1,2] In 2020, 17 million adults experienced co-occurring mental illness and SUD.[1] The coronavirus disease 2019 (COVID-19) pandemic led to additional widespread reports of psychological distress.[3] In addition, testing positive for COVID-19 was associated with a greater likelihood of suicide ideation in those who had other predisposing factors, such as major depressive disorder, alcohol use disorder, or generalized anxiety disorder.[4] The World Health Organization (WHO) recommends the application of approaches that promote mental healthcare, respond to mental health emergencies, and build supportive services for the future.[5] The spring 2022 cycle includes a review of BHSU measures addressing depression screening, response, and remission; follow-up care for medication management; and continuity of care for withdrawal management.

Depression Screening, Response, and Remission

Depression prevalence in the U.S. is 5.9 percent and is the second highest in the world with over 17 million cases.[6] Depression is often treatable but remains underdiagnosed and undertreated, which leads to a higher incidence of poor outcomes.[7–9] Detection and prevention efforts are critical to mitigating adverse outcomes and supporting the health and wellness of many Americans. Recommendations state that routine depression screening for adults over the age of 18 should be combined with suitable interpretation of screening results and implementation of interventions as needed for patient care.[10] The Standing Committee evaluated five measures this cycle that assessed the use of the Patient Health Questionnaire-9 (PHQ-9) and the Patient Health Questionnaire Modified for Teens (PHQ-9M) for screening for depression and determining response and remission rates over a period of time (NQF #0710e, NQF #0711, NQF #0712, NQF #1884, and NQF #1885).

Continuity of Care/Follow-Up Care

Fragmented healthcare leads to unplanned emergency room visits; increased admissions for serious illness; and lack of consistent management of chronic illness, including mental illness.[11] Continuity of care allows for a strong patient-provider relationship that allows for better health outcomes for patients battling behavioral health issues. In addition, patients who seek treatment and experience withdrawal symptoms are susceptible to relapse and returning to drug use without withdrawal management, which is designed to support patients through a potentially stressful transition to recovery.[12] This can include both psychological support and medication management to treat and manage persistent physical symptoms.[13] Consistent management of care allows providers to carefully manage symptoms and medication usage. Once a new medication regimen is begun, follow-up is needed to ensure a medication's effect is as desired.[11] The Standing Committee evaluated two measures this cycle that assess the continuity of care after medically managed withdrawal from alcohol and/or drugs (NQF #3312) and follow-up care for adult Medicaid beneficiaries who are newly prescribed an antipsychotic medication (NQF #3313).

NQF Portfolio of Performance Measures for Behavioral Health and Substance Use Conditions

The BHSU Standing Committee (Appendix C) oversees NQF's portfolio of BHSU measures (Appendix B), including measures for alcohol and drug use, care coordination, depression, medication use, experience of care, tobacco, and physical health. This portfolio contains 26 measures: 24 process measures, one outcome and resource use measure, and one patient-reported outcome performance measure (PRO-PM).

Additional measures have been assigned to other portfolios. These include healthcare-associated infection measures (Patient Safety), care coordination measures (Geriatrics and Palliative Care), imaging efficiency measures (Cost and Resource Use), and a variety of condition- or procedure-specific outcome measures (Cardiovascular, Cancer, Renal, etc.).

Behavioral Health and Substance Use Measure Evaluation

On June 30, 2022, the BHSU Standing Committee evaluated seven measures undergoing maintenance review against NQF's standard measure evaluation criteria.

Table 1. Behavioral Health and Substance Use Measure Evaluation Summary

Measure	Maintenance	New	Total
Measures under review for endorsement	7	0	7
Measures endorsed	4	0	4
Measures not endorsed	3	0	3
Reasons for not endorsing	Importance-1 Scientific Acceptibility-2	-	-

Cells marked by a dash (-) are intentionally left blank.

Comments Received Prior to Standing Committee Evaluation

NQF accepts comments on endorsed measures on an ongoing basis through the Quality Positioning System (QPS). In addition, NQF solicits comments for a continuous period during each evaluation cycle via an online tool located on the project webpage. For this evaluation cycle, the commenting period opened on May 18, 2022, and pre-meeting commenting closed on June 15, 2022. Prior to June 15, 16 comments were submitted and shared with the Standing Committee prior to the measure evaluation meeting(s) (Appendix F).

Comments Received After Standing Committee Evaluation

The continuous public commenting period with NQF member support closed on September 13, 2022. Following the Standing Committee's evaluation of the measures under review, NQF received 13 comments from four organizations (including two member organizations) and individuals pertaining to the draft report and the measures under review (Appendix G). All comments for each measure under review have also been summarized in Appendix A.

NQF members had the opportunity to express their support ("support" or "do not support") for each measure submitted for endorsement consideration to inform the Standing Committee's recommendations during the commenting period. One NQF member expressed "do not support" for NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712.

Overarching Themes

During the Standing Committee's discussion of the measures, one overarching issue emerged that was factored into the Standing Committee's ratings and recommendations for multiple measures and is not repeated in detail with each individual measure.

Lack of Telehealth in Behavioral Health and Substance Use Measures

The Standing Committee raised concerns regarding the lack of telehealth services included in the specifications of many of the measures. The Standing Committee expressed concerns about whether measures took telephonic or telemedicine follow-up into account, given the increase of these services during the COVID-19 pandemic. Some developers confirmed that telehealth services were included in their measures, and other developers clarified that since the data provided were pre-pandemic data, telehealth codes were not currently included in the measure; however, they planned to track Medicaid telehealth claims as a feature of the measure in the future. The Standing Committee asked NQF whether there was a precedent that new information or updated data could trigger an early review of a measure. NQF staff clarified the situations in which an endorsed measure might be reviewed earlier than scheduled but emphasized that the Standing Committee must review the measure now as currently submitted.

Summary of Measure Evaluation

The following brief summaries of the measure evaluation highlight the major issues that the Standing Committee considered. Details of the Standing Committee's discussion and ratings of the criteria for each measure are included in Appendix A.

Continuity of Care/Follow-Up Care

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS (CENTERS FOR MEDICARE & MEDICAID SERVICES [CMS]/LEWIN GROUP): ENDORSED

Description: Percentage of discharges from a medically managed withdrawal episode for adult Medicaid beneficiaries, ages 18–64, that were followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder [pharmacotherapy]) within 7 or 14 days after discharge; **Measure Type**: Process; **Level of Analysis**: Regional and State; Population: Population; **Setting of Care**: Inpatient/Hospital, Outpatient Services; **Data Source**: Claims

This population-level measure was originally endorsed in 2018 and was previously used in the Medicaid Innovation Accelerator Program (IAP), which ended in 2020. It is planned for voluntary use for state Medicaid programs but is not currently used in an accountability program.

The Standing Committee noted that the developer added additional evidence supporting the measure and that the data showed a clear gap in performance. The Standing Committee passed the measure on both criteria.

The Standing Committee agreed that the measure was reliable based on the reliability testing results. As mentioned in the *Overarching Issues* section, the Standing Committee expressed concerns about whether the measure took telephonic or telemedicine follow-up into account, given the increase of these services during the COVID-19 pandemic. The developer noted that this measure did not currently include telehealth services. During the discussion of validity, the Standing Committee asked whether patients who take monthly medications would affect the 14-day rate compared to the seven-day rate. The developer responded by explaining that seven-day follow-up is considered the standard of care, but they allowed for some flexibility by including the 14-day follow-up as well. The developer also explained that encounters that occur on the day of discharge also count toward the measure; therefore, receiving a medication at discharge that is renewed every 30 days would still be counted in the numerator. The Standing Committee did not have any further questions and passed the measure on validity.

The Standing Committee had no concerns with feasibility and passed the measure on this criterion. The Standing Committee asked how many states currently use this measure, and the developer indicated that these data are not available via Medicaid because states may not choose to use this measure since they are allowed to choose what they want to measure under Medicaid. The data used for testing came from nine states that are using the measure. The Standing Committee passed the measure on use. The Standing Committee also had no concerns with usability and passed the measure on usability and overall suitability for endorsement.

Two public or member comments were received during the commenting period. One comment expressed concern that the measure in its current state is not consistent with the World Health Organization's (WHO) recommendations to apply approaches that build supportive services for the future. The other comment, which applied to both this measure and NQF #3313, expressed concerns regarding the age range and payer population of the measure. The Standing Committee reviewed the developer's responses and determined the measure required no further discussion on the public comments.

During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and endorsed the measure. No appeals were received.

NQF #3313 FOLLOW-UP CARE FOR ADULT MEDICAID BENEFICIARIES WHO ARE NEWLY PRESCRIBED AN ANTIPSYCHOTIC MEDICATION (CMS/LEWIN GROUP): NOT ENDORSED

Description: Percentage of new antipsychotic prescriptions for Medicaid beneficiaries age 18 years and older who have completed a follow-up visit with a provider with prescribing authority within four weeks (28 days) of prescription of an antipsychotic medication; **Measure Type**: Process; **Level of Analysis**: Regional and State; Population: Population; **Setting of Care**: Outpatient Services; **Data Source**: Claims

This population-level measure was originally endorsed in 2018 and was previously used in the Medicaid IAP, which ended in 2020. It is planned for voluntary use for state Medicaid programs but is not currently used in an accountability program.

The Standing Committee had no concerns with the evidence provided; it agreed that a notable performance gap remains and passed the measure on evidence and performance gap. Regarding the measure specifications, a Standing Committee member asked about the lookback time frame that

determines whether this measure captures a new or a reinitiated prescription. The developer clarified that after 120 days, or four months, the patient would be considered a new user again. The Standing Committee raised a concern that the measure looks at whether a follow-up visit occurred rather than a follow-up visit specific to the antipsychotic prescription the patient received. The developer acknowledged a limitation of claims data: They do not distinguish whether the follow-up visit was specific to antipsychotic use. The Standing Committee asked how telehealth was accounted for in this measure. The developer noted that telehealth codes are not accounted for in the value sets. The Standing Committee recommended these be considered in the future and passed the measure on reliability.

It also found the validity testing sufficient but questioned whether a follow-up visit within 28 days can address potential physical health issues, such as metabolic syndrome, or whether a more extended period is more likely to show these effects effectively. There was an additional concern that community-based workers who do not bill under a provider's National Provider Identifier (NPI) number would not be captured by the measure. In response, the developer stated that they would discuss these considerations with their technical expert advisory panel. The Standing Committee noted that the measure's inability to capture certain types of follow-up visits, such as community health workers or registered nurses, means it is not counting certain types of progress. Due to these concerns, the Standing Committee was not able to reach consensus on validity. However, the Standing Committee was able to continue the discussion and voting of the remaining criteria except for overall suitability.

The measure is readily captured by Medicaid claims data. The Standing Committee had no concerns and passed the measure on feasibility. The Standing Committee also asked how states use the measure since this information was not available in the submission. The developer clarified that this measure is flagged for use in the CMS section 1115 Medicaid waiver. The Standing Committee did not have any further questions and passed the measure on use. In addition, the Standing Committee had no concerns regarding usability and passed the measure on this criterion.

The Standing Committee did not vote on the recommendation for endorsement because it did not reach consensus on validity—a must-pass criterion. The Standing Committee revoted on validity during the post-comment web meeting.

During the post-comment meeting, the Standing Committee reviewed one public comment relating to NQF #3313, which expressed concerns regarding the age range and payer population within the measure. The Standing Committee reviewed the developer's response and determined the measure required no further discussion on the public comment. The Standing Committee discussed validity concerns about aspects of the measure that may result in inaccurate results, including telemedicine visits not being included in the measure, sufficient follow-up time, and whether providers who conduct follow-up visits would be captured. Based on this discussion and review of public comments, the Standing Committee re-voted and did not pass the measure on validity and therefore did not vote on overall suitability for endorsement.

During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and did not endorse the measure.

Depression Screening, Response, and Remission

NQF #0710E DEPRESSION REMISSION AT 12 MONTHS (MN COMMUNITY MEASUREMENT): ENDORSED

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission twelve months (+/- 60 days) after an index visit; **Measure Type**: Outcome: PRO-PM; **Level of Analysis**: Clinician: Group/Practice; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

This clinician group-level measure was originally endorsed in 2011 and retained endorsement in 2016. It is currently used in the CMS Quality Payment Programs Merit-Based Incentive Payment System (MIPS) Track. This measure is also publicly reported on Minnesota Community Measurement's (MNCM) consumer-facing website MN HealthScores.

During the discussion on evidence, the Standing Committee asked for clarification regarding the developer's choice to use the PHQ-9 instead of other tools that focus on suicidality in relation to depression, such as the Ask Suicide-Screening Questions (ASQ) screening tool. The developer clarified that they did examine 21 other tools with comparable cut points for when the patient was in remission to the PHQ-9 and noted that the suicidality is not in scope for this measure; therefore, a tool that covers suicidality was not chosen. The Standing Committee passed the measure on evidence. The Standing Committee further agreed that there is a substantial gap to warrant this measure.

The Standing Committee had no concerns with the specifications, nor were any concerns raised with respect to the reliability testing, aside from the telehealth concern noted in the *Overarching Issues* section. Therefore, the Standing Committee passed the measure on reliability. For validity, a Standing Committee member asked for more information about the performance score of 22 percent of adults showing remission. The developer clarified that the numerator data only included patients with whom a follow-up visit was conducted, noting that some remission cases may be lost due to lack of follow-up. One Standing Committee member noted that it is difficult to interpret the validity data, given the high rate of patients who are lost to follow-up in the denominator that may be artificially lowering the performance. Other Standing Committee members agreed with the concern regarding patients who are lost to follow-up, noting that the goal of the measure is to improve care and that systems can be developed to aid in better follow-up. The developer noted that one goal of NQF #0710e was to address this known gap in care related to patients with depression who are lost to follow-up, which the developer stated is estimated to be as high as 80 percent in the measure. Having a measure that removes patients who are lost to follow-up creates the status quo. Due to the concerns regarding the loss to follow-up, the Standing Committee did not reach consensus on validity. However, the Standing Committee was able to continue the discussion and voting of the remaining criteria except for overall suitability.

A Standing Committee member asked for clarification regarding whether clinics get reimbursed for sending the measure data to the developer and what organization is paying for establishing and managing the data repository. The developer clarified that the PHQ-9 data are extracted from the Epic electronic health records (EHRs) and that there is no cost for participation. In addition, several health plans include these measures in their pay-for-performance contracts; however, MNCM does not reward providers for participation. The Standing Committee asked for clarification on the maintenance and staffing of the MNCM registry, to which the developer clarified that MNCM was funded by health plan

and medical group-member dues, state government contracts, grant funding, and other various sources of funding. When asked whether this measure would be reported on a national level, the developer clarified that they are not aware of other states collecting this measure statewide, but it is in CMS' pay-for-performance programming. The developer suggested that other states may be collecting this measure as part of value-based payment contracts. A Standing Committee member raised the following concern: This measure may be difficult to collect at the national level, given that not everyone has the infrastructure in place. The Standing Committee had no further concerns and passed the measure on use.

There were no usability data due to recent changes to the measure specifications since the last endorsement. The Standing Committee asked NQF to clarify whether this is a requirement. NQF staff replied that there are times when trend data may not be available for a measure and the Standing Committee would need to deliberate on whether the rationale the developer has provided is acceptable. The Standing Committee passed the measure on usability. The Standing Committee did not vote on the recommendation for endorsement because it did not reach consensus on validity—a must-pass criterion.

During the post-comment meeting, the Standing Committee reviewed several submitted comments that applied to the suite of depression measures, including two comments submitted by the developer responding to concerns mentioned by the Standing Committee during the measure evaluation meeting, two non-supportive comments, one supportive comment, and one comment requesting clarity from NQF on whether tools are endorsed by NQF. The developer also provided two comments with clarifications regarding data elements, measure construct, and telehealth services. The Standing Committee agreed that the developer's clarifications alleviated any concerns it had about validity and voted to pass the measure on validity and overall suitability for endorsement. During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and endorsed the measure. No appeals were received.

NQF #0711 DEPRESSION REMISSION AT SIX MONTHS (MN COMMUNITY MEASUREMENT): ENDORSED

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission six months (+/- 60 days) after an index visit; **Measure Type**: Outcome: PRO-PM; **Level of Analysis**: Clinician: Group/Practice; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

This clinician group-level measure was originally endorsed in 2011 and retained endorsement in 2016. It is currently used in the Minnesota Department of Health's Statewide Quality Reporting and Measurement System (SQRMS). This measure is also publicly reported on MNCM's consumer-facing website MN HealthScores.

The Standing Committee noted that the evidence for this measure was largely the same as the evidence for NQF #0710e. A Standing Committee member noted that an advantage of remission being examined at six months versus 12 months is the ability for clinicians to have more control over the outcomes. The Standing Committee had no concerns and passed the measure on evidence for both the clinician group-levels of analysis.

During the discussion on performance gap, a Standing Committee member noted that, similar to NQF #0710e, performance may be underestimated due to missing data that are still counted in the denominator, which is a validity concern. However, in evaluating the performance gap data, the Standing Committee noted that a gap existed under the measure's current specifications and passed the measure on this criterion.

The Standing Committee did not have any new concerns and passed the measure on reliability. For validity, the Standing Committee noted similar concerns with the previous measure (NQF #0710e), namely that patients, who are lost to follow-up, are included in the denominator, which may underestimate the performance results. Therefore, the Standing Committee did not reach consensus on validity.

The Standing Committee had similar feasibility concerns to the ones discussed during the review of NQF #0710e, specifically that the measure is largely dependent on having a registry; however, the Standing Committee noted that the measure is currently captured in EHRs and has no associated fees. The Standing Committee had no concerns with the measure's use and agreed that the benefits for patients being treated for depression outweighed any possible unintended consequences. Therefore, the Standing Committee passed the measure on feasibility, use, and usability. The Standing Committee did not vote on the recommendation for endorsement because it did not reach consensus on validity—a must-pass criterion.

As noted above in the summary of NQF #0710e, the Standing Committee reviewed several comments during the post-comment meeting that applied to the suite of depression measures. The Standing Committee agreed that the developer's clarifications alleviated any concerns it had about validity and voted to pass the measure on validity and overall suitability for endorsement. During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and endorsed the measure. No appeals were received.

NQF #1884 DEPRESSION RESPONSE AT SIX MONTHS – PROGRESS TOWARDS REMISSION (MN COMMUNITY MEASUREMENT): ENDORSED

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission twelve months (+/- 60 days) after an index visit; **Measure Type**: Outcome: PRO-PM; **Level of Analysis**: Clinician: Group/Practice; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

This clinician group-level measure was originally endorsed in 2014. It is currently used in the Core Quality Measures Collaborative (CQMC) 2020 core measure set for the Behavioral Health specialty. This measure is also publicly reported on MNCM's consumer-facing website MN HealthScores.

The Standing Committee discussed that two points in time are needed to calculate the measure: the evidence regarding targeting a score decrease of five points versus a 50 percent reduction in the score. Being satisfied with the developer's responses, the Standing Committee did not raise any additional concerns and passed the measure on evidence. The Standing Committee agreed that the measure performance data showed a sufficient gap and passed the measure on performance gap.

The Standing Committee had no concerns with the specifications, nor were any concerns raised with respect to the reliability testing, aside from the telehealth concern noted in the *Overarching Issues* section. Therefore, the Standing Committee passed the measure on reliability. With respect to validity, the Standing Committee raised similar concerns as with NQF #0710e and did not reach consensus on validity. However, the Standing Committee was able to continue the discussion and voting of the remaining criteria except for overall suitability.

The Standing Committee noted that this measure is captured in EHRs. It also raised a similar concern: The measure might only be feasible with a registry, and it also might be challenging for providers to report on a six-month time frame. Other Standing Committee members noted that the six-month time frame was an advantage toward detecting treatment resistance earlier. The Standing Committee passed the measure on feasibility. It noted that on a national scale, states may choose to use one measure in this suite of measures rather than all of them, which may pose some challenges; however, the measure is currently used in Minnesota. The Standing Committee passed the measure on use. A Standing Committee member raised a concern: Psychosocially complex patients may take longer than six months to show improvement with treatment, and as a result, they may not be captured in the numerator of this measure. Despite this concern, the Standing Committee passed the measure on usability.

The Standing Committee did not vote on the recommendation for endorsement because it did not reach consensus on validity—a must-pass criterion.

As noted above in the summary of NQF #0710e, the Standing Committee reviewed several comments during the post-comment meeting that applied to the suite of depression measures. The Standing Committee agreed that the developer's clarifications alleviated any concerns it had about validity and voted to pass the measure on validity and overall suitability for endorsement. During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and endorsed the measure. No appeals were received.

NQF #1885 DEPRESSION RESPONSE AT 12 MONTHS – PROGRESS TOWARDS REMISSION (MN COMMUNITY MEASUREMENT): NOT ENDORSED

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who demonstrated a response to treatment twelve months (+/- 60 days) after an index visit; **Measure Type**: Outcome: PRO-PM; **Level of Analysis**: Clinician: Group/Practice; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

This clinician-group level measure was originally endorsed in 2014 and is currently used as a quality metric in CMS' Center for Medicare & Medicaid Innovation (CMMI) Model Kidney Care First. This measure is also publicly reported on MNCMs consumer-facing website MN HealthScores.

The Standing Committee agreed that strong empirical evidence was provided and passed the measure on this criterion. The Standing Committee did not raise any concerns during the discussion on performance gap, noting that it was similar to the previous measure (NQF #1884). Therefore, the Standing Committee passed the measure on performance gap.

The Standing Committee also found the reliability testing to be similar to NQF #1884, noting the telehealth concerns as noted in the *Overarching Issues* section. The Standing Committee passed the

measure on reliability. Regarding validity, the Standing Committee noted that the exclusions were clinically appropriate, and the risk adjustment was appropriately conducted. Another Standing Committee member raised concerns with this measure's failure to account for the progress-relapse-progress course of the disease, noting that the time frame may capture a period of disease relapse instead of progress. Given these concerns, the Standing Committee did not reach consensus on validity.

The Standing Committee noted that some of the same discussion on feasibility from NQF #1884 applied to this measure, namely that it would be easier to report for entities that maintained a data registry. The Standing Committee had no other concerns and passed the measure on feasibility. The measure is also currently in use, and the Standing Committee noted that the benefits of this measure greatly outweigh the potential harms. The Standing Committee passed the measure on use and usability. The Standing Committee did not vote on the recommendation for endorsement because it did not reach consensus on validity—a must-pass criterion.

As noted above in the summary of NQF #0710e, the Standing Committee reviewed several comments during the post-comment meeting that applied to the suite of depression measures. Based on the developer's clarification and public comments, , the Standing Committee re-voted and passed the measure on validity. During the discussion on overall suitability for endorsement, an additional concern was raised: This measure may not be as clinically important as the other measures. The Standing Committee expressed that this measure did not demonstrate sufficient effort to achieve remission, particularly at 12 months versus six months. Specifically, if there is no progress toward remission at 12 months, there is a major issue, whereas six months would be a more reasonable expectation for progress. Based on this discussion, the Standing Committee did not pass the measure on overall suitability for endorsement. During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and did not endorse the measure.

NQF #0712 DEPRESSION ASSESSMENT WITH PHQ-9/PHQ-9M (MN COMMUNITY MEASUREMENT): NOT ENDORSED

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with a diagnosis of major depression or dysthymia who have a completed PHQ-9 or PHQ-9M tool during the measurement period; **Measure Type**: Process; **Level of Analysis**: Clinician: Group/Practice; **Setting of Care**: Outpatient Services; **Data Source**: Electronic Health Records

This clinician group-level measure was originally endorsed in 2011 and retained endorsement in 2016. It is also publicly reported on MNCMs consumer-facing website MN HealthScores.

The Standing Committee noted that the evidence for this measure would be stronger if it were linked to improved outcomes, which would also allow for more meaningful quality improvement. The Standing Committee also recognized the challenges of providing such evidence and discussed that not administering a PHQ-9 could result in missed diagnoses. The Standing Committee was not able to reach consensus on evidence. The Standing Committee also noted that the developer provided sufficient data showing a performance gap. Therefore, the Standing Committee passed the measure on performance gap.

The Standing Committee had no concerns with the specifications, nor were any concerns raised with respect to the reliability testing, aside from the telehealth concern noted in the *Overarching Issues* section. Therefore, the Standing Committee passed the measure on reliability. The Standing Committee also asked for clarification regarding the effect of frequency of PHQ-9 assessment. The developer clarified that patients were divided into those who had received only one to three PHQ-9 assessments versus those who received four to 12 PHQ-9 assessments. For patients who received one to three PHQ-9 assessments, remissions rates were 6.3 compared to 15.8 for those assessed more frequently (i.e., four to 12 times). A Standing Committee member asked whether the completion data element was reported as a "yes" or "no" for completion of the tool, or whether it is a full metric of the scoring. The developer clarified that the expectation has always been completion of the tool and that an incomplete tool does not count. The developer also clarified that other related tools, such as the PHQ-8, PHQ-2, or PHQ-3, cannot be reported for this measure. The Standing Committee had no concerns regarding the reliability testing and passed the measure on this criterion. The Standing Committee also noted that the validity testing at the accountable-entity level showed a relatively weak correlation between NQF #0712 and NQF #0711. For adults, the correlation between assessment with PHQ-9/PHQ-9M and *Depression Remission at Six Months* was the R-squared value of 0.1754, and for adolescents, the R-squared value was 0.2744. The Standing Committee passed the measure on validity.

The Standing Committee noted that the data are collected during the regular course of care and are well integrated into most EHRs; it also noted that the PHQ-9 screening tool is free and publicly available. The Standing Committee passed the measure on feasibility. The Standing Committee also noted that the measure is in use and publicly reported in all primary care clinics in Minnesota and in many bordering communities; it passed the measure on use. A Standing Committee member noted that the PHQ-9 may not be a standard part of care in many settings since some accountability organizations allow for use of a variety of validated screening tools; however, another member noted that the measure has shown gradual improvement over time. The Standing Committee had no concerns about unintended consequences and passed the measure on usability. The Standing Committee did not vote on the recommendation for endorsement because it did not reach consensus on evidence—a must-pass criterion.

As noted above in the summary of NQF #0710e, the Standing Committee reviewed several comments during the post-comment meeting that applied to the suite of depression measures. In response to the Standing Committee's concerns about the evidence, the developer provided new data from 26,000 patients to assess whether the frequency of the PHQ-9 assessment was associated with outcomes, noting that patients with three to 12 PHQ-9 assessments were three times more likely to achieve depression remission and response to treatment. The Standing Committee noted that the additional evidence the developer provided did not directly address the question about whether a single assessment of PHQ-9 was associated with improved outcomes. Specifically, the measure does not assess the number of assessments, only that an assessment was conducted. The Standing Committee suggested that creating a composite of this measure and the other suite of remission and response depression measures would strengthen this measure. The developer responded by explaining that while this is a companion measure to the suite of measures, they did not think a composite would be appropriate. The developer noted that this measure was created to generate information for the other depression measures, and without it, participation in the associated outcome measures may be reduced. Based on this discussion and review of public comments, the Standing Committee remained

concerned regarding the evidence for this measure. The Standing Committee re-voted and did not pass the measure on evidence and therefore did not vote on overall suitability for endorsement.

During the CSAC meeting on December 9, 2022, the CSAC upheld the Standing Committee's recommendation and did not endorse the measure.

Measures Withdrawn From Consideration

Two measures previously endorsed by NQF were withdrawn during the endorsement evaluation process. Endorsement for these measures has been removed.

Table 2. Measures Withdrawn From Consideration

Measure	Reason for Withdrawal
NQF #1365e child and adolescent major depressive disorder (mdd): suicide risk assessment	Developer requested removal of endorsement.
NQF #0104e adult major depressive disorder (mdd): suicide risk assessment	Developer requested removal of endorsement because they do not have the resources to support NQF endorsement activities.

References

1 Substance Abuse and Mental Health Services Administration. *Key Substance Use and Mental Health Indicators in the United States: Results from the 2019 National Survey on Drug Use and Health*. Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2020. https://www.samhsa.gov/data/sites/default/files/reports/rpt29393/2019NSDUHFFRPDFWHTML/2019NSDUHFFR1PDFW090120.pdf.

2 Kuntz L. Psychiatric Care in the US: Are We Facing a Crisis? *Psychiatric Times*. April 2022. https://www.psychiatrictimes.com/view/psychiatric-care-in-the-us-are-we-facing-a-crisis. Last accessed July 2022.

3 Wu YY, Walkover M, Zhang W. Trends in psychological distress and COVID-19 incidence across 15 U.S. metropolitan statistical areas in 2020. *Soc Sci Med*. 2021;285:114285.

4 Tsai J, Elbogen EB, Huang M, et al. Psychological distress and alcohol use disorder during the COVID-19 era among middle- and low-income U.S. adults. *J Affect Disord*. 2021;288:41-49.

5 The impact of COVID-19 on mental health cannot be made light of. https://www.who.int/news-room/feature-stories/detail/the-impact-of-covid-19-on-mental-health-cannot-be-made-light-of. Last accessed July 2022.

6 Depression Rates by Country 2022. https://worldpopulationreview.com/country-rankings/depression-rates-by-country. Last accessed July 2022.

7 Stewart JC, Perkins AJ, Callahan CM. Effect of Collaborative Care for Depression on Risk of Cardiovascular Events: Data from the IMPACT Randomized Controlled Trial. *Psychosom Med*. 2014;76(1):29-37.

8 Dhar AK, Barton DA. Depression and the Link with Cardiovascular Disease. *Front Psychiatry*. 2016;7:33.

9 May HT, Bair TL, Reiss-Brennan B, et al. The association of antidepressant and statin use with death and incident cardiovascular disease varies by depression severity. *Psychol Health Med*. 2017;22(8):919-931.

10 Maurer DM, Raymond TJ, Davis BN. Depression: Screening and Diagnosis. *afp*. 2018;98(8):508-515.

11 Continuity of care vital for patients with serious mental illness. https://www.healio.com/news/psychiatry/20191106/continuity-of-care-vital-for-patients-with-serious-mental-illness. Last accessed July 2022.

12 *Withdrawal Management*. World Health Organization; 2009. https://www.ncbi.nlm.nih.gov/books/NBK310652/. Last accessed July 2022.

13 Foundation G. Benefits of Withdrawal Management for Addiction in Chicago, IL. Gateway Foundation. https://www.gatewayfoundation.org/locations/chicago-kedzie/withdrawal-management-benefits/. Last accessed July 2022.

Appendix A: Details of Measure Evaluation

Rating Scale: H=High; **M**=Moderate; **L**=Low; **I**=Insufficient; **NA**=Not Applicable

NQF ensures that quorum is maintained for all live voting. Quorum is 66 percent of active Standing Committee members minus any recused Standing Committee members. Due to the exclusion of recused Standing Committee members from the quorum calculation, the required quorum for live voting may vary among measures. There were no recusals on the Standing Committee pertaining to any of the measures under review. Quorum (16 out of 23 Standing Committee members) was met and maintained throughout the review of the evidence and performance gap for NQF #3312. Quorum was lost during the discussion of reliability of NQF #3312. Therefore, the Standing Committee discussed all remaining criteria for measures NQF #3312, NQF #3313, NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712 and voted after the meeting using an online voting tool. The Standing Committee received a recording of the meeting and a link to submit online votes. For the post-comment call on December 2, 2022, quorum was reached and maintained throughout the meeting. Voting closed after 48 hours with at least the number of votes required for quorum. Voting results are provided below.

A measure is recommended for endorsement by the Standing Committee when greater than 60 percent of voting members select a passing vote option (i.e., Pass, High and Moderate, or Yes) on all must-pass criteria and overall suitability for endorsement. A measure is not recommended for endorsement when less than 40 percent of voting members select a passing vote option on any must-pass criterion or overall suitability for endorsement.

Measures Endorsed

NQF #3312 Continuity of Care After Medically Managed Withdrawal From Alcohol and/or Drugs (Centers for Medicare & Medicaid Services [CMS]/Lewin Group)

Measure Worksheet | Specifications

Description: Percentage of discharges from a medically managed withdrawal episode for adult Medicaid Beneficiaries, age 18-64, that was followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder (pharmacotherapy) within 7 or 14 days after discharge). This measure is reported across all medically managed withdrawal settings.

Numerator Statement: Discharges in the denominator who have an inpatient, intensive outpatient, partial hospitalization, outpatient visit, residential, or drug prescription or procedure within 7 or 14 days after discharge from an inpatient hospital, residential addiction program, or ambulatory medically managed withdrawal.

Denominator Statement: Adult Medicaid beneficiary discharges from medically managed withdrawal from January 1 to December 15 of the measurement year.

Exclusions: N/A

Adjustment/Stratification: N/A

Level of Analysis: Population: Regional and State

Setting of Care: Inpatient/Hospital/Outpatient Services

Type of Measure: Process

Data Source: Claims

Measure Steward: Centers for Medicare & Medicaid Services (CMS)

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total Votes- 16; H-0; M-14; L-2; I-0**; 1b. Performance Gap: **Total votes- 16; H-6; M-10; L-0; I-0**

Rationale:

- The Standing Committee noted that this is a maintenance process measure at the population level of analysis that assesses the continuity of care after medically managed withdrawal from alcohol and/or drugs.
- The Standing Committee noted that the developer added additional supportive evidence that aligns with previous evidence, as well as evidence demonstrating that patients who are followed up with treatment after detox have lower odds of two-year mortality as well as the cost of SUDs on the U.S. economy and the return-on-investment for funding addiction treatment programs. The Standing Committee had no concerns and passed the measure on evidence.
- The Standing Committee agreed that that data showed a clear gap in performance, noting that the seven-day follow-up mean performance rate was 35.7 percent, with a standard deviation of 12.29 percent. The lowest state had follow-up visits with 19.96 percent of patients, and the highest rate was 52.78 percent. The 14-day follow-up mean performance rate was 41.01 percent with a standard deviation of 11.78 percent. The lowest state had follow-up visits with 24.91 percent of patients, and the highest rate was 57.11 percent.
- The Standing Committee also noted that younger patients, White non-Hispanic and Black non-Hispanic patients, Medicaid-only patients, and males were more likely to have a follow-up visit than their counterparts.
- The Standing Committee passed the measure on performance gap.

2. Scientific Acceptability of Measure Properties:

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Total votes-17; H-3; M-9; L-3; I-2**; 2b. Validity: **Total votes-17; H-1; M-10; L-5; I-1**

Rationale:

- The Scientific Methods Panel (SMP) did not review this measure.
- The Standing Committee noted that the measure specifications are mostly clear and precise; however, the previous Standing Committee expressed concerns about the measure reporting both seven-day and 14-day follow-up.
- With respect to the measure's specifications, the Standing Committee expressed concerns about whether the measure took telephonic or telemedicine follow-up into account, given the increase of these services during the COVID-19 pandemic. The developer clarified that the data provided are pre-pandemic data, and they plan to track Medicaid telemedicine claims as a feature of the measure in the future; however, they mentioned that telemedicine codes are not currently included in the measure. The developer also clarified that 2020 data would be available in 2023, and these data might influence whether telemedicine is added to the specifications.
- One Standing Committee member asked whether the data came from Medicare Advantage. In response, the developer clarified that the data came from Medicaid fee-for-service and Medicaid-Managed Care. The Standing Committee asked NQF whether there was a precedent that new information or updated data could trigger an early review of a measure. NQF staff clarified the situations in which an endorsed measure might be reviewed earlier than scheduled but emphasized that the Standing Committee must review the measure now as currently submitted.
- A Standing Committee member also asked whether a developer providing updated data would trigger an early review. NQF staff clarified that the developer could submit annual updates, and if there were changes to the specifications, the measure might be presented for early review.
- The Standing Committee asked for the developer to speak regarding their plan to obtain telehealth data sooner than 2023. The developer clarified that they plan to identify literature and stakeholder feedback prior to quantitative tests. The developer noted that data will only be available as the states make them available.
- A Standing Committee member asked whether telehealth visits were excluded because they did not have an associated code or because they were not included in the analysis. The developer indicated that telehealth visits had codes but were not included in the analysis. The developer noted that if a code had a modifier related to phone or video visits, it would not be excluded and would not have to wait the three years to be included in the measure due to annual specification updates.

- A Standing Committee member asked NQF for clarification on whether the measure could still be used if the Standing Committee did not pass the measure on reliability. NQF staff noted that while measure users prefer endorsed measures, measures can still be used in accountability or quality improvement programs without endorsement; therefore, it is difficult to determine what the impact on actual use would be if endorsement is lost.
- The developer noted that endorsement aids users in seeing what quality measures are available and losing endorsement may decrease uptake of the measure.
- The Standing Committee asked whether the developer would resubmit with renewed specifications if endorsement was lost, and the developer noted that this decision would be made by the measure's steward, CMS.
- At the Standing Committee's request, NQF clarified that there is no middle ground between endorsement and non-endorsement and that NQF would document any suggestions regarding the measure in the draft report, noting that the Standing Committee would like the measure to return for review as soon as possible.
- The Standing Committee passed this measure on reliability.
- The Standing Committee noted that the developer examined validity through examining correlations between: (1) NQF #3312 *Seven-Day Follow-Up* and NQF #3314 *14-Day Follow-Up* and (2) NQF #3312 *Healthcare Effectiveness Data and Information Set (HEDIS) Follow-Up (Seven and 30 Days) After Emergency Department Visit for Alcohol and Other Drug Abuse Dependence*. For the correlation with seven- and 30-day follow-up after emergency department visits for alcohol and other drug abuse dependence (seven/30 days), NQF #3312 and NQF #3314 both had a strong positive correlation ($r=0.98$, $p<0.001$), but NQF #3312 (seven-day and 14-day follow up) had a weakly positive correlation ($r=0.371$ and $r=0.257$, respectively). The developer stated that the results were not statistically significant, which was likely due to the small sample size ($n=6$).
- During the discussion of validity, the Standing Committee asked about patients who might be receiving a monthly medication and how that would affect 14-day rates versus seven-day rates. The developer replied that seven-day follow-up is considered a standard of care, but they allowed for some flexibility by including a 14-day follow-up as well; the developer also noted that encounters that occur on the day of discharge also count toward the measure; therefore, receiving a treatment at discharge that is renewed every 30 days would still be counted in the numerator.
- The Standing Committee asked for clarification about whether the developer examined correlations with readmissions, death, or relapse as an outcome, and the developer indicated that they did not.
- The Standing Committee did not have any further questions and passed the measure on validity.

3. Feasibility: Total votes-17; H-6; M-10; L-1; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that all the data elements are available in defined fields in electronic claims and are coded by someone other than the person obtaining the original information.
- The Standing Committee had no questions or concerns and passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use**: Total votes-17; Pass-15; No Pass-2**; 4b. Usability: **Total votes-17; H-6; M-6; L-4; I-1**

Rationale:

- The Standing Committee noted that this measure is not currently used in an accountability program but is available for voluntary use for certain states.
- The Standing Committee asked how many states currently use this measure, and the developer indicated that these data are not available via Medicaid since states are allowed to choose what they want to measure under Medicaid. The data used for testing came from nine states that are using the measure.

- The Standing Committee noted that the developer provided assistance with measure implementation, including steps to calculate performance, but that no feedback has been received on performance from the measure entities as of yet; it also noted that no unexpected findings or harms have been identified so far. The Standing Committee passed the measure on use and usability.

5. Related and Competing Measures

- This measure is related to the following measures:
    - NQF #0004 Initiation and Engagement of Alcohol and Other Drug Abuse or Dependence Treatment
    - NQF #2605 Follow-Up After Emergency Department Visit for Mental Illness and Other Drug Abuse or Dependence
    - NQF #3453 Continuity of Care After Inpatient or Residential Treatment for Substance Abuse Disorder (SUD)
- During the post-comment call, the Standing Committee discussed the differences and similarities between the related measures but did not make any explicit recommendations for harmonization for NQF #3312.

6. Standing Committee Recommendation for Endorsement: **Total votes-17; Yes-11; No-6**

7. Public and Member Comment

- No pre-evaluation public comments were received.
- Two post-evaluation public comments were received, both of which did not express support for this measure.
    - One of the comments raised concerns about the age range and payer population proposed within the measures. The commenter noted that the measure is currently limited to Medicaid patients within specific age ranges and encouraged the developer to expand the measures to a larger patient population and payer mix. The commenter also raised the concern that the claims-based data may not be as accurate for these measures due to a relatively high volume of dual-eligible patients, and thus, the outcomes of these measures may not reveal a complete view of the patients who receive, or should receive, this best-practice standard of care.
        - The developer responded to the comment, noting that for NQF #3312, the upper limit of 64 years was chosen based on evidence from the literature, input from experts, feasibility of data collection, and findings from measure testing. Regarding the commenter's concerns about the accuracy of claims-based data for individuals eligible for dually enrolled participants, the developer clarified that CMS reviewed data sources to ensure the most accurate and complete data were used for measure calculation and testing. The primary data used for testing in the NQF #3312 and NQF #3313 submissions were Transformed Medicaid Statistical Information System (T-MSIS) Analytical Files (TAF); for participants dually enrolled in both Medicare and Medicaid, Medicare Parts A, B, C, and D claims data were also used.
    - The second comment expressed concern about the exclusion of telemedicine codes in the current set of measure specifications, given the dramatic increase in telemedicine services during the COVID-19 pandemic.
        - The developer provided a response, indicating that the measure was specified and tested using data obtained prior to the COVID-19 pandemic, and that in the future, when more recent Medicaid administrative claims data are available that encompass this time frame, the measure's technical specifications will be reconsidered.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-15; No-0 (December 9, 2022: Endorsed)

- The CSAC upheld the Standing Committee's decision to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

NQF #0710e Depression Remission at 12 Months (MN Community Measurement)

Measure Worksheet | Specifications

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission twelve months (+/- 60 days) after an index visit.

Numerator Statement: The number of patients in the denominator who reached remission, with a PHQ-9 or PHQ-9M result less than five, twelve months (+/- 60 days) after an index visit.

Denominator Statement: Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

Exclusions: Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

Adjustment/Stratification: Statistical Risk Model

Level of Analysis: Clinician: Group/Practice

Setting of Care: Outpatient Services

Type of Measure: Outcome: PRO-PM

Data Source: Electronic Health Records

Measure Steward: MN Community Measurement

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total Votes-17; Pass-17; No Pass-0**; 1b. Performance Gap: **Total votes-17; H-6; M-9; L-1; I-0;**

Rationale:

- The Standing Committee noted that the developer cited updated evidence from the Institute for Clinical Systems Improvement's (ICSI) Health Care Guideline titled *Depression in Primary Care* and the Veterans Affairs (VA) Department of Defense (DOD) Clinical Practice Guidelines for Depression for the treatment algorithm related to major depressive disorder (MDD) and persistent depressive disorder.
- The Standing Committee asked for clarification regarding the developer's choice to use the PHQ-9 instead of other tools, such as the Ask Suicide-Screening Questions (ASQ) screening tool. The developer clarified that they did examine 21 other tools with comparable cut points for when the patient was in remission to the PHQ-9 and noted that the suicidality is not in scope for this measure.
- The Standing Committee passed the measure on evidence.
- The Standing Committee noted that because the measure was respecified in 2020, trend data are available across the lifetime of the measure; however, the developer presented an analysis of the impact of COVID-19 on measure outcomes, which utilizes trend data. In the state of Minnesota, adult remission at 12 months fell from 10.3% in 2019 to 10.2% in 2020, and adolescent remission at 12 months fell from 7.8% in 2019 to 7% in 2020. The Standing Committee agreed that there is a substantial gap to warrant this measure and passed the measure on performance gap.

2. Scientific Acceptability of Measure Properties:

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Total votes- 17; H-0; M-15; L-1; I-1**; 2b. Validity: **Total votes-17; H-1; M-9; L-5; I-2;** Post-Comment Validity Revote**: Total Votes-15; H-1, M-12, L-2, I-0**

Rationale:

- The Standing Committee noted that the specifications were updated to include the PHQ-9M.
- One Standing Committee member asked for clarification on knowing which PHQ-9 survey to choose. The developer clarified that the most recent PHQ-9 is chosen within a four-month period.
- The Standing Committee had no concerns with the specifications or the reliability testing and passed the measure on reliability.
- For validity, a Standing Committee member asked for more information about the data point of 22 percent of individuals showing remission. The developer clarified that the data only included patients with

whom a follow-up visit was conducted, noting that some remission cases may be lost due to lack of follow-up.

- The Standing Committee noted that for missing data, the variability of rates among medical groups around the statewide average for adults were 17.0% (range of 0% to 37.2%) and 14.5% (range of 0% to 29.1%) for adolescents.
- The Standing Committee also noted that the 12 month follow-up has the widest variation among medical groups, and those overall rates are low.
- One Standing Committee member noted that it is difficult to interpret the validity data, given the high rate of missing data in the denominator that are artificially lowering performance.
- Other Standing Committee members disagreed, noting that the goal of the measure is to improve care and that systems can be developed to aid in better follow-up.
- The developer noted that one goal of these measures was to address this known gap in care related to patients with depression who are lost to follow-up, estimated to be as high as 80 percent.
- Due to the concerns noted, the Standing Committee did not reach consensus on validity.
- During the public comment period, the developer responded to the Standing Committee's concerns:
    - Regarding data element validity, the developer clarified that the data elements for these measures are contained in structured fields extracted directly from the EHR and are not abstracted and confirmed that agreement with the source is strong.
    - Regarding missing data, the developer clarified that the measure construct purposely includes patients in the denominator who do not have a follow–up in order to avoid bias in the measure and noted that the lack of a follow-up assessment is not missing data, as it represents a gap in care.
    - Regarding telehealth services, the developer clarified that telehealth services are included in the use and specifications of the measures and have been included as part of the denominator definition for several years.
- The Standing Committee discussed the developer's responses and was satisfied with the clarification of the concerns raised during the measure evaluation meeting.
- The Standing Committee passed the measure on validity.

3. Feasibility: Total votes-17; H-2; M-11; L-3; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that the measure is captured in EHRs, and groups can successfully extract the information from their EHR.
- The Standing Committee also noted that Minnesota Community Measurement (MNCM) developed a direct data submission process whereby medical groups submit a patient level data file for rate calculation and public reporting. MNCM is implementing a new data collection method that serves as a warehouse of clinical data in which measures are calculated centrally. No fees are associated with this program.
- A Standing Committee member asked for clarification regarding whether clinics get reimbursed for sending the measure data and what organization is paying for establishing and managing the data repository. The developer clarified that the PHQ-9 data are extracted from the Epic EHR and that there is no cost for participation. In addition, several health plans include these measures in their pay-for-performance contracts; however, MNCM does not reward providers for participation.
- The Standing Committee asked for clarification on maintenance and staffing of the MNCM registry, and the developer clarified that MNCM was funded by health plan and medical group member dues, state government contracts, grant funding, and other various sources of funding.
- When asked whether this measure would be reported on a national level, the developer clarified that they are not aware of other states collecting this measure statewide, but it is in CMS' pay-for-performance programming. The developer suggested that other states may be collecting this measure as part of value-based payment contracts. A Standing Committee member raised the following concern: This measure may be difficult to collect at the national level, given that not everyone has the infrastructure in place; however, when evaluating the current measure and the data presented, the Standing Committee had no serious concerns and passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use: Total votes-17; Pass-15; No Pass-2; 4b. Usability: Total votes-17; H-0; M-14; L-3; I-0

Rationale:

- The Standing Committee noted that the measure is publicly reported and included in two CMS Quality Payment Programs: MIPS and eCQM.
- The measure is publicly reported on the MN HealthScores website and as a part of the MNCM Annual Health Care Quality Report and the Annual Disparities by Insurance Type and Disparities by Race, Ethnicity, Language, and Country of Origin. It is also the focus of several issue briefs.
- The measure is also used in all primary care clinics in MN and bordering communities in Wisconsin, North Dakota, South Dakota, and Iowa.
- The Standing Committee passed the measure on use.
- The Standing Committee noted that prior to changes in the specifications, follow-up at six months improved from 17.0% in 2010 to 41.8% in 2019 for adults. Adolescents have a 2019 follow-up rate of 38.9%.
- Trend data for usability were not available across the lifetime of the measure due to recent changes in the specifications; therefore, the Standing Committee asked NQF to clarify whether trend data are a requirement. NQF replied that there are times when trend data may not be available for a measure and that the Standing Committee would need to deliberate on whether the rationale provided by the developer is acceptable.
- The Standing Committee passed the measure on usability.

5. Related and Competing Measures

- This measure is related to the following measures:
    - NQF #1884 Depression Response at Six Months – Progress Towards Remission
    - NQF #1885 Depression Response at 12 Months – Progress Towards Remission
    - NQF #0711 Depression Remission at Six Months
    - NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M
- During the post-comment call, a discussion of harmonization took place regarding the depression suite of the measures that the Standing Committee recommended for endorsement: NQF #1884, #710e, and #711. The Standing Committee agreed that the measures were harmonized and suggested that a composite of the suite of the measures may be stronger and could strengthen the measures that were not recommended for endorsement (NQF #0712 and #1885).

6. Standing Committee Recommendation for Endorsement: Total Votes-15; Yes-14; No-1

- During the post-comment meeting, the Standing Committee discussed the measure developer's responses and public comments, voted to pass this measure on validity, and subsequently voted to recommend it for endorsement.

7. Public and Member Comment

- Two pre-evaluation public comments were received for this measure:
    - One comment was submitted by the measure developer to clarify validity testing.
    - One comment was submitted by a member of the public: It raised concerns about the evidence to support the measure, validity testing, and feasibility.
- Six post-evaluation public comments were received for this measure:
    - Two comments were submitted by the measure developer.
        - In these comments, which apply to NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712 as well, the developer clarified measure details and addressed the Standing Committee's concerns about evidence, data element validity testing, missing data, and telemedicine that led to the "consensus not reached" status.
        - The Standing Committee discussed the developer's responses was satisfied with the clarification of the concerns raised during the measure evaluation meeting.
    - Two comments did not express support for the measure.

- One public comment, which applies to the depression measures (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712), was submitted, raising concerns that there was insufficient evidence demonstrating that depression scores can be successfully reduced by at least 50 percent across the defined patient population within a 12-month time frame, nor was any evidence provided supporting this requirement of 50 percent. The commenter asked for clarification on whether this measure has met all of the requirements for electronic clinical quality measures (eCQMs) since the complimentary measure, #0710e, cites the same data but uses different data sources and specifications. The comment also echoes the Standing Committee's concerns over the omission of telehealth services and inclusion of patients lost to follow-up in this measure.
    - The developer replied, noting that response measures are considered as an interim outcome and aim to show progress towards the ultimate outcome of the remission of symptoms. They stated that it is a reasonable expectation to have symptoms reduced at 12 months with the 60-day window, extending the assessment time frame out to 14 months. The developer noted that the acute treatment phase of depression is six to 12 weeks, so an assessment of symptoms during the measure's time frame is well into the continuation phase of treatment and an interim goal of symptom reduction by 50 percent or greater during this time frame is not unreasonable. The developer also cited two studies, which used a response of greater than or equal to 50 percent decrease as outcomes. To clarify the commenter's concern about eCQM development, the developer noted that this measure is not an eCQM, and instead is a digital quality measure for which all components are captured from discrete data elements in the electronic record. The developer stated that it has been capturing this information in a digital format via EHR extraction for over 10 years. Regarding the commenter's concerns about telehealth services, the developer clarified the telehealth is included in the measure specifications.
- Another public comment raised concerns about the lack of telehealth inclusion.
    - In response to the commenter's concerns about telehealth services, the developer reiterated the explanation from the other non-supportive comment.

- One comment expressed support for the measure.
    - The comment attributes the success of developing the clinic's care coordination system to the implementation of these measures and encouraged the Standing Committee to endorse the entire suite of measures.
- One comment raised concerns with NQF's process.
    - The commenter raised concerns that re-endorsement of the depression measure set is also endorsement of the PHQ-9 as NQF's tool of choice in measuring outcomes for depressed patients and noted that there were other tools used by providers at different facilities.
        - NQF provided a response thanking the commenter for their comment. NQF noted that NQF's endorsement of a measure may appear to endorse a certain tool due to its use in recommended measures but assured the commenter that NQF does not endorse measurement tools. Any measure developer may submit quality measures for the Standing Committee to review against the measure evaluation criteria, and no preference is given to measures using specific tools.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-15; No-0 (December 9, 2022: Endorsed)

- The CSAC upheld the Standing Committee's decision to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

NQF #0711 Depression Remission at Six Months (MN Community Measurement)

Measure Worksheet | Specifications

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission six months (+/-60 days) after an index visit.

Numerator Statement: The number of patients in the denominator who reached remission, with a PHQ-9 or PHQ-9M result less than five, six months (+/-60 days) after an index visit.

Denominator Statement: Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

Exclusions: Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

Adjustment/Stratification: Statistical Risk Model

Level of Analysis: Clinician: Group/Practice

Setting of Care: Outpatient Services

Type of Measure: Outcome: PRO-PM

Data Source: Electronic Health Records

Measure Steward: MN Community Measurement

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total votes-17; Pass-17; No Pass-0** 1b. Performance Gap: **Total votes-17; H-3; M-14; L-0; I-0;**

Rationale:

- The Standing Committee noted that the evidence for this measure was largely the same as the evidence for NQF #0710e.
- A Standing Committee member noted that an advantage of remission being examined at six months versus 12 is the ability for clinicians to have more control over the outcomes.
- The Standing Committee had no concerns and passed the measure on evidence.
- The Standing Committee noted that because the measure was respecified in 2020, trend data are not available across the lifetime of the measure; however, the developer presented an analysis of the impact of COVID-19 on measure outcomes, which utilizes trend data. In the state of Minnesota, adult remission at six months fell from 11.3% in 2019 to 11% in 2020, and adolescent remission at six months grew from 8% in 2019 to 8.5% in 2020. The Standing Committee agreed that there was a gap to warrant this measure's existence and passed the measure on performance gap.
- The Standing Committee agreed that a sufficient gap exists and passed the measure on performance gap.

2. Scientific Acceptability of Measure Properties:

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Total votes-17; H-0; M-15; L-1; I-1**; 2b. Validity: **Total Votes-17; H-0, M-10, L-4, I-3;** Post-Comment Validity Revote**: Total Votes-15; H-2, M-10, L-3, I-0**

Rationale:

- The Standing Committee noted that reliability testing was conducted at the accountable-entity level, and that for adults, the signal-to-noise (SNR) ratio was 0.940. For adolescents, the score was 0.898.
- The Standing Committee had no concerns and passed the measure on reliability.
- The Standing Committee noted that validity testing was conducted at the accountable-entity level and correlated with depression response at six months for the following results: R-squared value = 0.8922 (for adults) and R-squared value = 0.8386 (for adolescents).
- For validity, the Standing Committee highlighted similar concerns with the previous measure (NQF #0710e), namely how missing data were counted within the measure, considering those patients who are lost to follow-up remained in the denominator.

- Due to the concerns listed above, the Standing Committee did not reach consensus on validity.
- During the post-comment meeting, the Standing Committee noted that this measure had the same concerns as NQF #0710e and agreed that the developer's responses alleviated those concerns.
- The Standing Committee passed the measure on validity.

3. Feasibility: Total votes-17; H-1; M-12; L-3; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that the measure is currently captured in an EHR and that groups can successfully extract the information from their EHR.
- The Standing Committee raised concern that the measure seems more feasible for systems that have a registry but noted that the measure has no fees associated with it.
- The Standing Committee reiterated the earlier concern that the measure only allows the use of one depression screener when there are other tools that may also be acceptable; however, it accepted the developer's previous statements, clarifying the reasons why the PHQ-9 and PHQ-9M were chosen and passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use: Total votes-17; Pass-16; No Pass-1; 4b. Usability: Total votes-17; H-0; M-15; L-2; I-0

Rationale:

- The Standing Committee noted that this measure is used in Minnesota on the MN HealthScores website.
- The Standing Committee had no concerns with the measure's use and passed the measure on this criterion.
- The Standing Committee agreed that the benefits for patients being treated for depression outweighed any possible unintended consequences and passed the measure on usability.

5. Related and Competing Measures

- This measure is related to the following measures:
    - NQF #1884 Depression Response at Six Months – Progress Towards Remission
    - NQF #1885 Depression Response at 12 Months – Progress Towards Remission
    - NQF #0710e Depression Remission at 12 Months
    - NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M

- During the post-comment call, a discussion of harmonization took place regarding the depression suite of the measures that the Standing Committee recommended for endorsement: NQF #1884, #710e, and #711. The Standing Committee agreed that the measures were harmonized and suggested that a composite of the suite of the measures may be stronger and could strengthen the measures that were not recommended for endorsement (NQF #0712 and #1885).

6. Standing Committee Recommendation for Endorsement: Total Votes-15; Yes-14; No-1

Rationale:

- During the post-comment meeting, the Standing Committee discussed the measure developer's responses and public comments, voted to pass this measure on validity, and subsequently voted to recommend it for endorsement.

7. Public and Member Comment

- Two public comments were submitted:
    - One comment was submitted by the measure developer to clarify validity testing.
    - One comment was submitted by a member of the public: It raised concerns about the evidence to support the measure, validity testing, and feasibility.
- Six post-evaluation public comments were received for this measure as noted in detail in the discussion of NQF #0710e:

- Two comments were submitted by the measure developer.
  - In these comments, which apply to NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712 as well, the developer clarified measure details and addressed the Standing Committee's concerns about evidence, data element validity testing, missing data, and telemedicine that led to the "consensus not reached" status.
  - The Standing Committee discussed the developer's responses was satisfied with the clarification of the concerns raised during the measure evaluation meeting.
- Two comments did not express support for the measure.
  - One public comment, which applies to the depression measures (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712), was submitted raising concerns with evidence, data sources, and the omission of telehealth as noted in detail above.
    - The developer replied to each of the commenter's points.
  - Another public comment raised concerns about the lack of telehealth inclusion.
    - In response to the commenter's concerns about telehealth services, the developer reiterated the explanation from the other non-supportive comment.
- One comment expressed support for the measure.
  - The comment attributes the success of developing the clinic's care coordination system to the implementation of these measures and encouraged the Standing Committee to endorse the entire suite of measures.
- One comment raised concerns with NQF's process.
  - The commenter raised concerns regarding NQF's endorsement of a patient-reported outcome (PRO) tool.
    - NQF provided a response, clarifying that NQF does not endorse measurement tools.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-15; No-0 (December 9, 2022: Endorsed)

- The CSAC upheld the Standing Committee's decision to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

NQF #1884 Depression Response at Six Months – Progress Towards Remission (MN Community Measurement)

Measure Worksheet | Specifications

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who are progressing towards remission by achieving a response (PHQ-9 or PHQ-9M score reduced by 50% or greater) six months (+/- 60 days) after an index visit.

Numerator Statement: The number of patients in the denominator who achieved a response as demonstrated by a PHQ-9 or PHQ-9M score reduced by 50% or greater six months (+/- 60 days) after an index visit.

Denominator Statement: Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

Exclusions: Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

Adjustment/Stratification: Statistical Risk Model

Level of Analysis: Clinician: Group/Practice

Setting of Care: Outpatient Services

Type of Measure: Outcome: PRO-PM

Data Source: Electronic Health Records

Measure Steward: MN Community Measurement

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total votes-17; Pass-14; No Pass-3**; 1b. Performance Gap: **Total votes-17; H-2; M-12; L-2; I-1;**

Rationale:

- The Standing Committee noted that the developer provided updated evidence for this measure, which included updated evidence from the ICSI's Health Care Guideline titled *Depression in Primary Care* and the VA DOD Clinical Practice Guidelines for Depression for the treatment algorithm related to major depressive disorder (MDD) and persistent depressive disorder.
- The Standing Committee raised a concern that two points in time are needed to calculate the measure and asked the developer whether there were data showing the effect of follow-up frequency at six months versus 12 months. The developer clarified that individuals with four to 12 PHQ-9 assessments during the assessment period were three times more likely to achieve response or remission at 12 months compared to patients with only one to three PHQ-9 assessments.
- One Standing Committee member inquired about the evidence regarding targeting a score decrease of five points versus a 50% reduction in score. The developer emphasized strong evidence related to the cut points of the PHQ-9 tool, showing zero to four as remission with mild to no depression symptoms; a five-point drop in score is considered clinically significant.
- The Standing Committee member followed up by asking why the developer chose a 50% reduction instead of a five-point reduction. The developer stated that for higher initial PHQ-9 scores, while five points may be clinically significant, 50% reduction indicates more meaningful progress toward remission.
- The Standing Committee did not raise any additional concerns and passed the measure on evidence.
- The Standing Committee agreed that the measure performance data showed a sufficient gap: In total, 19.4 percent (range of 0–37.2%) of patients with an index event had a depression response (>=50% from baseline) at six months. The Standing Committee also noted that the developer included stratification information in the submission that was not noted in earlier discussions.
- The Standing Committee passed the measure on performance gap.

2. Scientific Acceptability of Measure Properties:

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Total votes-17; H-1; M-11; L-4; I-1**; 2b. Validity: **Total Votes-17; H-0; M-8; L-5; I-4**); Post-Comment Validity Revote**: Total Votes-14; H-1; M-11; L-2; I-0**

Rationale:

- The Standing Committee noted that the developer made several changes to the measure specifications, including incorporating adolescents ages 12 to 17, adding the PHQ-9M (modified for teens) PRO tool, expanding the assessment window to +/- 60 days, modifying an exclusion value set for personality disorder, adding exclusions for schizophrenia and pervasive developmental disorder, and removing the requirement of depression diagnosis being in the primary position for a behavioral specialty.
- The Standing Committee noted that the developer conducted reliability testing at the accountable-entity level and found an SNR of 0.92 for adults and 0.83 for adolescents.
- The Standing Committee had no concerns regarding the specifications or reliability testing and passed the measure on reliability.
- The Standing Committee noted that the developer provided validity testing at both the patient/encounter and accountable-entity levels. Patient/encounter validity testing was conducted through data quality checks and audits, and 94% of those audited passed the audit. At the accountable-entity level, correlation between depression response at six months and depression remission at six months had an R-squared of 0.8922 for adults and 0.8386 for adolescents.
- For validity, the Standing Committee raised similar concerns with the previous measure (NQF #0710e and NQF #0711), namely how missing data were counted within the measure, considering those patients who are lost to follow-up remained in the denominator.
- Due to the concerns noted, the Standing Committee did not reach consensus on validity.
- During the post-comment meeting, the Standing Committee noted that this measure had the same concerns as NQF #0710e and agreed that the developer's responses alleviated those concerns.

- The Standing Committee passed the measure on validity.

3. Feasibility: Total votes-17; H-0; M-14; L-2; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that the measure is currently captured in an EHR in which groups can successfully extract the information from their EHR.
- The Standing Committee raised the following concern: The measure seems more feasible for systems that have a registry, and it also might be challenging for providers to report on a six-month time frame.
- Other Standing Committee members noted that the six-month time frame was an advantage toward detecting treatment resistance earlier.
- The Standing Committee reiterated the earlier concern that the measure only allows the use of one depression screener when there are other tools that may also be acceptable; despite this concern, the Standing Committee accepted the developer's previous statements clarifying the reasons why the PHQ-9 and PHQ-9M were chosen and passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use: Total votes-17; Pass-13; No Pass-4; 4b. Usability: Total votes-17; H-0; M-12; L-5; I-0

Rationale:

- The Standing Committee noted that the measure is currently used in Minnesota as part of MN HealthScores.
- The Standing Committee raised the following concern: On a national scale, states may choose to use one measure in this suite of measures rather than all of them, which may pose some challenges.
- The Standing Committee passed the measure on use.
- The Standing Committee noted that the developer could not provide trend data over the lifetime of the measure.
- A Standing Committee member raised a concern: Psychosocially complex patients may take longer to show improvement with treatment than six months, and as a result, they may not be captured in this measure. The Standing Committee determined that the measure was usable and that its benefits outweighed any potential harm and passed the measure on usability.

5. Related and Competing Measures

- This measure is related to the following measures:
    - NQF #1885 Depression Response at 12 Months – Progress Towards Remission
    - NQF #0711 Depression Remission at Six Months
    - NQF #0710e Depression Remission at 12 Months
    - NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M

- During the post-comment call, a discussion of harmonization took place regarding the depression suite of the measures that the Standing Committee recommended for endorsement: NQF #1884, #710e, and #711. The Standing Committee agreed that the measures were harmonized and suggested that a composite of the suite of the measures may be stronger and could strengthen the measures that were not recommended for endorsement (NQF #0712 and #1885).

6. Standing Committee Recommendation for Endorsement: Total Votes-14; Yes-13; No-1

- During the post-comment meeting, the Standing Committee discussed the measure developer's responses, voted to pass this measure on validity, and subsequently voted to recommend it for endorsement.

7. Public and Member Comment

- Two public comments were submitted:
    - One comment was submitted by the measure developer to clarify validity testing.

    - One comment was submitted by a member of the public: It raised concerns about the evidence to support the measure, validity testing, and feasibility.
- Six post-evaluation public comments were received for this measure as noted in detail in the discussion of NQF #0710e:
    - Two comments were submitted by the measure developer.
        - In these comments, which apply to NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712 as well, the developer clarified measure details and addressed the Standing Committee's concerns about evidence, data element validity testing, missing data, and telemedicine that led to the "consensus not reached" status.
        - The Standing Committee discussed the developer's responses was satisfied with the clarification of the concerns raised during the measure evaluation meeting.
    - Two comments did not express support for the measure.
        - One public comment, which applies to the depression measures (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712), was submitted raising concerns with evidence, data sources, and the omission of telehealth as noted in detail above.
            - The developer replied to each of the commenter's points.
        - Another public comment raised concerns about the lack of telehealth inclusion.
            - In response to the commenter's concerns about telehealth services, the developer reiterated the explanation from the other non-supportive comment.
    - One comment expressed support for the measure.
        - The comment attributes the success of developing the clinic's care coordination system to the implementation of these measures and encouraged the Standing Committee to endorse the entire suite of measures.
    - One comment raised concerns with NQF's process.
        - The commenter raised concerns regarding NQF's endorsement of a PRO tool.
            - NQF provided a response clarifying that NQF does not endorse measurement tools.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-15; No-0 (December 9, 2022: Endorsed)

- The CSAC upheld the Standing Committee's decision to recommend the measure for endorsement.

9. Appeals

- No appeals were received.

Measures Not Endorsed

NQF #3313 Follow-Up Care for Adult Medicaid Beneficiaries Who Are Newly Prescribed an Antipsychotic Medication (CMS/Lewin Group)

Measure Worksheet | Specifications

Description: Percentage of new antipsychotic prescriptions for Medicaid beneficiaries aged 18 years and older who have completed a follow-up visit with a provider with prescribing authority within four weeks (28 days) of prescription of an antipsychotic medication.

Numerator Statement: The percentage of Medicaid beneficiaries aged 18 years and older with new antipsychotic prescriptions who completed an outpatient follow-up visit with a provider with prescribing authority within 28 days of the new antipsychotic prescription fill.

Denominator Statement: New antipsychotic prescriptions for Medicaid beneficiaries aged 18 years and older.

Exclusions: Medicaid beneficiaries with an acute inpatient admission during the 28-day follow-up period after prescription of an antipsychotic medication.

Adjustment/Stratification: N/A

Level of Analysis: Population: Regional and State

Setting of Care: Outpatient Services

Type of Measure: Process

Data Source: Claims
Measure Steward: CMS

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total votes-17; H-3; M-13; L-0; I-1**; 1b. Performance Gap: **Total votes-17; H-3; M-13; L-1; I-0;**

Rationale:

- The Standing Committee noted that this is a maintenance process measure at the population level of analysis that assesses follow-up care for adult Medicaid beneficiaries who are newly prescribed an antipsychotic medication.
- The Standing Committee noted that the developer added additional evidence supporting the measure, including mention of an updated guideline from 2020 that instructs the following: "Patients may take between 2 and 4 weeks to show an initial response and longer periods of time to show full or optimal response. Monitoring of the patient's clinical status for 2-4 weeks is warranted on a therapeutic dose unless the patient is having uncomfortable side effects." The Standing Committee had no concerns and passed the measure on evidence.
- The Standing Committee agreed that the data showed a clear gap in performance with a mean performance rate of 46.72%, while the lowest rate was 35.86% and the highest rate was 58.72%. The Standing Committee also noted that older patients, White non-Hispanic and Hispanic patients, dual-eligible patients, and females were all more likely to have a follow-up than their counterparts.
- The Standing Committee passed the measure on performance gap.

2. Scientific Acceptability of Measure Properties:

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Total votes- 17; H-2; M-10; L-4; I-1**; 2b. Validity: **Total Votes-15; H-0, M-9, L-4, I-2;** Post-Comment Validity Revote**: Total Votes-15; H-0, M-1, L-12, I-2**

Rationale:

- Regarding specifications, a Standing Committee member asked about the lookback time frame that determines whether this measure captures a new or a reinitiated prescription. The developer clarified that after 120 days, or four months, the patient would be considered a new user again.
- The Standing Committee asked why the measure did not apply to populations outside of Medicaid, and the developer noted that Medicaid is the scope for the contract; however, they are open to partnering with other developers who serve commercial and Medicare clients if they have interest.
- The Standing Committee raised a concern that the measure looks at whether a follow-up visit occurred at all rather than a follow-up visit specific to the antipsychotic prescription the patient received. For instance, if an individual receives an antipsychotic prescription from a psychiatric visit and 20 days later sees a primary care doctor, would the primary care visit count toward the measure since a primary care provider has prescribing authority? The developer confirmed that it would and acknowledged that a limitation of claims data is that it does not distinguish whether the follow-up visit covered antipsychotic use.
- The Standing Committee asked how telehealth was handled in this measure, and the developer noted that as with NQF #3312, telehealth codes are not accounted for in the value sets. The Standing Committee recommended that these be considered in the future and decided to pass the measure on reliability.
- The Standing Committee passed the measure on reliability.
- The Standing Committee noted that validity was examined through correlations between NQF #3313 and two HEDIS measures: (1) *Follow-Up After Hospitalization (Seven and 30 Days) for Mental Illness* and (2) *Follow-Up Care for Children Prescribed ADHD Medication*. NQF #3313 had a moderately positive correlation with both HEDIS measures: r=0.42 and r=0.38 for seven and 30 days, respectively, for the first measure, and r=0.58 and r=0.53 for the initiation phase and continuation and maintenance phase, respectively. The results were not statistically significant, which was likely due to the small sample size (n=9).

- The Standing Committee discussed whether a follow-up visit within 28 days can actually address potential physical health issues, such as metabolic syndrome, or whether a longer period of time is more likely to effectively show these effects. There was an additional concern that community-based workers who do not bill under a provider's NPI number would not be captured by the measure. The developer replied that they would discuss these considerations with their technical expert advisory panel.
- Due to the concerns noted, the Standing Committee did not reach consensus on validity.
- During the post-comment meeting, the Standing Committee reiterated validity concerns about aspects of the measure that may result in inaccurate results, including telemedicine visits not being included in the measure, sufficient follow-up time, and whether providers who conduct follow-up visits would be captured, and did not pass the measure on validity.

3. Feasibility: Total votes-17; H-4; M-12; L-1; I-0

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that all the data elements are available in Medicaid claims (Transformed Medicaid Statistical Information System [TMSIS]) and Medicare Part A, B, C and D administrative claims.
- The Standing Committee had no questions or concerns and passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use: Total votes- 17; Pass-15; No Pass-2; 4b. Usability: Total votes-17; H-2; M-11; L-3; I-1

Rationale:

- The Standing Committee noted that this measure has been approved for states to use to assess the continuum of mental healthcare for antipsychotic use as part of the Section 1115 Demonstration for Serious Mental Illness (SMI) and Serious Emotional Disturbance (SED) launched by CMS.
- The Standing Committee also noted that seven states will likely report NQF #3313 in the near future: District of Columbia, Idaho, Indiana, Oklahoma, Utah, Vermont, and Washington.
- The Standing Committee asked how states use the measure since this information was not available in the submission, and the developer clarified that this measure is flagged for use in the CMS section 1115 Medicaid waiver.
- The Standing Committee voted to pass the measure on use.
- For usability, the Standing Committee again noted the need to consider capturing certain types of follow-up visits, such as community health workers and registered nurses, in order to report many types of progress; however, it decided that the current measure is still usable and voted to pass the measure on usability.

5. Related and Competing Measures

- This measure is related to the following measures:
    - NQF #0108 Follow-Up Care for Children Prescribed ADHD Medication (ADD)
    - NQF #3539e Use of Antipsychotics in Older Adults in the Inpatient Hospital Screening
- The Standing Committee did not recommend this measure for endorsement and therefore did not discuss related and competing measures during the post-comment call.

7. Public and Member Comment

- No pre-evaluation public comments were received.
- One post-evaluation public comment was received.
    - This comment did not express support for this measure.
        - This non-supportive comment raised concerns about the age range and payer population proposed within the measures. The commenter noted that the measure is currently limited to Medicaid patients within specific age ranges and encouraged the developer to expand the measures to a larger patient population and payer mix. The commenter also raised the concern that the claims-based data may not be as accurate

for these measures due to a relatively high volume of dual-eligible patients, and thus, the outcomes of these measures may not reveal a complete view of the patients who receive, or should receive, this best-practice standard of care.

- The developer provided a comment in response to the comment, noting that for NQF #3313, the current specifications currently do not set an upper limit for the age of individuals eligible for inclusion in the measure; only pediatric cases are excluded from assessment. Regarding the commenter's concerns about the accuracy of claims-based data for individuals eligible for dually enrolled participants, the developer clarified that CMS reviewed data sources to ensure the most accurate and complete data were used for measure calculation and testing. The primary data used for testing in the NQF #3312 and NQF #3313 submissions were Transformed Medicaid Statistical Information System (T-MSIS) Analytical Files (TAF); for participants dually enrolled in both Medicare and Medicaid, Medicare Parts A, B, C, and D claims data were also used.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-14; No-1 (December 9, 2022: Not Endorsed)

- The CSAC upheld the Standing Committee's decision to not recommend the measure for endorsement.

NQF #1885 Depression Response at 12 Months – Progress Towards Remission (MN Community Measurement)

Measure Worksheet | Specifications

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who are progressing towards remission by achieving a response (PHQ-9 or PHQ-9M score reduced by 50% or greater) twelve months (+/- 60 days) after an index visit.

Numerator Statement: The number of patients in the denominator who achieved a response as demonstrated by a PHQ-9 or PHQ-9M score reduced by 50% or greater twelve months (+/- 60 days) after an index visit.

Denominator Statement: Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

Exclusions: Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

Adjustment/Stratification: Statistical Risk Model

Level of Analysis: Clinician: Group/Practice

Setting of Care: Outpatient Services

Type of Measure: Outcome: PRO-PM

Data Source: Electronic Health Records

Measure Steward: MN Community Measurement

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total votes-17; Pass-14; No Pass-3**; 1b. Performance Gap: **Total votes-17; H-3; M-11; L-2; I-1;**

Rationale:

- The Standing Committee noted that the developer added updated evidence for this measure, including a guideline from the ICSI concluding that clinicians should establish and maintain follow-up with patients, which is highly correlated with improved response and remission scores.
- The Standing Committee agreed that the empirical evidence provided was strong and very similar to the evidence was provided for the previous measure and passed the measure on evidence.
- The Standing Committee noted rates of follow-up among adults at 12 months decreased from 41.8% in 2019 to 39.6% in 2020, likely influenced by COVID-19. For adolescents, rates of follow-up at 12 months

decreased from 38.9% in 2019 to 35.6% in 2020, also likely influenced by COVID-19. The Standing Committee passed the measure on performance gap

2. Scientific Acceptability of Measure Properties:

(2a. Reliability – precise specifications, testing; 2b. Validity – testing, threats to validity)

2a. Reliability: **Total votes-17; H-0; M-12; L-4; I-1**; 2b. Validity: **Total votes-17; H-0; M-10; L-4; I-3;** Post-Comment Validity Revote**: Total Votes-14; H-1; M-9; L-3; I-1**

Rationale:

- The Standing Committee noted that the reliability was similar to the previous measure (NQF #1884) and the reliability testing was conducted at the accountable-entity level. For adults, the SNR was 0.92 and 0.84 for adolescents.
- The Standing Committee had no concerns and passed the measure on reliability.
- The Standing Committee noted that validity testing was conducted at both the patient/encounter level using data quality checks and audits and at the accountable-entity level by comparing correlation to other constructs. Data audits found 94% of those audited passed the audited. The Standing Committee expressed concerns that not all critical data elements were tested, as well as whether "response at 12 months" was a valid goal compared to remission at 12 months, following the "response at six months" measure.
- Another Standing Committee member raised concerns that this measure does not account for the progress-relapse-progress nature of life, noting that the time frame may capture relapse instead of progress.
- Due to the concerns noted, the Standing Committee did not reach consensus on validity.
- During the post-comment meeting, the Standing Committee noted that this measure had the same concerns as NQF #0710e and agreed that the developer's responses alleviated those concerns.
- The Standing Committee passed the measure on validity.

3. Feasibility: Total votes-17; H-1; M-13; L-2; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that the measure is currently captured in an EHR in which groups can successfully extract the information from their EHR.
- The Standing Committee members noted that the six-month time frame was an advantage over the 12-month time frame toward detecting treatment resistance earlier but had no concerns and passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use: Total votes-17; Pass-13; No Pass-4; 4b. Usability: Total votes-17; H-1; M-10; L-5; I-1

Rationale:

- The Standing Committee noted that the measure is currently used in all primary care clinics in Minnesota as well as bordering communities in other states. The measure is also publicly reported on the MN HealthScores website and as part of the MNCM Annual Health Care Quality Report. It has also been selected for inclusion as a quality metric for CMS' Center for Medicare & Medicaid Innovation (CMMI) Innovation Model Kidney Care First. The Standing Committee had no concerns and passed the measure on use.
- The Standing Committee noted that promising improvements were noted between 2010 and 2019 for adults, although no trend data were available over time due to changes in the measure's specifications. The Standing Committee also noted that the benefits of this measure greatly outweigh the potential harms and passed it on usability.

5. Related and Competing Measures

- This measure is related to the following measures:
  - NQF #1884 Depression Response at Six Months – Progress Towards Remission
  - NQF #0711 Depression Remission at Six Months
  - NQF #0710e Depression Remission at 12 Months
  - NQF #0712 Depression Assessment 36 PHQ-9/PHQ-9M
- The Standing Committee did not recommend this measure for endorsement and therefore did not discuss related and competing measures during the post-comment call.

6. Standing Committee Recommendation for Endorsement: Total Votes-14; Yes-8; No-6

- During the post-comment meeting discussion on overall suitability for endorsement, the Standing Committee raised an additional concern: This measure may not be as clinically important as the other measures in the depression suite. The Standing Committee expressed that this measure did not demonstrate sufficient effort to achieve remission, particularly at 12 months versus six months. Specifically, if there is no progress toward remission at 12 months, there is a major issue, whereas six months would be a more reasonable expectation for progress.
- The Standing Committee voted to not recommend the measure for endorsement.

7. Public and Member Comment

- Two pre-evaluation public comments were received:
  - One comment was submitted by the measure developer to clarify validity testing.
  - One comment was submitted by a member of the public, which raised concerns about the evidence to support the measure, validity testing, and feasibility.
- Six post-evaluation public comments were received for this measure as noted in detail in the discussion of NQF #0710e:
  - Two comments were submitted by the measure developer.
    - In these comments, which apply to NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712 as well, the developer clarified measure details and addressed the Standing Committee's concerns about evidence, data element validity testing, missing data, and telemedicine that led to the "consensus not reached" status.
    - The Standing Committee discussed the developer's responses was satisfied with the clarification of the concerns raised during the measure evaluation meeting.
  - Two comments did not express support for the measure.
    - One public comment, which applies to the depression measures (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712), was submitted raising concerns with evidence, data sources, and the omission of telehealth as noted in detail above.
      - The developer replied to each of the commenter's points.
    - Another public comment raised concerns about the lack of telehealth inclusion.
      - In response to the commenter's concerns about telehealth services, the developer reiterated the explanation from the other non-supportive comment.
  - One comment expressed support for the measure.
    - The comment attributes the success of developing the clinic's care coordination system to the implementation of these measures and encouraged the Standing Committee to endorse the entire suite of measures.
  - One comment raised concerns with NQF's process.
    - The commenter raised concerns regarding NQF's endorsement of a PRO tool.
      - NQF provided a response clarifying that NQF does not endorse measurement tools.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-13; No-2 (December 9, 2022: Not Endorsed)

- The CSAC upheld the Standing Committee's decision to not recommend the measure for endorsement.

NQF #0712 Depression Assessment 36 With PHQ-9/PHQ-9M (MN Community Measurement)

Measure Worksheet | Specifications

Description: The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with a diagnosis of major depression or dysthymia who have a completed PHQ-9 or PHQ-9M tool during a four-month measurement period.
Numerator Statement: Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) included in the denominator who have at least one PHQ-9 or PHQ-9M tool administered and completed during a four-month measurement period.
Denominator Statement: Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with a diagnosis of major depression or dysthymia.
Exclusions: Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.
Adjustment/Stratification: N/A
Level of Analysis: Clinician: Group/Practice
Setting of Care: Outpatient Services
Type of Measure: Process
Data Source: Electronic Health Records
Measure Steward: MN Community Measurement

STANDING COMMITTEE MEETING June 30, 2022

1. Importance to Measure and Report:

(1a. Evidence, 1b. Performance Gap)

1a. Evidence: **Total votes-17; H-1; M-8; L-3; I-5**; 1b. Performance Gap: **Total Votes-17; H-1; M-8; L-3; I-5;** Post-Comment Evidence Revote**: Total Votes-15; H-1; M-7; L-6; I-1**

Rationale:

- The Standing Committee noted that the developer provided updated evidence for this measure that is related to the PHQ-9 being a validated tool and the use of the PHQ-9 being part of collaborative care, of which PHQ-9 performance is one component.
- The Standing Committee raised the following concern: The evidence for the measure would be stronger if it were linked to improved outcomes, which would also allow for more meaningful quality improvement. The Standing Committee recognized the challenges of providing such evidence and discussed that not administering a PHQ-9 would result in missed diagnoses.
- Due to the concerns noted, the Standing Committee did not reach consensus on evidence.
- During the post-comment meeting, the Standing Committee discussed the additional evidence provided by the developer.
- The developer provided new data from 26,000 patients to assess whether the frequency of the PHQ-9 assessment was associated with outcomes, noting that patients with three to 12 PHQ-9 assessments were three times more likely to achieve depression remission and response to treatment.
- The Standing Committee noted that it did not directly address the question about whether a single assessment of PHQ-9 was associated with improved outcomes. Specifically, the measure does not assess the number of assessments, only that an assessment was conducted.
- As a result of this discussion, the Standing Committee did not pass the measure on evidence.
- The Standing Committee noted that in 2020, the average performance score was 77.7% for adults and 78.4% for adolescents in Minnesota.
- The Standing Committee noted that the developer provided sufficient data showing a gap. It also noted that disparities exist and passed the measure on performance gap.

2. Scientific Acceptability of Measure Properties:

(2a. Reliability - precise specifications, testing; 2b. Validity - testing, threats to validity)

2a. Reliability: **Total votes-17; H-1; M-13; L-3; I-0**; 2b. Validity: **Total votes-17; H-0; M-14; L-3; I-0**

Rationale:

- The Standing Committee highlighted that the specifications have been updated since the measure's last endorsement, noting that changes include incorporating adolescents ages 12-17 into the measure, adding

the PHQ-9M PRO tool to the measure (modified for teens), modifying the exclusion value set of personality disorder, and adding exclusions for schizophrenia and pervasive developmental disorder.

- The Standing Committee asked for clarification regarding the statistic that patients who were frequently assessed with the PHQ-9 were about three times more likely to reach remission and response. The developer clarified that patients were divided into those who had received only one to three PHQ-9 assessments versus those who received four to 12 PHQ-9 assessments. For patients who received one to three PHQ-9 assessments, remissions rates were 6.3 compared to 15.8 for those assessed more frequently (i.e., four to 12 times).
- A Standing Committee member asked for specification clarifications regarding a patient who receives a PHQ-9 within four months, who then receives a depression diagnosis. They asked whether the PHQ-9 must be indexed within four months within that diagnosis or whether it is the follow-up to the first PHQ-9. The developer clarified that this measure is not related to identifying patients for index; rather, it focuses on a four-month period.
- The Standing Committee raised a concern: Not all EHR systems may record PHQ-9s uniformly or adequately; therefore, accurate completion of the metric may be difficult.
- A Standing Committee member asked whether the completion data element was reported as a "yes" or "no" for completion of the tool or whether it is a full metric of the scoring. The developer clarified that the expectation has always been completion of the tool and that an incomplete tool does not count. The developer also clarified that other related tools, such as the PHQ-8, PHQ-2, or PHQ-3, cannot be reported for this measure.
- The Standing Committee noted that reliability testing was conducted at the accountable-entity level using a beta-binomial model and found a reliability score of 0.932903 for adults and an average reliability score of 0.878959 for adolescents.
- The Standing Committee had no concerns and passed the measure on reliability.
- The Standing Committee noted that the correlation was performed against a depression outcome measure to test the hypothesis that clinics that do well assessing their patients with a diagnosis of depression frequently with the PHQ-9/PHQ-9M will also perform better in achieving remission (PHQ-9<5) at six months.
- The Standing Committee noted that the convergent validity testing at the accountable-entity level showed a positive but relatively weak correlation; it passed the measure on validity.
- While the Standing Committee did express that the correlation between NQF #0712 and NQF #0711 could be stronger, it agreed that a positive correlation existed and passed the measure on validity.

3. Feasibility: Total votes-17; H-2; M-13; L-1; I-1

(3a. Clinical data generated during care delivery; 3b. Electronic sources; 3c. Susceptibility to inaccuracies/ unintended consequences identified; 3d. Data collection strategy can be implemented)

Rationale:

- The Standing Committee noted that the data are collected during the regular course of care and are well integrated into most EHRs. It also noted that the PHQ-9 screening tool is free and publicly available.
- The Standing Committee passed the measure on feasibility.

4. Usability and Use:

(Used and useful to the intended audiences for 4a. Accountability and Transparency; 4b. Improvement; and 4c. Benefits outweigh evidence of unintended consequences)

4a. Use: Total votes-17; Pass-16; No Pass-1; 4b. Usability: Total votes-17; H-4; M-10; L-3; I-0

Rationale:

- The Standing Committee noted that the measure is used in Minnesota on the MN HealthScores website and publicly reported in all primary care clinics in Minnesota and many bordering communities. The Standing Committee passed the measure on use.
- A Standing Committee member noted that the PHQ-9 may not be a standard part of care in many settings since some accountability organizations allow for use of a variety of validated screening tools; however, another member noted that the measure has shown gradual improvement over time.
- The Standing Committee had no concerns about unintended consequences and passed the measure on usability.

5. Related and Competing Measures

- This measure is related to the following measures:
  - NQF #1884 Depression Response at Six Months – Progress Towards Remission
  - NQF #1885 Depression Response at 12 Months – Progress Towards Remission
  - NQF #0711 Depression Remission at Six Months
  - NQF #0710e Depression Remission at 12 Months
- The Standing Committee did not recommend this measure for endorsement and therefore did not discuss related and competing measures during the post-comment call.

6. Standing Committee Recommendation for Endorsement: Vote Not Taken

- The Standing Committee did not pass the measure on evidence during the post-comment meeting and therefore did not vote on overall suitability for endorsement.

7. Public and Member Comment

- Three pre-evaluation public comments were received:
  - One comment was submitted by the measure developer to clarify the evidence that supports the measure.
  - One comment was submitted by a member of the public, which raised concerns about whether this measure is intended to be an eCQM and the fact that this measure is very similar to #0710e.
  - Another comment was submitted by a member of the public, which supported this measure as well as #0710e, #0711, #0712, #1884 and #1885.
- Six post-evaluation public comments were received for this measure as noted in detail in the discussion of NQF #0710e:
  - Two comments were submitted by the measure developer.
    - In these comments, which apply to NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712 as well, the developer clarified measure details and addressed the Standing Committee's concerns about evidence, data element validity testing, missing data, and telemedicine that led to the "consensus not reached" status.
    - The Standing Committee discussed the developer's responses was satisfied with the clarification of the validity concerns raised during the measure evaluation meeting. The Standing Committee remained concerned regarding the evidence provided for this measure.
  - Two comments did not express support for the measure.
    - One public comment, which applies to the depression measures (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712), was submitted raising concerns with evidence, data sources, and the omission of telehealth as noted in detail above.
      - The developer replied to each of the commenter's points.
    - Another public comment raised concerns about the lack of telehealth inclusion.
      - In response to the commenter's concerns about telehealth services, the developer reiterated the explanation from the other non-supportive comment.
  - One comment expressed support for the measure.
    - The comment attributes the success of developing the clinic's care coordination system to the implementation of these measures and encouraged the Standing Committee to endorse the entire suite of measures.
  - One comment raised concerns with NQF's process.
    - The commenter raised concerns regarding NQF's endorsement of a PRO tool.
      - NQF provided a response clarifying that NQF does not endorse measurement tools.

8. Consensus Standards Approval Committee (CSAC) Endorsement Decision: Total votes-15; Yes-12; No-3 (December 9, 2022: Not Endorsed)

- The CSAC upheld the Standing Committee's decision to not recommend the measure for endorsement.

Appendix B: Behavioral Health and Substance Use Portfolio—Use in Federal Programs*

NQF #	Title	Federal Programs (Finalized or Implemented)
0004	Initiation and Engagement of Alcohol and Other Drug Dependence Treatment	None
0028	Preventive Care & Screening: Tobacco Use: Screening & Cessation Intervention	None
0105	Antidepressant Medication Management (AMM)	Medicaid: Adult Core Set
0108	Follow-Up Care for Children Prescribed ADHD Medication (ADD)	HEDIS Quality Measure Rating System
0576	Follow-Up After Hospitalization for Mental Illness (FUH)	None
0640	HBIPS– Two Hours of Physical Restraint Use	Hospital Compare
0641	HBIPS– Three Hours of Seclusion Use	Hospital Compare
0710e	Depression Remission at 12 Months (eMeasure)	Merit-Based Incentive Payment System (MIPS)
0711	Depression Remission at Six Months	None
1879	Adherence to Antipsychotic Medications for Individuals With Schizophrenia	Merit-Based Incentive Payment System (MIPS) Medicaid: Adult Core Set HEDIS Quality Measure Rating System
1884	Depression Response at Six Months – Progress Towards Remission	None
1932	Diabetes Screening for People With Schizophrenia or Bipolar Disorder Who Are Using Antipsychotic Medications (SSD)	None
2152	Preventive Care and Screening: Unhealthy Alcohol Use: Screening & Brief Counseling	Merit-Based Incentive Payment System (MIPS) Doctors & Clinicians Compare
2607	Diabetes Care for People With Serious Mental Illness: Hemoglobin A1c (HbA1c) Poor Control (>9.0%)	Medicaid: Adult Core Set
2800	Metabolic Monitoring for Children and Adolescents on Antipsychotics	None
2801	Use of First-Line Psychosocial Care for Children and Adolescents on Antipsychotics	None
3175	Continuity of Pharmacotherapy for Opioid Use Disorder	None

NQF #	Title	Federal Programs (Finalized or Implemented)
3312	Continuity of Care After Medically Managed Withdrawal From Alcohol and/or Drugs	None
3332	Psychosocial Screening Using the Pediatric Symptom Checklist-Tool (PSC-Tool)	None
3389	Concurrent Use of Opioids and Benzodiazepines (COB)	Medicaid: Adult Core Set
3400	Use of Pharmacotherapy for Opioid Use Disorder (OUD)	HEDIS Quality Measure Rating System Medicaid: IAP Medicaid: Adult Core Set
3453	Continuity of Care After Inpatient or Residential Treatment for Substance Use Disorder (SUD)	None
3488	Follow-Up After Emergency Department Visit for Alcohol and Other Drug Abuse or Dependence	None
3489	Follow-Up After Emergency Department Visit for Mental Illness	None
3539e	Use of Antipsychotics in Older Adults in the Inpatient Hospital Setting	None
3541	Annual Monitoring for Persons on Long-Term Opioid Therapy (AMO)	Marketplace Quality Rating System (QRS)
3589	Prescription or Administration of Pharmacotherapy to Treat Opioid Use Disorder (OUD)	None
3590	Continuity of Care After Receiving Hospital or Residential Substance Use Disorder (SUD) Treatment	None

* CMS Measures Inventory Tool Last Accessed on January 24, 2023.

Appendix C: Behavioral Health and Substance Use Standing Committee and NQF Staff

STANDING COMMITTEE

Harold Pincus, MD (Co-Chair)
Professor and Vice Chair of Psychiatry and Co-Director of the Irving Institute for Clinical and Translational Research, Columbia University
New York, New York

Michael Trangle, MD (Co-Chair)
Senior Fellow, HealthPartners Institute for Research and Education
Bloomington, Minnesota

Loretta (Ann) Bostic, DNP, MBA, APRN, CRNA, PMHNP-BC
Program Director, MedSurg Behavioral Health Services
Charleston, West Virginia

Caroline Carney, MD, MSc, FAMP, CPHQ
Chief Medical Officer, Magellan Rx Management, Magellan Health, Inc.
Phoenix, Arizona

Vitka Eisen, MSW, Ed.D
President and Chief Executive Officer, HealthRIGHT 360
San Francisco, California

Julie Goldstein Grumet, PhD
Director of Zero Suicide Institute/Senior Health Care Advisor, Suicide Prevention Resource Center
Education Development Center
Washington, District of Columbia

Benjamin Han, MD, MPH
Assistant Professor, University of California, San Diego
La Jolla, California

Morissa Henn, MPH, DrPH
Community Health Director, Intermountain Healthcare
Salt Lake City, Utah

Lisa Jensen, DNP, APRN
Director of Clinical Operations, Specialty Care Services, Veterans Health Administration
North Salt Lake, Utah

Caitlyn Kjolhede, BSN, MBA
Vice President of Clinical Performance, CHESS Health
East Rochester, New York

Kraig Knudsen, PhD
Chief, Bureau of Research and Evaluation, Ohio Department of Mental Health and Addiction Services
Columbus, Ohio

Barbara Lang, LPC, LISAC
Deputy Chief of Standards and Compliance, Community Bridges, Inc.
Yuma, Arizona

Michael R. Lardieri, LCSW
Assistant Vice President Strategic Program Development, Northwell Health, Behavioral Health Services Line
Glen Oaks, New York

Raquel Mazon Jeffers, MPH, MIA
Senior Director, U.S. Initiatives, Community Health Acceleration Partnership
Hopewell, New Jersey

Brooke Parish, MD
Executive Medical Director, Health Care Service Corporation, Blue Cross Blue Shield of New Mexico
Albuquerque, New Mexico

David Pating, MD
Associate Clinical Professor, Addiction Psychiatry – Substance Use Services, San Francisco Department of Public Health
San Francisco, California

Vanita Pindolia, PharmD
Vice President, Ambulatory Clinical Pharmacy Programs, Henry Ford Health System (HFHS)/Health Alliance Plan (HAP)
Detroit, Michigan

Chantelle Rice Collins, OTD, OTR/L, CDCES
Associate Professor, University of Southern California
Los Angeles, California

Andrew Sperling, JD
Director of Federal Advocacy, National Alliance on Mental Illness
Arlington, Virginia

Jeffery Susman, MD
Dean and VPHA, Emeritus, Northeast Ohio Medical University
Rootstown, Ohio

Allen Tien, MD
President and Director of Applied Research, Medical Decision Logic, Inc.
Baltimore, Maryland

Patrick Triplett, MD
Clinical Director and Physician Advisor, Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine
Baltimore, MD

Heidi Waters, MBA, PhD
Director, Health Economics and Outcomes Research, Otsuka Pharmaceutical Development and Commercialization, Inc.
Rockville, Maryland

Bonnie Zima, MD, MPH
Professor in Residence, Child and Adolescent Psychiatry, Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles (UCLA)
Los Angeles, California

NQF STAFF

Elizabeth Drye, MD, MS
Chief Scientific Officer, Measurement Science and Application

Tricia Elliott, DHA, MBA, CPHQ, FNAHQ
Vice President, Measurement Science and Application *(Former)*

Matthew K. Pickering, PharmD
Managing Director, Measurement Science and Application

Poonam Bal, MHSA
Senior Director, Measurement Science and Application (*Former*)

Elizabeth Freedman, MPH
Senior Director, Measurement Science and Application

Leah Chambers, MHA
Director, Measurement Science and Application

Tamara H. Funk, MPH
Director, Measurement Science and Application (*Former*)

Erin Buchanan, MPH
Senior Manager, Measurement Science and Application

Hannah Ingber, MPH
Manager, Measurement Science and Application (*Former*)

Sean Sullivan, MA
Analyst, Measurement Science and Application

Yemsrach Kidane, PMP
Senior Project Manager, Program Operations

Jesse Pines, MD, MBA, MSCE
Consultant

Appendix D: Measure Specifications

NQF #0710e Depression Remission at 12 Months

STEWARD

MN Community Measurement

DESCRIPTION

The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission twelve months (+/- 60 days) after an index visit.

TYPE

Outcome: PRO-PM

DATA SOURCE

Electronic Health Records

The data source is the medical group's/clinic's medical record information, most frequently from an EMR. A CSV file is created by each medical group and uploaded to a password protected, HIPAA secure data portal which performs rate calculation. Selected Patient Reported Data, not because it is necessarily a separate data source, but because this measure is based on a patient reported outcome tool, a PRO-PM measure. Frequently this PRO tool, the PHQ-9, is housed within a clinic's EMR, or in paper charts is a part of the patient's medical record.

Data collection guides are located on our corporate website:

- Direct data submission https://helpdesk.mncm.org/helpdesk/KB/View/24186732-depression-care-guides
- PIPE data files https://helpdesk.mncm.org/helpdesk/KB/View/41088290-pipe-data-files

PROM

The PHQ-9 depression assessment tool is a patient reported outcome tool that is in the public domain and can be obtained for free use on the Patient Health Questionnaire (PHQ) Screeners website at www.phqscreeners.com. Modes of administration include traditional paper, mail, electronic and telephonic. The tool is available on the website with 79 language translations available.

The PHQ-9 tool is validated for use as a measure to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (to determine treatment response). [Löwe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42:1194-1201 and Kroenke

K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010]

The PHQ-9M is a modified version of the PHQ-9 tool for adolescents. Please refer to discussion in question sp.27

LEVEL

Clinician: Group/Practice

SETTING

Outpatient Services

NUMERATOR STATEMENT

The number of patients in the denominator who reached remission, with a PHQ-9 or PHQ-9M result less than five, twelve months (+/- 60 days) after an index visit.

NUMERATOR DETAILS

This PROM-PM outcome measure is longitudinal, seeking to measure the absence of depression symptoms (remission) within twelve months (+/- 60 days) for the patients with an index event of depression, measured as an elevated PHQ-9 or PHQ-9M.

The numerator is defined as patients with a twelve-month (+/- 60 days) PHQ-9 or PHQ-9M score of less than five.

The numerator rate is calculated as follows:

pts with major depression or dysthymia with a PHQ-9 or PHQ-9M score < 5 at 12 months(+/- 60 days)/

pts with major depression or dysthymia with index contact PHQ-9 > 9

Patients who do not have a twelve month +/- 60 day PHQ-9 or PHQ-9M score obtained remain in the denominator and are counted as not being in remission. Not having a PHQ-9 or PHQ-9M score within the 120 day window is considered a numerator miss.

Time period for data collection: there is a set index period for this measure, typically patients who have an index visit within a calendar period (e.g. index dates between 11/1/2017 and 10/31/2018) and then allowing enough time to pass to accommodate the timeframe for assessment. (e.g. for remission at twelve months +/- 60 days with index dates of service ending 10/31/2018, the assessment period for twelve month remission would go through 12/30/2019). Technically, the six- and twelve-month remission measures are collected together in the MN program, and the index assessment period is fourteen months in duration.

Denominator identification period (index) 11/1/2017 to 10/31/2018

Measure assessment period through 12/30/2019; reported in 2020

DENOMINATOR STATEMENT

Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

DENOMINATOR DETAILS

The target population, patients age 12 and older with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine, is identified as follows:

Patients age 12 and older at the time of the index visit AND Index visit

An index visit occurs when ALL of the following criteria are met during a face-to-face visit or contact with an eligible provider:

- a PHQ-9 or PHQ-9M result greater than nine
- an active diagnosis of Major Depression or Dysthymia (Major Depression or Dysthymia Value Set)
- the patient is NOT in a prior index period

An index period begins with an index visit and is 14 months in duration.

Denominator is stratified by age range for adolescents (12 to 17 years of age) and adults (18 years of age and older).

Patients who do not have a twelve month +/- 60 day follow-up PHQ-9 or PHQ-9M score obtained remain in the denominator for this measure.

Please refer to the attached data dictionary for an inclusive list of all ICD-9/ICD-10 codes and data element definitions.

EXCLUSIONS

Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

EXCLUSION DETAILS

Required exclusions:

- Patient had a diagnosis of Bipolar Disorder (Bipolar Disorder Value Set) any time prior to the end of their measure assessment period
- Patient had an active diagnosis of Schizophrenia or Psychotic Disorder (Schizophrenia Psychotic Disorder Value Set) any time prior to the end of their measure assessment period

Allowable exclusions:

- Patient had an active diagnosis of Personality Disorder – Emotionally Labile Conditions (Personality Disorder – Emotionally Labile Value Set) any time prior to the end of their measurement assessment period
- Patient had an active diagnosis of Pervasive Developmental Disorder (Pervasive Disorder Value Set) any time prior to the end of the measurement assessment period
- Patient was a permanent nursing home resident at any time during the denominator identification period or measure assessment period
- Patient was in hospice or receiving palliative care (Palliative Care Value Set) at any time during the denominator identification or measure assessment period
- Patient died prior to the end of their measurement assessment period

The direct data submission process in MN allows for both up-front exclusions of the population and, because this is a longitudinal outcome measure, processes are in place to allow exclusions that may occur after index during the course of the measurement assessment period. Please see field specifications in the attached data dictionary.

RISK ADJUSTMENT

Yes - Additional risk adjustment analysis is included Statistical risk model with risk factors (specify number of risk factors) MNCM uses Logistic Regression Modeling to create values supporting a method of Indirect Standardization Risk Adjustment, commonly referred to as Expected Value. Indirect standardization does not change the actual performance rates, rather answers the question: "If all providers had this medical group/ clinic's mix of patients, what would the statewide average be?". This method compares the provider's actual performance to the expected rate.

Example Clinic X	Unadjusted	Standardized to Clinic X Patient Mix
Statewide	39%	32%
Clinic X	35%	35%

Clinic X vs Statewide Below Above (Actual : Expected = 1.09

Risk variables used for this measure include age, initial PHQ-9/ PHQ-9M score, insurance product and patient neighborhood deprivation index (based on zip-code). Deprivation index includes use of SNAP benefits, living under the poverty level, unemployed status, public assistance, and single female with children. In MN, the ratio ranges are -6.41 (Red Lake) to +1.42 (Flom) with a mean of zero. "A measure of census-tract neighborhood deprivation is likely a good proxy for a range of individual-level and true area-level constructs relevant to outcomes of interest and feasible to obtain." [National Academies of Sciences, Engineering, and Medicine, 2017: Accounting for Social Risk Factors in Medicare Payment]

2021 Submission

12 Month Remission- Adults

Analysis of Maximum Likelihood Estimates

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.333	0.0312	5599.0401	<.0001
pt_age	1	0.0128	0.000661	373.07	<.0001
mdcr	1	-0.342	0.0309	122.3173	<.0001
mhcp	1	-0.5529	0.0279	392.2078	<.0001
unins	1	-0.3618	0.0586	38.0577	<.0001
undt	1	-0.4279	0.0349	150.4565	<.0001
mod_severe	1	0.2974	0.0219	184.1034	<.0001
severe	1	-0.5849	0.0297	386.796	<.0001
dep_idx	1	0.1336	0.013	106.4387	<.0001

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; 12 Month Remission Adults

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

12 Month Remission- Adolescents

Analysis of Maximum Likelihood Estimates 12 Month Remission- Adolescents

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6113	0.3501	21.1784	<.0001
pt_age	1	-0.0354	0.0229	2.3882	0.1223
mdcr	1	-0.1082	0.304	0.126	0.7221
mhcp	1	-0.2457	0.0856	8.2411	0.0041
unins	1	-1.0255	0.3249	9.9604	0.0016
undt	1	-0.1429	0.1139	1.5742	0.2096
mod_severe	1	-0.3665	0.0777	22.2451	<.0001
severe	1	-0.6361	0.1024	38.571	<.0001
dep_idx	1	0.1039	0.0541	3.6952	0.0546

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; 12 Month Remission Adolescents

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

Definitions for Logistic Model

AIC – This is the Akaike Information Criterion. It is calculated as AIC = -2 Log L + 2((k-1) + s), where k is the number of levels of the dependent variable and s is the number of predictors in the model. AIC is used for the comparison of nonnested models on the same sample. Ultimately, the model with the smallest AIC is considered the best, although the AIC value itself is not meaningful.

SC – This is the Schwarz Criterion. It is defined as – 2 Log L + ((k-1) + s)*log(Σ fi), where fi's are the frequency values of the i^th observation, and k and s were defined previously. Like AIC, SC penalizes for the number of predictors in the model and the smallest SC is most desirable and the value itself is not meaningful.

-2 Log L – This is negative two times the log-likelihood. The -2 Log L is used in hypothesis tests for nested models and the value in itself is not meaningful.

Intercept Only – This column refers to the respective criterion statistics with no predictors in the model, i.e., just the response variable.

Intercept and Covariates – This column corresponds to the respective criterion statistics for the fitted model. A fitted model includes all independent variables and the intercept. We can compare the values in this column with the criteria corresponding Intercept Only value to assess model fit/significance.

Test – These are three asymptotically equivalent Chi-Square tests. They test against the null hypothesis that at least one of the predictors' regression coefficient is not equal to zero in the model. The difference between them are where on the log-likelihood function they are evaluated.

Likelihood Ratio – This is the Likelihood Ratio (LR) Chi-Square test that at least one of the predictors' regression coefficient is not equal to zero in the model. The LR Chi-Square statistic can be calculated by -2 Log L(null model) – 2 Log L(fitted model) = 231.289-160.236 = 71.05, where L(null model) refers to the Intercept Only model and L(fitted model) refers to the Intercept and Covariates model.

Score – This is the Score Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Wald – This is the Wald Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Chi-Square, DF and Pr > ChiSq – These are the Chi-Square test statistic, Degrees of Freedom (DF) and associated p-value (PR>ChiSq) corresponding to the specific test that all of the predictors are simultaneously equal to zero. We are testing the probability (PR>ChiSq) of observing a Chi-Square statistic as extreme as, or more so, than the observed one under the null hypothesis; the null hypothesis is that all of the regression coefficients in the model are equal to zero. The DF defines the distribution of the Chi-Square test statistics and is defined by the number of predictors in the model. Typically, PR>ChiSq is compared to a specified alpha level, our willingness to accept a type I error, which is often set at 0.05 or 0.01. The small p-value from all three tests would lead us to conclude that at least one of the regression coefficients in the model is not equal to zero.

STRATIFICATION

This measure is stratified by age range and results are reported separately by age: Adolescents (12-17 years of age) and Adults (18 years of age and older).

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

COPYRIGHT / DISCLAIMER

PHQ-9: Copyright

Developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke and colleagues, with an educational grant from Pfizer Inc. No permission required to reproduce, translate, display or distribute. N/A

NQF #0711 Depression Remission at Six Months

STEWARD

MN Community Measurement

DESCRIPTION

The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who reach remission six months (+/- 60 days) after an index visit.

TYPE

Outcome: PRO-PM

DATA SOURCE

Electronic Health Records, Electronic Health Records: Electronic Health Records, Other The data source is the medical group's/ clinic's medical record information, most frequently from an EMR. A CSV file is created by each medical group and uploaded to a password protected, HIPAA secure data portal which performs rate calculation. Selected Patient Reported Data, not because it is necessarily a separate data source, but because this measure is based on a patient reported outcome tool, a PRO-PM measure. Frequently this PRO tool, the PHQ-9, is housed within a clinic's EMR, or in paper charts is a part of the patient's medical record.

Data collection guides are located on our corporate website:

- direct data submission https://helpdesk.mncm.org/helpdesk/KB/View/24186732-depression-care-guides
- PIPE data files https://helpdesk.mncm.org/helpdesk/KB/View/41088290-pipe-data-files

PROM

The PHQ-9 depression assessment tool is a patient reported outcome tool that is in the public domain and can be obtained for free use on the Patient Health Questionnaire (PHQ) Screeners website at www.phqscreeners.com. Modes of administration include traditional paper, mail, electronic and telephonic. The tool is available on the website with 79 language translations available.

The PHQ-9 tool is validated for use as a measure to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (to determine treatment response). [Löwe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42:1194-1201 and Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010]

The PHQ-9M is a modified version of the PHQ-9 tool for adolescents. Please refer to discussion in question sp.27

LEVEL

Clinician: Group/Practice

SETTING

Outpatient Services

NUMERATOR STATEMENT

The number of patients in the denominator who reached remission, with a PHQ-9 or PHQ-9M result less than five, six months (+/- 60 days) after an index visit.

NUMERATOR DETAILS

This PROM-PM outcome measure is longitudinal, seeking to measure the absence of depression symptoms (remission) within six months (+/- 60 days) for the patients with an index event of depression, measured as an elevated PHQ-9 or PHQ-9M.

The numerator is defined as patients with a six-month (+/- 60 days) PHQ-9 or PHQ-9M score of less than five.

The numerator rate is calculated as follows:

pts with major depression or dysthymia with a PHQ-9 or PHQ-9M score < 5 at 6 months(+/- 60 days)/ # pts with major depression or dysthymia with index contact PHQ-9 > 9

Patients who do not have a six month +/- 60 day PHQ-9 or PHQ-9M score obtained remain in the denominator and are counted as not being in remission. Not having a PHQ-9 or PHQ-9M score within the 120 day window is considered a numerator miss.

Time period for data collection: there is a set index period for this measure, typically patients who have an index visit within a calendar period (e.g. index dates between 11/1/2017 and 10/31/2018) and then allowing enough time to pass to accommodate the timeframe for assessment. (e.g. for remission at six months +/- 60 days with index dates of service ending 10/31/2018, the assessment period for twelve month remission [to also capture 12 month remission rates] would go through 12/30/2019). Technically, the six- and twelve-month remission measures are collected together in the MN program, and the index assessment period is fourteen months in duration.

Denominator identification period (index) 11/1/2017 to 10/31/2018

Measure assessment period through 12/30/2019; reported in 2020

DENOMINATOR STATEMENT

Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

DENOMINATOR DETAILS

The target population, patients age 12 and older with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine, is identified as follows:

Patients age 12 and older at the time of the index visit AND Index visit

An index visit occurs when ALL of the following criteria are met during a face-to-face visit or contact with an eligible provider:

- a PHQ-9 or PHQ-9M result greater than nine
- an active diagnosis of Major Depression or Dysthymia (Major Depression or Dysthymia Value Set)
- the patient is NOT in a prior index period

An index period begins with an index visit and is 14 months in duration.

Denominator is stratified by age range for adolescents (12 to 17 years of age) and adults (18 years of age and older).

Patients who do not have a six month +/- 60 day follow-up PHQ-9 or PHQ-9M score obtained remain in the denominator for this measure.

Please refer to the attached data dictionary for an inclusive list of all ICD-9/ICD-10 codes and data element definitions.

EXCLUSIONS

Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

EXCLUSION DETAILS

Required exclusions:

- Patient had a diagnosis of Bipolar Disorder (Bipolar Disorder Value Set) any time prior to the end of their measure assessment period
- Patient had an active diagnosis of Schizophrenia or Psychotic Disorder (Schizophrenia Psychotic Disorder Value Set) any time prior to the end of their measure assessment period

Allowable exclusions:

- Patient had an active diagnosis of Personality Disorder – Emotionally Labile Conditions (Personality Disorder – Emotionally Labile Value Set) any time prior to the end of their measurement assessment period
- Patient had an active diagnosis of Pervasive Developmental Disorder (Pervasive Disorder Value Set) any time prior to the end of the measurement assessment period
- Patient was a permanent nursing home resident at any time during the denominator identification period or measure assessment period
- Patient was in hospice or receiving palliative care (Palliative Care Value Set) at any time during the denominator identification or measure assessment period

- Patient died prior to the end of their measurement assessment period

The direct data submission process in MN allows for both up-front exclusions of the population and, because this is a longitudinal outcome measure, processes are in place to allow exclusions that may occur after index during the course of the measurement assessment period. Please see field specifications in the attached data dictionary.

RISK ADJUSTMENT

Yes - Additional risk adjustment analysis is included Statistical risk model with risk factors (specify number of risk factors) MNCM uses Logistic Regression Modeling to create values supporting a method of Indirect Standardization Risk Adjustment, commonly referred to as Expected Value. Indirect standardization does not change the actual performance rates, rather answers the question: "If all providers had this medical group/ clinic's mix of patients, what would the statewide average be?". This method compares the provider's actual performance to the expected rate.

Example Clinic X Unadjusted Standardized to Clinic X Patient Mix

Statewide 39% 32%

Clinic X 35% 35%

Clinic X vs Statewide Below Above (Actual : Expected = 1.09

Risk variables used for this measure include age, initial PHQ-9/ PHQ-9M score, insurance product and patient neighborhood deprivation index (based on zip-code). Deprivation index includes use of SNAP benefits, living under the poverty level, unemployed status, public assistance, and single female with children. In MN, the ratio ranges are -6.41 (Red Lake) to +1.42 (Flom) with a mean of zero. "A measure of census-tract neighborhood deprivation is likely a good proxy for a range of individual-level and true area-level constructs relevant to outcomes of interest and feasible to obtain." [National Academies of Sciences, Engineering, and Medicine, 2017: Accounting for Social Risk Factors in Medicare Payment]

2021 Submission

6 Month Remission- Adults

Analysis of Maximum Likelihood Estimates

Depression Remission at 6 Months- Adults

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.1358	0.0296	5216.7102	<.0001
pt_age	1	0.0105	0.00063	276.2196	<.0001
mdcr	1	-0.2731	0.0296	85.0788	<.0001

mhcp	1	-0.4861	0.0262	344.1504	<.0001
unins	1	-0.461	0.0584	62.2767	<.0001
undt	1	-0.3461	0.0326	112.9074	<.0001
mod_severe	1	-0.2918	0.0209	195.8009	<.0001
severe	1	-0.6009	0.0283	449.5633	<.0001
dep_idx	1	0.1384	0.0123	126.2771	<.0001

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; Adults

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

6 Month Remission- Adolescents

Analysis of Maximum Likelihood Estimates

Depression Remission at 6 Months- Adolescents

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.784	0.3411	5.2836	0.0215
pt_age	1	-0.0876	0.0224	15.2209	<.0001
mdcr	1	-0.235	0.3166	0.5539	0.4567
mhcp	1	-0.3018	0.0851	12.5672	0.0004
unins	1	-0.7365	0.2779	7.0237	0.008
undt	1	-0.1974	0.1142	2.9887	0.0838
mod_severe	1	-0.3322	0.0763	18.9391	<.0001
severe	1	-0.699	0.1038	45.3599	<.0001
dep_idx	1	0.0444	0.052	0.7299	0.3929

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; Adolescents

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

Definitions for Logistic Model

AIC – This is the Akaike Information Criterion. It is calculated as AIC = -2 Log L + 2((k-1) + s), where k is the number of levels of the dependent variable and s is the number of predictors in

the model. AIC is used for the comparison of nonnested models on the same sample. Ultimately, the model with the smallest AIC is considered the best, although the AIC value itself is not meaningful.

SC – This is the Schwarz Criterion. It is defined as – 2 Log L + ((k-1) + s)*log(Σ fi), where fi's are the frequency values of the i^th observation, and k and s were defined previously. Like AIC, SC penalizes for the number of predictors in the model and the smallest SC is most desirable and the value itself is not meaningful.

-2 Log L – This is negative two times the log-likelihood. The -2 Log L is used in hypothesis tests for nested models and the value in itself is not meaningful.

Intercept Only – This column refers to the respective criterion statistics with no predictors in the model, i.e., just the response variable.

Intercept and Covariates – This column corresponds to the respective criterion statistics for the fitted model. A fitted model includes all independent variables and the intercept. We can compare the values in this column with the criteria corresponding Intercept Only value to assess model fit/significance.

Test – These are three asymptotically equivalent Chi-Square tests. They test against the null hypothesis that at least one of the predictors' regression coefficient is not equal to zero in the model. The difference between them are where on the log-likelihood function they are evaluated.

Likelihood Ratio – This is the Likelihood Ratio (LR) Chi-Square test that at least one of the predictors' regression coefficient is not equal to zero in the model. The LR Chi-Square statistic can be calculated by -2 Log L(null model) – 2 Log L(fitted model) = 231.289-160.236 = 71.05, where L(null model) refers to the Intercept Only model and L(fitted model) refers to the Intercept and Covariates model.

Score – This is the Score Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Wald – This is the Wald Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Chi-Square, DF and Pr > ChiSq – These are the Chi-Square test statistic, Degrees of Freedom (DF) and associated p-value (PR>ChiSq) corresponding to the specific test that all of the predictors are simultaneously equal to zero. We are testing the probability (PR>ChiSq) of observing a Chi-Square statistic as extreme as, or more so, than the observed one under the null hypothesis; the null hypothesis is that all of the regression coefficients in the model are equal to zero. The DF defines the distribution of the Chi-Square test statistics and is defined by the number of predictors in the model. Typically, PR>ChiSq is compared to a specified alpha level, our willingness to accept a type I error, which is often set at 0.05 or 0.01. The small p-value from all three tests would lead us to conclude that at least one of the regression coefficients in the model is not equal to zero.

STRATIFICATION

This measure is stratified by age range and results are reported separately by age: Adolescents (12-17 years of age) and Adults (18 years of age and older).

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

COPYRIGHT / DISCLAIMER

PHQ-9: Copyright

Developed by Drs. Robert L. Spitzer, Janet B.W. Williams, Kurt Kroenke and colleagues, with an educational grant from Pfizer Inc. No permission required to reproduce, translate, display or distribute. N/A

NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M

STEWARD

MN Community Measurement

DESCRIPTION

The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with a diagnosis of major depression or dysthymia who have a completed PHQ-9 or PHQ-9M tool during a four month measurement period.

TYPE

Process

DATA SOURCE

Other, Electronic Health Records The data source is the medical group's/clinic's medical record information, most frequently from an EMR. A CSV file is created by each medical group and uploaded to a password protected, HIPAA secure data portal which performs rate calculation.

PROM

The PHQ-9 depression assessment tool is a patient reported outcome tool that is in the public domain and can be obtained for free use on the Patient Health Questionnaire (PHQ) Screeners website at www.phqscreeners.com. Modes of administration include traditional paper, mail, electronic and telephonic. The tool is available on the website with 79 language translations available.

The PHQ-9 tool is validated for use as a measure to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (to determine treatment response). [Löwe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42:1194-1201 and Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010]

The PHQ-9M is a modified version of the PHQ-9 tool for adolescents. Please refer to discussion in question sp.27

LEVEL

Clinician: Group/Practice

SETTING

Outpatient Services

NUMERATOR STATEMENT

Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) included in the denominator who have at least one PHQ-9 or PHQ-9M tool administered and completed during a four month measurement period.

NUMERATOR DETAILS

The total number of unique adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) in the denominator who had a least one PHQ-9 or PHQ-9M tool administered and completed during a four month measurement period in which a visit or contact with the patient has occurred.

Partially completed tools (e.g. answering 6 of the 9 questions) do not count as a completed tool. A valid PHQ-9 or PHQ-9M requires the completion of all nine questions for accurate scoring.

The numerator rate is calculated as follows:

pts with major depression or dysthymia with one or more completed PHQ-9 or PHQ-9M tools/

pts with major depression or dysthymia with a visit or contact during the measurement period

Rates are stratified by adolescents (12 to 17 years of age) and adults (18 years of age or older).

Time period for data collection: four month measurement periods (In the MN program 2/01 to 5/31, 6/01 to 9/30 and 10/01 to 1/31) with dates of service occurring within the four month period.

DENOMINATOR STATEMENT

Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with a diagnosis of major depression or dysthymia.

DENOMINATOR DETAILS

The target population, patients age 12 and older with the diagnosis of major depression or dysthymia, regardless of severity level of the PHQ-9 or PHQ-9M.

The number of unique patients who had a least one visit or contact with a provider during the measurement period with a diagnosis of major depression or dysthymia (Major Depression or Dysthymia Value Set). Contact is defined as visit, telephone call, e-visit or other contact that is associated with a PHQ-9 tool being completed by the patient.

EXCLUSIONS

Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

EXCLUSION DETAILS

Required exclusions:

- Patient had a diagnosis of Bipolar Disorder (Bipolar Disorder Value Set) any time prior to the end of the measurement period
- Patient had an active diagnosis of Schizophrenia or Psychotic Disorder (Schizophrenia Psychotic Disorder Value Set) any time prior to the end of the measurement period

Allowable exclusions:

- Patient died prior to the end of the measurement period
- Patient was a permanent nursing home resident at any time during the measurement period
- Patient was in hospice or receiving palliative care at any time during the measurement period (Palliative Care Value Set)
- Patient had a diagnosis of Personality Disorder – Emotionally Labile Conditions (Personality Disorder – Emotionally Labile Value Set) any time prior to the end of the measurement period
- Patient had an active diagnosis of Pervasive Developmental Disorder (Pervasive Disorder Value Set) any time prior to the end of the measurement period

- The direct data submission process in MN allows for both up-front exclusions of the population and, because this is a longitudinal outcome measure, processes are in place to allow exclusions that may occur after index during the course of the measurement assessment period. Please see field specifications in the attached data dictionary.

RISK ADJUSTMENT

No additional risk adjustment analysis included

No risk adjustment or stratification

Not applicable

STRATIFICATION

This measure is stratified by age range and results are reported separately by age: Adolescents (12-17 years of age) and Adults (18 years of age and older).

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

Historically, this measure is calculated by submitting a count of patients for the denominator and a count of patients in the numerator to a HIPAA secure data portal as part of the process in uploading a detailed patient file to calculate the six and twelve month remission outcome rates. MNCM is in the process of onboarding MN practice to a new warehouse (PIPE) and will calculate this measure centrally for practices based on encounter level data; full statewide transition to PIPE is planned for 2024.

The numerator rate is calculated as follows:

# of adolescent and adult pts with major depression or dysthymia with at least one PHQ-9 or PHQ-9M tool administered during the four month measurement period/

# of adolescent and adult pts with major depression or dysthymia

Query processes that medical groups follow to obtain counts:

During the four month measurement period (e.g. dates of service 6/1/2020 to 9/30/2020) how many patients had an office visit or other contact (phone, email) and diagnosis codes for major depression or dysthymia? (denominator)

Of these patients, how many had a PHQ-9 or PHQ-9M tool administered? (numerator)

The counting process is validated during the denominator certification process (where groups document all steps in identifying the depression population). Groups are asked to describe the process they use for obtaining the counts. Denominator documents are reviewed (certified) by MNCM staff prior to data collection and submission. This is to ensure that all groups are identifying their population correctly.

COPYRIGHT / DISCLAIMER

NQF #1884 Depression Response at Six Months – Progress Towards Remission

STEWARD

MN Community Measurement

DESCRIPTION

The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who are progressing towards remission by achieving a response (PHQ-9 or PHQ-9M score reduced by 50% or greater) six months (+/- 60 days) after an index visit.

TYPE

Outcome: PRO-PM

DATA SOURCE

Electronic Health Records: Electronic Health Records, Other, Electronic Health Records The data source is the medical group's/clinic's medical record information, most frequently from am EMR. A CSV file is created by each medical group and uploaded to a password protected, HIPAA secure data portal which performs rate calculation. Selected Patient Reported Data, not because it is necessarily a separate data source, but because this measure is based on a patient reported outcome tool, a PRO-PM measure. Frequently this PRO tool, the PHQ-9, is housed within a clinic's EMR, or in paper charts is a part of the patient's medical record.

PROM

The PHQ-9 depression assessment tool is a patient reported outcome tool that is in the public domain and can be obtained for free use on the Patient Health Questionnaire (PHQ) Screeners

website at www.phqscreeners.com. Modes of administration include traditional paper, mail, electronic and telephonic. The tool is available on the website with 79 language translations available.

The PHQ-9 tool is validated for use as a measure to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (to determine treatment response). [Löwe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42:1194-1201 and Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010]

The PHQ-9M is a modified version of the PHQ-9 tool for adolescents. Please refer to discussion in question sp.27

LEVEL

Clinician: Group/Practice

SETTING

Outpatient Services

NUMERATOR STATEMENT

The number of patients in the denominator who achieved a response as demonstrated by a PHQ-9 or PHQ-9M score reduced by 50% or greater six months (+/- 60 days) after an index visit.

NUMERATOR DETAILS

This PROM-PM outcome measure is longitudinal, seeking to measure improvement of depression symptoms with a PHQ-9 or PHQ-9M result reduced by 50% or greater (response) within six months (+/- 60 days) for the patients with an index event(depression and elevated PHQ-9 or PHQ-9M).

The numerator is defined as patients with a six-month (+/- 60 days) PHQ-9 or PHQ-9M score reduced by 50% or greater.

The numerator rate is calculated as follows:

pts with major depression or dysthymia with a PHQ-9 or PHQ-9M score reduced by 50% or greater at 6 months(+/- 60 days)/

pts with major depression or dysthymia with index contact PHQ-9 > 9

Patients who do not have a six month +/- 60 day PHQ-9 or PHQ-9M score obtained remain in the denominator and are counted as not having a response to treatment. Not having a PHQ-9 or PHQ-9M score within the 120 day window is considered a numerator miss.

Time period for data collection: there is a set index period for this measure, typically patients who have an index visit within a calendar period (e.g. index dates between 11/1/2017 and 10/31/2018) and then allowing enough time to pass to accommodate the timeframe for assessment. (e.g. for response at six months +/- 60 days with index dates of service ending 10/31/2018, the assessment period for twelve month remission and response [to also capture 12 month remission and response rates] would go through 12/30/2019). Technically, the six- and twelve-month remission and response measures are collected together in the MN program, and the index assessment period is fourteen months in duration.

Denominator identification period (index) 11/1/2017 to 10/31/2018

Measure assessment period through 12/30/2019; reported in 2020

DENOMINATOR STATEMENT

Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

DENOMINATOR DETAILS

The target population, patients age 12 and older with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine, is identified as follows:

Patients age 12 and older at the time of the index visit AND Index visit

An index visit occurs when ALL of the following criteria are met during a face-to-face visit or contact with an eligible provider:

- a PHQ-9 or PHQ-9M result greater than nine
- an active diagnosis of Major Depression or Dysthymia (Major Depression or Dysthymia Value Set)
- the patient is NOT in a prior index period

An index period begins with an index visit and is 14 months in duration.

Denominator is stratified by age range for adolescents (12 to 17 years of age) and adults (18 years of age and older).

Patients who do not have a six month +/- 60 day follow-up PHQ-9 or PHQ-9M score obtained remain in the denominator for this measure.

Please refer to the attached data dictionary for an inclusive list of all ICD-9/ICD-10 codes and data element definitions.

EXCLUSIONS

Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or

personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

EXCLUSION DETAILS

Required exclusions:

- Patient had a diagnosis of Bipolar Disorder (Bipolar Disorder Value Set) any time prior to the end of their measure assessment period
- Patient had an active diagnosis of Schizophrenia or Psychotic Disorder (Schizophrenia Psychotic Disorder Value Set) any time prior to the end of their measure assessment period

Allowable exclusions:

- Patient had an active diagnosis of Personality Disorder – Emotionally Labile Conditions (Personality Disorder – Emotionally Labile Value Set) any time prior to the end of their measurement assessment period
- Patient had an active diagnosis of Pervasive Developmental Disorder (Pervasive Disorder Value Set) any time prior to the end of the measurement assessment period
- Patient was a permanent nursing home resident at any time during the denominator identification period or measure assessment period
- Patient was in hospice or receiving palliative care (Palliative Care Value Set) at any time during the denominator identification or measure assessment period
- Patient died prior to the end of their measurement assessment period

The direct data submission process in MN allows for both up-front exclusions of the population and, because this is a longitudinal outcome measure, processes are in place to allow exclusions that may occur after index during the course of the measurement assessment period. Please see field specifications in the attached data dictionary.

RISK ADJUSTMENT

Yes - Additional risk adjustment analysis is included

Statistical risk model with risk factors (specify number of risk factors)

MNCM uses Logistic Regression Modeling to create values supporting a method of Indirect Standardization Risk Adjustment, commonly referred to as Expected Value. Indirect standardization does not change the actual performance rates, rather answers the question: “If all providers had this medical group/ clinic’s mix of patients, what would the statewide average be?”. This method compares the provider’s actual performance to the expected rate.

Example Clinic X	Unadjusted	Standardized to Clinic X Patient Mix
Statewide	39%	32%
Clinic X	35%	35%

Clinic X vs Statewide Below Above (Actual : Expected = 1.09

Risk variables used for this measure include age, initial PHQ-9/ PHQ-9M score, insurance product and patient neighborhood deprivation index (based on zip-code). Deprivation index includes use of SNAP benefits, living under the poverty level, unemployed status, public assistance, and single female with children. In MN, the ratio ranges are -6.41 (Red Lake) to +1.42 (Flom) with a mean of zero. "A measure of census-tract neighborhood deprivation is likely a good proxy for a range of individual-level and true area-level constructs relevant to outcomes of interest and feasible to obtain." [National Academies of Sciences, Engineering, and Medicine, 2017: Accounting for Social Risk Factors in Medicare Payment]

2021 Submission

6 Month Response- Adults

Analysis of Maximum Likelihood Estimates

Depression Response at 6 Months- Adults

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6416	0.0239	4727.3299	<.0001
pt_age	1	0.00866	0.000505	293.9879	<.0001
mdcr	1	-0.2636	0.0241	119.8347	<.0001
mhcp	1	-0.4468	0.0201	492.469	<.0001
unins	1	-0.5447	0.0461	139.6832	<.0001
undt	1	-0.3607	0.0257	196.7224	<.0001
mod_severe	1	0.0818	0.0167	24.1109	<.0001
severe	1	0.0775	0.0201	14.8663	0.0001
dep_idx	1	0.1399	0.0096	212.371	<.0001

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; Adults

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

6 Month Response- Adolescents

Analysis of Maximum Likelihood Estimates

Depression Response at 6 Months- Adolescents

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.5187	0.2586	4.022	0.0449
pt_age	1	-0.0772	0.0169	20.7373	<.0001
mdcr	1	0.0426	0.2112	0.0406	0.8403
mhcp	1	-0.2546	0.0629	16.3837	<.0001
unins	1	-0.4058	0.174	5.3997	0.0201
undt	1	-0.1813	0.0854	4.5033	0.0338
mod_severe	1	0.08	0.0586	1.8627	0.1723
severe	1	0.208	0.067	9.6377	0.0019
dep_idx	1	0.0445	0.0381	1.3673	0.2423

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; Adolescents

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

Definitions for Logistic Model

AIC – This is the Akaike Information Criterion. It is calculated as AIC = -2 Log L + 2((k-1) + s), where k is the number of levels of the dependent variable and s is the number of predictors in the model. AIC is used for the comparison of nonnested models on the same sample. Ultimately, the model with the smallest AIC is considered the best, although the AIC value itself is not meaningful.

SC – This is the Schwarz Criterion. It is defined as – 2 Log L + ((k-1) + s)*log(Σ fi), where fi's are the frequency values of the i^th observation, and k and s were defined previously. Like AIC, SC penalizes for the number of predictors in the model and the smallest SC is most desirable and the value itself is not meaningful.

-2 Log L – This is negative two times the log-likelihood. The -2 Log L is used in hypothesis tests for nested models and the value in itself is not meaningful.

Intercept Only – This column refers to the respective criterion statistics with no predictors in the model, i.e., just the response variable.

Intercept and Covariates – This column corresponds to the respective criterion statistics for the fitted model. A fitted model includes all independent variables and the intercept. We can compare the values in this column with the criteria corresponding Intercept Only value to assess model fit/significance.

Test – These are three asymptotically equivalent Chi-Square tests. They test against the null hypothesis that at least one of the predictors' regression coefficient is not equal to zero in the model. The difference between them are where on the log-likelihood function they are evaluated.

Likelihood Ratio – This is the Likelihood Ratio (LR) Chi-Square test that at least one of the predictors' regression coefficient is not equal to zero in the model. The LR Chi-Square statistic can be calculated by -2 Log L(null model) – 2 Log L(fitted model) = 231.289-160.236 = 71.05, where L(null model) refers to the Intercept Only model and L(fitted model) refers to the Intercept and Covariates model.

Score – This is the Score Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Wald – This is the Wald Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Chi-Square, DF and Pr > ChiSq – These are the Chi-Square test statistic, Degrees of Freedom (DF) and associated p-value (PR>ChiSq) corresponding to the specific test that all of the predictors are simultaneously equal to zero. We are testing the probability (PR>ChiSq) of observing a Chi-Square statistic as extreme as, or more so, than the observed one under the null hypothesis; the null hypothesis is that all of the regression coefficients in the model are equal to zero. The DF defines the distribution of the Chi-Square test statistics and is defined by the number of predictors in the model. Typically, PR>ChiSq is compared to a specified alpha level, our willingness to accept a type I error, which is often set at 0.05 or 0.01. The small p-value from all three tests would lead us to conclude that at least one of the regression coefficients in the model is not equal to zero.

STRATIFICATION

This measure is stratified by age range and results are reported separately by age: Adolescents (12-17 years of age) and Adults (18 years of age and older).

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

Measure Calculation Algorithms; Determining Depression Index and Calculation of Numerator

This measure is calculated by submitting a visit level file for the eligible patients. Each record in the file represents a contact with the patient and PHQ-9 or PHQ-9M score associated with this contact. Data files are submitted to a HIPAA secure data portal. Programming within the data portal determines the starting

point (index visit) and then calculates based on dates if a six month +/- 60 days PHQ-9 or PHQ-9M was obtained and the resulting score.

Calculation logic:

Is patient eligible for inclusion with diagnosis codes (Major Depression or Dysthymia Value Set)

and PHQ-9 or PHQ-9M > 9?

If yes, mark the visit as index (anchor) and include this patient in the denominator.

Does patient have a PHQ-9 or PHQ-9M score completed with a contact date that is six months +/- 60 days from the index date?

If yes, include this score to calculate rate. Programming logic includes the most recent score within the +/- 60 day window.

If no, patient is included in the denominator only. Not having a PHQ-9 or PHQ-9M score within the 120 day window is considered a numerator miss.

If the patient does have a six month +/- 60 day PHQ-9 or PHQ-9M and the score is it reduced by 50% or more from the index PHQ-9 or PHQ-9M score? [For example, a patient with an index PHQ-9/PHQ-9M score of 21 then at six months +/- 60 days has a most recent follow-up score of 9 would be considered a response and in the numerator]

If six month +/- 60 day PHQ-9 or PHQ-9M is reduced by 50% or greater; is considered a numerator case for rate calculation.

Reporting of this measure is currently at the clinic and medical group level.

Risk adjustment methodology uses individual patient level variables (age, insurance product depression severity level and zip code based deprivation index) to adjust for these variables at the clinic site and medical group practice level. Age is a continuous variable. Insurance product is Commercial, Medicare, Minnesota Health Care Plans (MHCP) and Cash or Uninsured patients. Depression severity level is based on the index PHQ-9 or PHQ-9M score, Moderate (PHQ9 below 15), Moderately Severe (PHQ9 15 to 19), Severe (PHQ9 over 19). The risk adjustment employs an actual to expected methodology where the actual measure result remains unaltered, instead a risk adjusted comparison is created based on same proportions of the risk factors that the clinic has. Our MNHealthscores website displays both the actual and expected rates in the detailed view.

COPYRIGHT / DISCLAIMER

N/A

NQF #1885 Depression Response at 12 Months – Progress Towards Remission

STEWARD

MN Community Measurement

DESCRIPTION

The percentage of adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia who are progressing towards remission by achieving a response (PHQ-9 or PHQ-9M score reduced by 50% or greater) twelve months (+/- 60 days) after an index visit.

TYPE

Outcome: PRO-PM

DATA SOURCE

Other, Electronic Health Records: Electronic Health Records, Electronic Health Records

The data source is the medical group's/clinic's medical record information, most frequently from am EMR. A CSV file is created by each medical group and uploaded to a password protected, HIPAA secure data portal which performs rate calculation. Selected Patient Reported Data, not because it is necessarily a separate data source, but because this measure is based on a patient reported outcome tool, a PRO-PM measure. Frequently this PRO tool, the PHQ-9, is housed within a clinic's EMR, or in paper charts is a part of the patient's medical record.

PROM

The PHQ-9 depression assessment tool is a patient reported outcome tool that is in the public domain and can be obtained for free use on the Patient Health Questionnaire (PHQ) Screeners website at www.phqscreeners.com. Modes of administration include traditional paper, mail, electronic and telephonic. The tool is available on the website with 79 language translations available.

The PHQ-9 tool is validated for use as a measure to assess the level of depression severity (for initial treatment decisions) as well as an outcome tool (to determine treatment response). [Löwe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004;42:1194-1201 and Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010]

The PHQ-9M is a modified version of the PHQ-9 tool for adolescents. Please refer to discussion in question sp.27

LEVEL

Clinician: Group/Practice

SETTING

Outpatient Services

NUMERATOR STATEMENT

The number of patients in the denominator who achieved a response as demonstrated by a PHQ-9 or PHQ-9M score reduced by 50% or greater twelve months (+/- 60 days) after an index visit.

NUMERATOR DETAILS

This PROM-PM outcome measure is longitudinal, seeking to measure improvement of depression symptoms with a PHQ-9 or PHQ-9M result reduced by 50% or greater (response) within twelve months (+/- 60 days) for the patients with an index event(depression and elevated PHQ-9 or PHQ-9M).

The numerator is defined as patients with a twelve-month (+/- 60 days) PHQ-9 or PHQ-9M score reduced by 50% or greater.

The numerator rate is calculated as follows:

pts with major depression or dysthymia with a PHQ-9 or PHQ-9M score reduced by 50% or greater at 12 months(+/- 60 days)/

pts with major depression or dysthymia with index contact PHQ-9 > 9

Patients who do not have a twelve month +/- 60 day PHQ-9 or PHQ-9M score obtained remain in the denominator and are counted as not having a response to treatment. Not having a PHQ-9 or PHQ-9M score within the 120 day window is considered a numerator miss.

Time period for data collection: there is a set index period for this measure, typically patients who have an index visit within a calendar period (e.g. index dates between 11/1/2017 and 10/31/2018) and then allowing enough time to pass to accommodate the timeframe for assessment. (e.g. for response at twelve months +/- 60 days with index dates of service ending 10/31/2018, the assessment period for twelve month remission and response [to also capture 12 month remission and response rates] would go through 12/30/2019). Technically, the six- and twelve-month remission and response measures are collected together in the MN program, and the index assessment period is fourteen months in duration.

Denominator identification period (index) 11/1/2017 to 10/31/2018

Measure assessment period through 12/30/2019; reported in 2020.

DENOMINATOR STATEMENT

Adolescent patients (12 to 17 years of age) and adult patients (18 years of age or older) with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine.

DENOMINATOR DETAILS

The target population, patients age 12 and older with major depression or dysthymia and an initial (index) PHQ-9 or PHQ-9M score greater than nine, is identified as follows:

Patients age 12 and older at the time of the index visit AND Index visit

An index visit occurs when ALL of the following criteria are met during a face-to-face visit or contact with an eligible provider:

- a PHQ-9 or PHQ-9M result greater than nine
- an active diagnosis of Major Depression or Dysthymia (Major Depression or Dysthymia Value Set)
- the patient is NOT in a prior index period

An index period begins with an index visit and is 14 months in duration.

Denominator is stratified by age range for adolescents (12 to 17 years of age) and adults (18 years of age and older).

Patients who do not have a twelve month +/- 60 day follow-up PHQ-9 or PHQ-9M score obtained remain in the denominator for this measure.

Please refer to the attached data dictionary for an inclusive list of all ICD-9/ICD-10 codes and data element definitions.

EXCLUSIONS

Patients who die, are a permanent resident of a nursing home or are enrolled in hospice are excluded from this measure. Additionally, patients who have a diagnosis of bipolar or personality disorder, schizophrenia or psychotic disorder, or pervasive developmental disorder are excluded.

EXCLUSION DETAILS

Required exclusions:

- Patient had a diagnosis of Bipolar Disorder (Bipolar Disorder Value Set) any time prior to the end of their measure assessment period
- Patient had an active diagnosis of Schizophrenia or Psychotic Disorder (Schizophrenia Psychotic Disorder Value Set) any time prior to the end of their measure assessment period

Allowable exclusions:

- Patient had an active diagnosis of Personality Disorder – Emotionally Labile Conditions (Personality Disorder – Emotionally Labile Value Set) any time prior to the end of their measurement assessment period

- Patient had an active diagnosis of Pervasive Developmental Disorder (Pervasive Disorder Value Set) any time prior to the end of the measurement assessment period
- Patient was a permanent nursing home resident at any time during the denominator identification period or measure assessment period
- Patient was in hospice or receiving palliative care (Palliative Care Value Set) at any time during the denominator identification or measure assessment period
- Patient died prior to the end of their measurement assessment period

The direct data submission process in MN allows for both up-front exclusions of the population and, because this is a longitudinal outcome measure, processes are in place to allow exclusions that may occur after index during the course of the measurement assessment period. Please see field specifications in the attached data dictionary.

RISK ADJUSTMENT

Yes - Additional risk adjustment analysis is included

Statistical risk model with risk factors (specify number of risk factors)

MNCM uses Logistic Regression Modeling to create values supporting a method of Indirect Standardization Risk Adjustment, commonly referred to as Expected Value. Indirect standardization does not change the actual performance rates, rather answers the question: "If all providers had this medical group/ clinic's mix of patients, what would the statewide average be?". This method compares the provider's actual performance to the expected rate.

Example Clinic X Unadjusted Standardized to Clinic X Patient Mix

Statewide 39% 32%

Clinic X 35% 35%

Clinic X vs Statewide Below Above (Actual : Expected = 1.09

Risk variables used for this measure include age, initial PHQ-9/ PHQ-9M score, insurance product and patient neighborhood deprivation index (based on zip-code). Deprivation index includes use of SNAP benefits, living under the poverty level, unemployed status, public assistance, and single female with children. In MN, the ratio ranges are -6.41 (Red Lake) to +1.42 (Flom) with a mean of zero. "A measure of census-tract neighborhood deprivation is likely a good proxy for a range of individual-level and true area-level constructs relevant to outcomes of interest and feasible to obtain." [National Academies of Sciences, Engineering, and Medicine, 2017: Accounting for Social Risk Factors in Medicare Payment]

2021 Submission

12 Month Response- Adults

Analysis of Maximum Likelihood Estimates

Depression Response at 12 Months- Adults

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.8846	0.0253	5560.51	<.0001
pt_age	1	0.0112	0.000531	440.8772	<.0001
mdcr	1	-0.2698	0.0251	115.9745	<.0001
mhcp	1	-0.4775	0.0215	495.1298	<.0001
unins	1	-0.4472	0.0473	89.4065	<.0001
undt	1	-0.4456	0.0278	256.3172	<.0001
mod_severe	1	0.0474	0.0175	7.3062	0.0069
severe	1	0.0425	0.0213	3.9953	0.0456
dep_idx	1	0.1376	0.0102	182.9268	<.0001

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; Adults

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

12 Month Response- Adolescents

Analysis of Maximum Likelihood Estimates

Depression Response at 12 Months- Adolescents

Compared to Patients with Commercial Insurance and Moderate Depression

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7016	0.266	6.9551	0.0084
pt_age	1	-0.0671	0.0174	14.8288	0.0001
mdcr	1	-0.5406	0.2661	4.126	0.0422
mhcp	1	-0.2282	0.0642	12.6212	0.0004
unins	1	-0.6443	0.1982	10.5702	0.0011
undt	1	-0.19	0.0879	4.6756	0.0306
mod_severe	1	-0.0108	0.0603	0.0323	0.8574
severe	1	0.1504	0.0687	4.8021	0.0284
dep_idx	1	0.1133	0.0405	7.829	0.0051

SAS Statistical Software Output Analysis of Variables Selected for Risk Adjustment; Adolescents

Table of results for data elements selected for the risk stratification model (age, insurance product, severity of depression at index event and deprivation index. All variables have a Chi-squared p value of less than .0001.

Definitions for Logistic Model

AIC – This is the Akaike Information Criterion. It is calculated as AIC = -2 Log L + 2((k-1) + s), where k is the number of levels of the dependent variable and s is the number of predictors in the model. AIC is used for the comparison of nonnested models on the same sample. Ultimately, the model with the smallest AIC is considered the best, although the AIC value itself is not meaningful.

SC – This is the Schwarz Criterion. It is defined as – 2 Log L + ((k-1) + s)*log(Σ fi), where fi's are the frequency values of the i^th observation, and k and s were defined previously. Like AIC, SC penalizes for the number of predictors in the model and the smallest SC is most desirable and the value itself is not meaningful.

-2 Log L – This is negative two times the log-likelihood. The -2 Log L is used in hypothesis tests for nested models and the value in itself is not meaningful.

Intercept Only – This column refers to the respective criterion statistics with no predictors in the model, i.e., just the response variable.

Intercept and Covariates – This column corresponds to the respective criterion statistics for the fitted model. A fitted model includes all independent variables and the intercept. We can compare the values in this column with the criteria corresponding Intercept Only value to assess model fit/significance.

Test – These are three asymptotically equivalent Chi-Square tests. They test against the null hypothesis that at least one of the predictors' regression coefficient is not equal to zero in the model. The difference between them are where on the log-likelihood function they are evaluated.

Likelihood Ratio – This is the Likelihood Ratio (LR) Chi-Square test that at least one of the predictors' regression coefficient is not equal to zero in the model. The LR Chi-Square statistic can be calculated by -2 Log L(null model) – 2 Log L(fitted model) = 231.289-160.236 = 71.05, where L(null model) refers to the Intercept Only model and L(fitted model) refers to the Intercept and Covariates model.

Score – This is the Score Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Wald – This is the Wald Chi-Square Test that at least one of the predictors' regression coefficient is not equal to zero in the model.

Chi-Square, DF and Pr > ChiSq – These are the Chi-Square test statistic, Degrees of Freedom (DF) and associated p-value (PR>ChiSq) corresponding to the specific test that all of the predictors

are simultaneously equal to zero. We are testing the probability (PR>ChiSq) of observing a Chi-Square statistic as extreme as, or more so, than the observed one under the null hypothesis; the null hypothesis is that all of the regression coefficients in the model are equal to zero. The DF defines the distribution of the Chi-Square test statistics and is defined by the number of predictors in the model. Typically, PR>ChiSq is compared to a specified alpha level, our willingness to accept a type I error, which is often set at 0.05 or 0.01. The small p-value from all three tests would lead us to conclude that at least one of the regression coefficients in the model is not equal to zero.

STRATIFICATION

This measure is stratified by age range and results are reported separately by age: Adolescents (12-17 years of age) and Adults (18 years of age and older).

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

Measure Calculation Algorithms; Determining Depression Index and Calculation of Numerator

This measure is calculated by submitting a visit level file for the eligible patients. Each record in the file represents a contact with the patient and PHQ-9 or PHQ-9M score associated with this contact. Data files are submitted to a HIPAA secure data portal. Programming within the data portal determines the starting point (index visit) and then calculates based on dates if a twelve month +/- 60 days PHQ-9 or PHQ-9M was obtained and the resulting score.

Calculation logic:

Is patient eligible for inclusion with diagnosis codes (Major Depression or Dysthymia Value Set)

and PHQ-9 or PHQ-9M > 9?

If yes, mark the visit as index (anchor) and include this patient in the denominator.

Does patient have a PHQ-9 or PHQ-9M score completed with a contact date that is twelve months +/- 60 days from the index date?

If yes, include this score to calculate rate. Programming logic includes the most recent score within the +/- 60 day window.

If no, patient is included in the denominator only. Not having a PHQ-9 or PHQ-9M score within the 120 day window is considered a numerator miss.

If the patient does have a twelve month +/- 60 day PHQ-9 or PHQ-9M and the score is it reduced by 50% or more from the index PHQ-9 or PHQ-9M score? [For example, a patient with an index PHQ-9/PHQ-9M score of 21 then at twelve months +/- 60 days has a most recent follow-up score of 9 would be considered a response and in the numerator]

If twelve month +/- 60 day PHQ-9 or PHQ-9M is reduced by 50% or greater; is considered a numerator case for rate calculation.

Reporting of this measure is currently at the clinic and medical group level.

Risk adjustment methodology uses individual patient level variables (age, insurance product depression severity level and zip code based deprivation index) to adjust for these variables at the clinic site and medical group practice level. Age is a continuous variable. Insurance product is Commercial, Medicare, Minnesota Health Care Plans (MHCP) and Cash or Uninsured patients. Depression severity level is based on the index PHQ-9 or PHQ-9M score, Moderate (PHQ9 below 15), Moderately Severe (PHQ9 15 to 19), Severe (PHQ9 over 19). The risk adjustment employs an actual to expected methodology where the actual measure result remains unaltered, instead a risk adjusted comparison is created based on same proportions of the risk factors that the clinic has. Our MNHealthscores website displays both the actual and expected rates in the detailed view.

COPYRIGHT / DISCLAIMER

NQF #3312 Continuity of Care After Medically Managed Withdrawal From Alcohol and/or Drugs

STEWARD

Centers for Medicare & Medicaid Services

DESCRIPTION

Percentage of discharges from a medically managed withdrawal episode for adult Medicaid Beneficiaries, age 18-64, that was followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder (pharmacotherapy) within 7 or 14 days after discharge). This measure is reported across all medically managed withdrawal settings.

TYPE

Process

DATA SOURCE

Claims

We used the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs) Research Identifiable Files (RIFs) which contains beneficiary, service utilization, administrative claims, and expenditure data for the Medicaid population, including those covered through both fee-for-service (FFS) and managed care payers. The Medicaid T-MSIS TAF RIFs for calendar year 2018 were used to assess importance, reliability and validity: Demographic and Eligibility (DE) file; Inpatient (IP) file; Other Services (OT) file; Long-Term Care (LT) file; and Pharmacy (RX) file. For beneficiaries dually enrolled in Medicare [Medicare Advantage (MA) or Medicare Fee-For-Service (FFS)], in addition to the T-MSIS TAF RIFs, we also used Medicare Part A, B, C and D administrative claims from the Center for Medicare and Medicaid's Chronic Conditions Warehouse (CCW): MA encounter and Medicare FFS carrier, outpatient (OP), inpatient (IP), skilled nursing facility (SNF), MedPAR and prescription drug event (PDE) files.

Measure validity was assessed through NCQA Healthcare Effectiveness Data and Information Set (HEDIS) data from measure year 2018. HEDIS data are collected from Medicaid Health Management Organizations and Preferred Provider Organizations via the NCQA Interactive Data Submission System (IDSS) portal.

LEVEL

Population: Regional and State, Population: Population

SETTING

Inpatient/Hospital, Outpatient Services

NUMERATOR STATEMENT

Discharges in the denominator who have an inpatient, intensive outpatient, partial hospitalization, outpatient visit, residential, or drug prescription or procedure within 7 or 14 days after discharge from an inpatient hospital, residential addiction program, or ambulatory medically managed withdrawal.

NUMERATOR DETAILS

The numerator includes individuals with any of the following within 14 days after discharge from medically managed withdrawal:

- Pharmacotherapy on day of discharge through day 7 or 14.
- Outpatient, intensive outpatient, partial hospitalization, or residential treatment procedure with a diagnosis of SUD on the day after discharge through day 7 or 14.

- Outpatient, intensive outpatient, partial hospitalization, or residential treatment with standalone SUD procedure on the day after discharge through day 7 or 14.
- Inpatient admission with an SUD diagnosis or procedure code on day after discharge through day 7 or 14.
- Long-term care institutional claims with an SUD diagnosis on day after discharge through day 7 or 14.

If an overdose diagnosis code appears on the same outpatient or inpatient claim that is being viewed as follow-up, that claim does not qualify as follow-up.

SUD diagnoses are used to identify procedures connected to SUD diagnoses. SUD diagnoses are identified through ICD-10 codes (available in the attached value set: NQF 3312—Tab 3). Procedures are defined using a combination of Healthcare Common Procedure Coding System (HCPCS) codes, Uniform Billing (UB) Revenue Codes and ICD-10 procedure codes (available in the attached value sets: NQF 3312—Tabs 4–8).

Pharmacotherapy includes naltrexone (short or long acting), acamprosate, or disulfiram for alcohol dependence treatment and buprenorphine for opioid dependence treatment, as well HCPCS codes to identify procedures related to injecting drugs (e.g., long-acting injectable naltrexone) (available in the attached value sets: NQF 3312—Tabs 9–10).

Code lists for this measure are in the attached value sets. States may need to adapt the list of codes to include state-specific codes.

DENOMINATOR STATEMENT

Adult Medicaid beneficiary discharges from medically managed withdrawal from January 1 to December 15 of the measurement year.

DENOMINATOR DETAILS

Measure data is reported annually (12 months). To account for the 14-day time period after discharge from medically managed withdrawal, the denominator period starts January 1 and ends December 15 of the measurement year.

Eligible population meets the following conditions:

- Medicaid beneficiaries aged 18-64 with at least one discharge from medically managed withdrawal during the year January 1 to December 15.
- Enrolled in Medicaid during the month of discharge medically managed withdrawal and the following month.

The denominator is based on discharges, not individual beneficiaries. A beneficiary may have more than one qualifying medically managed withdrawal episode.

The location of medically managed withdrawal can include hospital inpatient, inpatient residential addiction, other stayover treatment, and ambulatory medically managed withdrawal.

Medically managed withdrawal is identified using a combination of HCPCS codes, UB Revenue Codes, and ICD-10 procedure codes. A list of codes to identify medically managed withdrawal is posted in the value sets: Table NQF 3312—Tabs 1–2. States will likely need to modify the specifications to include their state-specific codes.

EXCLUSIONS

Not applicable: the measure does not have denominator exclusions.

EXCLUSION DETAILS

Not applicable: the measure does not have denominator exclusions.

RISK ADJUSTMENT

No additional risk adjustment analysis included

No risk adjustment or stratification

STRATIFICATION

Not applicable

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

The following step are used to identify the denominator, numerator, and calculation of the measure rate:

Step 1: Identify denominator

- Step 1A: Eligible population: Identify Medicaid beneficiaries ages 18–64 who have any medically managed withdrawal in inpatient hospital, residential addiction treatment program, or ambulatory medically managed withdrawal from January 1 to December 15 of the measurement year and are enrolled the month of medically managed withdrawal and the following month. Age is calculated as of January 1 of the measurement year.
- Step 1B: Among the Medicaid beneficiaries in Step 1A, identify all discharges from medically managed withdrawal using all inpatient, outpatient, and ambulatory claims files or tables that contain HCPCS or ICD-10 procedure codes and UB revenue codes (see NQF 3312 – Tab 1-2 for code lists). If more than one discharge from medically managed withdrawal in a year, treat each discharge from medically managed withdrawal as a separate episode, e.g., an inpatient hospital medically managed withdrawal in January and an ambulatory medically managed withdrawal in July counts as two episodes.

- Step 1B.1: Multiple medically managed withdrawal claims that are up to 2 days apart are combined into a single episode. Sort the inpatient, outpatient, and ambulatory discharge from medically managed withdrawals by Beneficiary ID and service dates to ensure the discharges from these multiple data sources are in chronological order. Then combine claims that are up to 2 days apart while retaining all clinical fields from each episode.

- Step 1C: Identify appropriate location of medically managed withdrawal services: hospital inpatient, inpatient residential addiction, outpatient residential outpatient addiction, other stayover treatment, and ambulatory medically managed withdrawal. Use HCPCS medically managed withdrawal procedure codes to assign medically managed withdrawal location whenever possible; revenue center medically managed withdrawal will map to the hospital inpatient location when the revenue codes appear on an inpatient claim or table (see attached value set: NQF 3312 – Tab 2). Revenue center medically managed withdrawal will map to other stayover treatment when the revenue codes appear on a non-inpatient claim. If there is more than 1 medically managed withdrawal location when episodes are combined, assign the location using the first claim´s location. If there is a tie between a medically managed withdrawal episode being identified via revenue center codes and a more specific category using HCPCS on the same claim, the HCPCS location prevails.

Step 2: Identify numerator

- Step 2A: From the denominator in Step 1B, identify those discharges from medically managed withdrawal in any setting with a qualifying continuity service within 7 or 14 days after discharge.
  - Step 2A.1: Identify SUD continuity services: Continuity services are assigned using clinical claims billing information (e.g., diagnosis, procedure, revenue codes; see attached value sets NQF 3312 – Tab 2-8). The measure includes all claims files or data tables that contain clinical fields (e.g., inpatient hospital, outpatient, other ambulatory, and long-term care). SUD diagnoses can be in any position – primary or secondary – for continuity services. Since multiple claims files or tables could each contain a continuity claim, this calls for creating continuity variables separately within each file type or table, sorting the files or tables by beneficiary ID and service dates, then putting them together in order to assign the set of variables that are "First" to occur relative to the medically managed withdrawal episode discharge date. Continuity services have to occur the day after discharge through day 7 or 14.
  - Step 2A.2: Identify pharmacotherapy which may occur in multiple files or tables (see attached value sets: NQF 3312 – Tab 9-10). For example, one claims file or data source may contain injectables, another claims file or table data source may contain oral medications. Consequently, pharmacotherapy variables are created separately in each source, the data sources are then sorted by beneficiary ID and service dates, then multiple pharmacotherapy data sources are put together so they will be in chronological order to assign "First" variables. Pharmacotherapy

services could be provided on the same day as the discharge from medically managed withdrawal through day 7 or 14.

- Step 2A.3: Co-occurring events: Emergency department visits, even with an SUD diagnosis, do not count as continuity. Also, other continuity services, e.g., an outpatient visit that occur on the same day as an emergency department visit with an SUD diagnosis do not count as continuity. If an overdose diagnosis code appears on the same claim as the continuity service, then the service does not count as continuity. If an inpatient continuity claim has an emergency department visit meaning that the beneficiary was admitted through the emergency department, it is allowed to remain a continuity service.

Step 3: Calculate rate

- Step 3A: Calculate the overall 7- or 14-day continuity rates by dividing the number of discharges with a qualifying continuity service (Step 2A) by the denominator (Step 1B).
- Step 3B: Calculate the rates separately for each medically managed withdrawal location by dividing the respective number of discharges by each location with a qualifying continuity service (Step 2A) by the denominator (Step 1C).

COPYRIGHT / DISCLAIMER

NQF #3313 Follow-Up Care for Adult Medicaid Beneficiaries Who Are Newly Prescribed an Antipsychotic Medication

STEWARD

Centers for Medicare & Medicaid Services

DESCRIPTION

Percentage of new antipsychotic prescriptions for Medicaid beneficiaries age 18 years and older who have completed a follow-up visit with a provider with prescribing authority within four weeks (28 days) of prescription of an antipsychotic medication.

TYPE

Process

DATA SOURCE

Claims

We used the Transformed Medicaid Statistical Information System (T-MSIS) Analytic Files (TAFs) Research Identifiable Files (RIFs) which contains beneficiary, service utilization, administrative claims, and expenditure data for the Medicaid population, including those covered through both fee-for-service (FFS) and managed care payers. The Medicaid T-MSIS TAF RIFs from September 1, 2017 through December 31, 2018 were used to assess importance, reliability and validity: Demographic and Eligibility (DE), Other Services (OT), Inpatient (IP) and Pharmacy (RX). For beneficiaries dually enrolled in Medicare [Medicare Advantage (MA) or Medicare Fee-For-Service (FFS)], in addition to the T-MSIS TAF RIFs data, we also used Medicare Part A, B, C and D administrative claims (September 1, 2017 through December 31, 2018) from the Center for Medicare and Medicaid's Chronic Conditions Warehouse (CCW): MA encounter and Medicare FFS carrier, Outpatient (OP), Inpatient (IP), MedPAR and Prescription Drug Event (PDE) files.

Measure validity was assessed by conducting a convergent validity analysis using NCQA Healthcare Effectiveness Data and Information Set (HEDIS) data from measure year 2018. HEDIS data are collected from Medicaid Health Management Organizations and Preferred Provider Organizations via the NCQA Interactive Data Submission System (IDSS) portal.

LEVEL

Population: Regional and State

SETTING

Outpatient Services

NUMERATOR STATEMENT

The percentage of Medicaid beneficiaries aged 18 years and older with new antipsychotic prescriptions who completed an outpatient follow-up visit with a provider with prescribing authority within 28 days of the new antipsychotic prescription fill.

NUMERATOR DETAILS

The numerator uses a 28-day (4 week) follow-up period based on clinical guidelines for appropriate follow-up after prescription of new antipsychotic medications. The optimal follow-up period was determined through testing and consultation with the Clinical Advisory Work group. The day after the prescription is filled is counted as day 1 of the follow-up period. The date of the follow-up visit with a provider is determined by using the service date on the medical claim.

See attached Excel file for CPT and HCPCS codes that qualify for the numerator.

DENOMINATOR STATEMENT

New antipsychotic prescriptions for Medicaid beneficiaries aged 18 years and older.

DENOMINATOR DETAILS

Target population meets the following conditions:

1. Medicaid beneficiary aged 18 years and older (including dual-eligible and Medicaid-only enrollees)
2. Newly prescribed an antipsychotic medication
3. Enrolled in Medicaid during the 120 days (four months) prior to and the 28 days (4 weeks) following a new prescription of an antipsychotic medication

Beneficiaries with "newly filled prescriptions" are those who have had no antipsychotic medications dispensed for either new or refill antipsychotic prescription during a period of 120 days (four months) prior to the prescription fill date. The measure focuses on new prescriptions of antipsychotic medications.

We used National Drug Codes to identify the following antipsychotic medications for this measure:

- Amitriptyline Hydrochloride
- Aripiprazole (Abilify)
- Aripiprazole lauroxil (Aristada)
- Asenapine Maleate (Saphris)
- Brexpiprazole (Rexulti)
- Cariprazine hydrochloride (Vraylar)
- Chlorpromazine hydrochloride
- Clozapine (Clozaril, Fazaclo, Versacloz)
- Droperidol (Inapsine)
- Fentanyl citrate/droperidol
- Fluoxetine Hydrochloride-Olanzapine (Symbyax, Fluoxetine, Sarafem, Selfaemra)
- Fluoxetine-olanzapine
- Fluphenazine
- Haloperidol (Haldol)
- Iloperidone (Fanapt)
- Loxapine succinate (Loxitane, Adasuve)
- Lurisadone hydrochloride (Latuda)

- Molindone hydrochloride (Moban)
- Olanzapine (Zyprexa)
- Paliperidone (Invega)
- Perphenazine
- Pimozide (Orap)
- Prochlorperazine maleate (Compazine, Compro)
- Quetiapine fumarate (Seroquel)
- Risperidone (Risperdal)
- Thioridazine hydrochloride
- Thiothixene (Navane)
- Trifluoperazine hydrochloride
- Ziprasidone (Geodon)

See attached Excel file for NDCs that qualify for the denominator.

EXCLUSIONS

- Medicaid beneficiaries with an acute inpatient admission during the 28-day follow-up period after prescription of an antipsychotic medication
- Patients who expired within 28 days of new antipsychotic prescription fill date.

EXCLUSION DETAILS

Acute inpatient admission during the 28-day follow-up period: Beneficiaries with an inpatient admission during the 28-day follow-up period are excluded from the measure.

Death: Patients with a date of death during the 28-day follow-up period are excluded from the measure.

RISK ADJUSTMENT

No additional risk adjustment analysis included

No risk adjustment or stratification

STRATIFICATION

Not applicable; this measure is not stratified.

TYPE SCORE

Rate/proportion

Better quality = Higher score

ALGORITHM

Step 1: Identify Eligible Population

Step 1A. Identify Medicaid beneficiaries (both dual-eligible and Medicaid-only enrollees) aged 18 years and older.

Step 1B. From this group, identify new prescriptions of one or more antipsychotic medications.

Step 2: Apply Continuous Enrollment Requirement

From the population identified in step 1:

Step 2A. Remove any prescriptions for beneficiaries who were not continuously enrolled for at least 120 days before or 28 days following the new prescription.

Step 3: Apply Exclusions:

From the population identified in step 2:

Step 3A. Remove any prescriptions for beneficiaries who had an acute inpatient admission during the 28 days following the new prescription.

Step 3B. Remove any prescriptions for beneficiaries who expired during the 28 days following the new prescription.

Step 4: Numerator

From the prescriptions within the denominator (after denominator exclusions have been applied):

Step 4A. Identify the number of prescriptions for beneficiaries who had a qualifying outpatient encounter within 28 days of the prescription date of the antipsychotic medication.

Step 5: Calculate the measure score

Step 5A. Divide the total number of prescriptions in the numerator by the total number of prescriptions in the denominator, after denominator exclusions have been applied.

Step 5B. Multiply this number by 100 to determine the performance rate.

COPYRIGHT / DISCLAIMER

Appendix E: Related and Competing Measures

Comparison of NQF #3312 and NQF #0004

STEWARD/DEVELOPER

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Centers for Medicare & Medicaid Services

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

National Committee for Quality Assurance

DESCRIPTION

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Percentage of discharges from a medically managed withdrawal episode for adult Medicaid Beneficiaries, age 18-64, that was followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder (pharmacotherapy) within 7 or 14 days after discharge). This measure is reported across all medically managed withdrawal settings.

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

This measure assesses the degree to which the organization initiates and engages members identified with a need for alcohol and other drug (AOD) abuse and dependence services and the degree to which members initiate and continue treatment once the need has been identified. Two rates are reported:

- Initiation of AOD Treatment. The percentage of adolescent and adult members with a new episode of AOD abuse or dependence who initiate treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter, partial hospitalization, telehealth or medication assisted treatment (MAT) within 14 days of the diagnosis.
- Engagement of AOD Treatment. The percentage of adolescent and adult members with a new episode of AOD abuse or dependence who initiated treatment and who had two or more additional AOD services or MAT within 34 days of the initiation visit.

NUMERATOR

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Discharges in the denominator who have an inpatient, intensive outpatient, partial hospitalization, outpatient visit, residential, or drug prescription or procedure within 7 or 14 days after discharge from an inpatient hospital, residential addiction program, or ambulatory medically managed withdrawal.

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Initiation of AOD Treatment:

Initiation of treatment through an inpatient AOD admission, outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth or medication treatment within 14 days of the diagnosis. ---- Engagement of AOD Treatment: Initiation of AOD treatment and two or more additional AOD services or medication treatment within 34 days of the initiation visit.

DENOMINATOR

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Adult Medicaid beneficiary discharges from medically managed withdrawal from January 1 to December 15 of the measurement year.

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Patients age 13 years of age and older as of December 31 of the measurement year who were diagnosed with a new episode of alcohol or other drug dependency (AOD) during the first 10 and ½ months of the measurement year (e.g., January 1-November 15).

MEASURE TYPE

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Process

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Process

DATA SOURCE

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Claims

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Claims

TARGET POPULATION

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Populations at Risk: Populations at Risk, Adults (Age >= 18), Populations at Risk: Dual eligible beneficiaries of Medicare and Medicaid

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Elderly, Populations at Risk, Children

CARE SETTING

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Inpatient/Hospital, Outpatient Services

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Outpatient Services, Inpatient/Hospital, Emergency Department and Services

LEVEL OF ANALYSIS

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Population: Regional and State, Population: Population

NQF #0004 INITIATION AND ENGAGEMENT OF ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE TREATMENT

Health Plan

Comparison of NQF #3312 and NQF #2605

STEWARD/DEVELOPER

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Centers for Medicare & Medicaid Services

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

National Committee for Quality Assurance

DESCRIPTION

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Percentage of discharges from a medically managed withdrawal episode for adult Medicaid Beneficiaries, age 18-64, that was followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder (pharmacotherapy) within 7 or 14 days after discharge). This measure is reported across all medically managed withdrawal settings.

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

The percentage of discharges for patients 18 years of age and older who had a visit to the emergency department with a primary diagnosis of mental health or alcohol or other drug dependence during the measurement year AND who had a follow-up visit with any provider with a corresponding primary diagnosis of mental health or alcohol or other drug dependence within 7- and 30-days of discharge. Four rates are reported:

- The percentage of emergency department visits for mental health for which the patient received follow-up within 7 days of discharge.
- The percentage of emergency department visits for mental health for which the patient received follow-up within 30 days of discharge.
- The percentage of emergency department visits for alcohol or other drug dependence for which the patient received follow-up within 7 days of discharge.
- The percentage of emergency department visits for alcohol or other drug dependence for which the patient received follow-up within 30 days of discharge.

NUMERATOR

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Discharges in the denominator who have an inpatient, intensive outpatient, partial hospitalization, outpatient visit, residential, or drug prescription or procedure within 7 or 14 days after discharge from an inpatient hospital, residential addiction program, or ambulatory medically managed withdrawal.

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

THE NUMERATOR FOR EACH DENOMINATOR POPULATION CONSISTS OF TWO RATES:

MENTAL HEALTH

Rate 1: an outpatient visit, intensive outpatient encounter or partial hospitalization with any provider with a primary diagnosis of mental health within 7 days after emergency department discharge

Rate 2: an outpatient visit, intensive outpatient encounter or partial hospitalization with any provider with a primary diagnosis of mental health within 30 days after emergency department discharge

ALCOHOL OR OTHER DRUG DEPENDENCE

Rate 1: an outpatient visit, intensive outpatient encounter or partial hospitalization with any provider with a primary diagnosis of alcohol or other drug dependence within 7 days after emergency department discharge

Rate 2: an outpatient visit, intensive outpatient encounter or partial hospitalization with any provider with a primary diagnosis of alcohol or other drug dependence within 30 days after emergency department discharge

DENOMINATOR

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Adult Medicaid beneficiary discharges from medically managed withdrawal from January 1 to December 15 of the measurement year.

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

Patients who were treated and discharged from an emergency department with a primary diagnosis of mental health or alcohol or other drug dependence on or between January 1 and December 1 of the measurement year.

MEASURE TYPE

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Process

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

Process

DATA SOURCE

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Claims

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

Claims

TARGET POPULATION

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Populations at Risk: Populations at Risk, Adults (Age >= 18), Populations at Risk: Dual eligible beneficiaries of Medicare and Medicaid

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

Populations at Risk

CARE SETTING

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Inpatient/Hospital, Outpatient Services

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

Inpatient/Hospital, Outpatient Services

LEVEL OF ANALYSIS

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Population: Regional and State, Population: Population

NQF #2605 FOLLOW-UP AFTER EMERGENCY DEPARTMENT VISIT FOR MENTAL ILLNESS OR ALCOHOL AND OTHER DRUG ABUSE OR DEPENDENCE

Health Plan, Population: Regional and State

Comparison of NQF #3312 and NQF #3453

STEWARD/DEVELOPER

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Centers for Medicare & Medicaid Services

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Centers for Medicare & Medicaid Services

DESCRIPTION

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Percentage of discharges from a medically managed withdrawal episode for adult Medicaid Beneficiaries, age 18-64, that was followed by a treatment service for substance use disorder (including the prescription or receipt of a medication to treat a substance use disorder (pharmacotherapy) within 7 or 14 days after discharge). This measure is reported across all medically managed withdrawal settings.

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Percentage of discharges from an inpatient or residential treatment for substance use disorder (SUD) for Medicaid beneficiaries, ages 18 to 64, which was followed by a treatment service for SUD. SUD treatment includes having an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth encounter, or filling a prescription or being administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

NUMERATOR

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Discharges in the denominator who have an inpatient, intensive outpatient, partial hospitalization, outpatient visit, residential, or drug prescription or procedure within 7 or 14 days after discharge from an inpatient hospital, residential addiction program, or ambulatory medically managed withdrawal.

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Discharges in the denominator with an outpatient visit, intensive outpatient encounter or partial hospitalization, telehealth or filled a prescription for or were administered or ordered a medication for SUD. (After an inpatient discharge only, residential treatment also counts as continuity of care.) Two rates are reported, continuity within 7 and 14 days after discharge.

DENOMINATOR

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Adult Medicaid beneficiary discharges from medically managed withdrawal from January 1 to December 15 of the measurement year.

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Adult Medicaid beneficiary discharges from inpatient or residential treatment for SUD with a principal diagnosis of SUD during from January 1 to December 15 of the measurement year.

MEASURE TYPE

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Process

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Process

DATA SOURCE

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Claims

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Claims

TARGET POPULATION

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Populations at Risk: Populations at Risk, Adults (Age >= 18), Populations at Risk: Dual eligible beneficiaries of Medicare and Medicaid

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

N/A

CARE SETTING

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Inpatient/Hospital, Outpatient Services

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Emergency Department and Services, Home Care, Inpatient/Hospital, Outpatient Services

LEVEL OF ANALYSIS

NQF #3312 CONTINUITY OF CARE AFTER MEDICALLY MANAGED WITHDRAWAL FROM ALCOHOL AND/OR DRUGS

Population: Regional and State, Population: Population

NQF #3453 CONTINUITY OF CARE AFTER INPATIENT OR RESIDENTIAL TREATMENT FOR SUBSTANCE USE DISORDER (SUD)

Population: Regional and State

Appendix F: Pre-Evaluation Comments

Comments received as of June 15, 2022.

NQF #0710e Depression Remission at 12 Months

Commenter

Collette Cole, on behalf of MN Community Measurement

Comment

Hello, During the process of submitting our scientific testing for this measure NQF# 0710e Depression Remission at Twelve Months, we inadvertently did not include the c-statistic for this measure. This statistic was calculated during the logistic regression procedure but the clinical staff completing the application did not recognize the c-statistic in part due to the large number of pairs and the spacing of the table. The calculated concordance (c-statistic) for this measure was 0.616 (adults) and 0.592 (adolescents) which meet the criteria for a well calibrated model. Association of Predicted Probabilities and Observed Responses Adults Percent Concordant 61.6Somers' D 0.233 Percent Discordant 38.4Gamma 0.233 Percent Tied 0.0Tau-a 0.042 Pairs 1319452743c 0.616 Association of Predicted Probabilities and Observed Responses Adolescents Percent Concordant 59.1Somers' D 0.183 Percent Discordant 40.8Gamma 0.183 Percent Tied 0.0Tau-a 0.026 Pairs 9810185c 0.592 Please consider this additional information in the standing committee's assessment of the risk adjustment model. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer, MN Community Measurement

NQF #0710e Depression Remission at 12 Months

Commenter

Koryn Rubin, on behalf of American Medical Association

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to express our concerns on the evidence and testing provided in support of this measure. While the AMA agrees that it is useful to understand the rate of remissions for individuals diagnosed with depression, we do not believe that the developer provided sufficient evidence demonstrating that remission can be successfully achieved across the defined patient population within a twelve-month timeframe. It is important that the data demonstrate that practices can implement structures or processes that lead to improved outcomes and the measure results in rates that truly reflect the quality of care delivered by a practice rather than differences in patient mix or other factors outside of the practice's control. We also seek clarification on whether this measure has met all of the requirements for electronic clinical quality measures (eCQMs) since the National Quality Forum measure evaluation criteria require patient-encounter-level validity testing. We did not see any information that would satisfy this requirement. We also were unable to find the feasibility scorecard results and therefore were unable to fully evaluate this criterion. The AMA requests that these gaps in evidence and testing be

addressed prior to continued endorsement of this measure. We appreciate the Committee's consideration of our comments.

NQF #0711 Depression Remission at Six Months

Commenter

Collette Cole, on behalf of MN Community Measurement

Comment

Hello, During the process of submitting our scientific testing for this measure NQF# 0711 Depression Remission at Six Months, we inadvertently did not include the c-statistic for this measure. This statistic was calculated during the logistic regression procedure but the clinical staff completing the application did not recognize the c-statistic in part due to the large number of pairs and the spacing of the table. The calculated concordance (c-statistic) for this measure was 0.608 (adults) and 0.595 (adolescents) which meet the criteria for a well calibrated model. Association of Predicted Probabilities and Observed Responses Adults Percent Concordant 60.8Somers' D 0.217 Percent Discordant 39.2Gamma 0.217 Percent Tied 0.0Tau-a 0.043 Pairs 1449575559c 0.608 Association of Predicted Probabilities and Observed Responses Adolescents Percent Concordant 59.5Somers' D 0.190 Percent Discordant 40.5Gamma 0.190 Percent Tied 0.0Tau-a 0.028 Pairs 10006425c 0.595 Please consider this additional information in the standing committee's assessment of the risk adjustment model. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer, MN Community Measurement

NQF #0711 Depression Remission at Six Months

Commenter

Koryn Rubin, on behalf of American Medical Association

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to express our concerns and request clarification on one item with this measure. While the AMA agrees that it is useful to understand the rate of remissions for individuals diagnosed with depression, we do not believe that the developer provided sufficient evidence demonstrating that remission can be successfully achieved across the defined patient population within a six-month timeframe. It is important that the data demonstrate that practices can implement structures or processes that lead to improved outcomes and the measure results in rates that truly reflect the quality of care delivered by a practice rather than differences in patient mix or other factors outside of the practice's control. We also seek clarification on whether this measure is intended to be captured as an electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing and missing feasibility scorecard for NQF #710e would also apply to this measure. The AMA requests that the gaps in evidence and clarification on whether the measure

is intended to be an eCQM be addressed prior to continued endorsement of this measure. We appreciate the Committee's consideration of our comments.

NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M

Commenter

Collette Cole, on behalf of MN Community Measurement

Comment

Hello, MN Community Measurement is submitting this comment in response to NQF staff feedback about insufficient evidence for this measure #0712 Depression Assessment with the PHQ-9/PHQ9-M. It was noted that there was no empirical evidence to demonstrate that performing the PHQ-9 in and of itself in isolation is associated with any improvement in outcomes. In terms of measurement and assessing outcomes, frequent and ongoing assessment with the PHQ-9/PHQ-9M is key to understanding the patient's progress towards the reduction of depression symptoms. Administering the PHQ-9 is like taking a blood pressure- you need to do something with the information to affect the outcome of hypertension. Depression is now being considered the sixth vital sign by many and assessing patients is critical to identifying depression and improving outcomes. [Trivedi, M., Jha, M. et al VitalSign6: a Primary Care First (PCP-First) Model for Universal Screening and Measurement-Based Care for Depression. Pharmaceuticals 2019, 12, 71; doi:10.3390/ph12020071] Supportive evidence was provided in the context of the Institute for Clinical Systems Improvement Depression Care guidelines for 1) Comprehensive Treatment Plan with Shared Decision Making- Collaborative Care Model and 2) Establish Follow-up Plan. In addition, PubMed lists 14,699 studies associated with the use of the PHQ-9 for measuring or monitoring depression including a 2021 meta-analysis that supports the use of this tool to determine outcomes [Negeri, Z.F., Levis, B., Sun, Y. et al Accuracy of the Patient Health Questionnaire-9 for screening to detect major depression: updated systemic review and individual participant data meta-analysis BMJ 2021 Oct 5;375:n2183. Doi:10.1136/bmj.n.2183]. This measure is an important companion to outcome measures of response and remission, serving two purposes: the first is to understand how well a practice does at assessing their patients who have a diagnosis of depression, and the second is to guard against gaming of the outcome measures through selective administration of the PHQ-9. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer, MN Community Measurement

NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M

Commenter

Koryn Rubin, on behalf of American Medical Association

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to request clarification on one item with this measure. We seek clarification on whether this measure is intended to be captured as an electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed

that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing and missing feasibility scorecard for NQF #710e would also apply to this measure. The AMA requests clarification on whether the measure is intended to be an eCQM be addressed prior to continued endorsement of this measure. We appreciate the Committee's consideration of our comments.

NQF #0712 Depression Assessment With PHQ-9/ PHQ-9M

Commenter

Steven Inman, on behalf of Children's Health Network of Minnesota

Comment

Dear NQF: I represent Children's Health Network, the Network affiliated with Children's Hospitals and Clinics of Minnesota, and I support the re-endorsement of the suite of depression measures currently under review by the Behavioral Health Standing Committee. We have been using these measures for many years at our organization to understand and support positive care and outcomes for patients with depression. Additionally we have pay-for-performance contracts with insurance payers and recognition programs that utilize the rates for these measures. Using the PHQ-9 helps clinics in screening, diagnosing and ongoing monitoring of symptoms of depression. Our organization has increased focus on depression care and value these measures that support our focus. The outcome measures, work together in measuring outcomes at multiple points in time using the same information to measure remission or progress towards remission. The use of this measures on a statewide basis in Minnesota helps to focus attention on these outcomes for an important health problem that impacts many people. I support the continued endorsement of all five measures (NQF#s 0710e, 0711, 0712, 1884 and 1885). Respectfully; Steven Inman, MD Pediatrician Medical Director - Children's Health Network of Minnesota

NQF #1884 Depression Response at Six Months – Progress Towards Remission

Commenter

Collette Cole, on behalf of MN Community Measurement

Comment

Hello, During the process of submitting our scientific testing for this measure NQF# 1884 Depression Response at Six Months, we inadvertently did not include the c-statistic for this measure. This statistic was calculated during the logistic regression procedure but the clinical staff completing the application did not recognize the c-statistic in part due to the large number of pairs and the spacing of the table. The calculated concordance (c-statistic) for this measure was 0.578 (adults) and 0.552 (adolescents) which meet the criteria for a well calibrated model. Association of Predicted Probabilities and Observed Responses Adults Percent Concordant 57.8Somers' D 0.156 Percent Discordant 42.2Gamma 0.156 Percent Tied 0.0Tau-a 0.049 Pairs 2261788815c 0.578 Association of Predicted Probabilities and Observed Responses Adolescents Percent Concordant 55.2Somers' D 0.104 Percent Discordant 44.8Gamma 0.104 Percent Tied 0.0Tau-a 0.027 Pairs 17784665c 0.552 Please consider this additional

information in the standing committee's assessment of the risk adjustment model. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer, MN Community Measurement

NQF #1884 Depression Response at Six Months – Progress Towards Remission

Commenter

Koryn Rubin, on behalf of American Medical Association

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to express our concerns on the evidence and testing provided in support of this measure. While the AMA agrees that it is useful to understand the rate of response for individuals diagnosed with depression, we do not believe that the developer provided sufficient evidence demonstrating that depression scores can be successfully reduced by at least 50% across the defined patient population within a six-month timeframe nor was any evidence provided supporting this requirement of 50%. For example, would the measure better capture clinical care and patient outcomes if it measured a minimal clinically significant difference in the depression score. It is important that the data demonstrate that practices can implement structures or processes that lead to improved outcomes and the measure results in rates that truly reflect the quality of care delivered by a practice rather than differences in patient mix or other factors outside of the practice's control. We also seek clarification on whether this measure is intended to be captured as an electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing and missing feasibility scorecard for NQF #710e would also apply to this measure. The AMA requests that the gaps in evidence and clarification on whether the measure is intended to be an eCQM be addressed prior to continued endorsement of this measure. We appreciate the Committee's consideration of our comments.

NQF #1885 Depression Response at 12 Months – Progress Towards Remission

Commenter

Collette Cole, on behalf of MN Community Measurement

Comment

Hello, During the process of submitting our scientific testing for this measure NQF# 1885 Depression Response at Twelve Months, we inadvertently did not include the c-statistic for this measure. This statistic was calculated during the logistic regression procedure but the clinical staff completing the application did not recognize the c-statistic in part due to the large number of pairs and the spacing of the table. The calculated concordance (c-statistic) for this measure was 0.587 (adults) and 0.556 (adolescents) which meet the criteria for a well calibrated model. Association of Predicted Probabilities and Observed Responses Adults Percent Concordant 58.7Somers' D 0.173 Percent Discordant

41.3Gamma 0.173 Percent Tied 0.0Tau-a 0.049 Pairs 2042832300c 0.587 Association of Predicted Probabilities and Observed Responses Adolescents Percent Concordant 55.6Somers' D 0.113 Percent Discordant 44.4Gamma 0.113 Percent Tied 0.0Tau-a 0.028 Pairs 16879016c 0.556 Please consider this additional information in the standing committee's assessment of the risk adjustment model. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer, MN Community Measurement

NQF #1885 Depression Response at 12 Months – Progress Towards Remission

Commenter

Koryn Rubin, on the behalf of American Medical Association

Comment

The American Medical Association (AMA) appreciates the opportunity to comment on this measure. We are writing to express our concerns on the evidence and testing provided in support of this measure. While the AMA agrees that it is useful to understand the rate of response for individuals diagnosed with depression, we do not believe that the developer provided sufficient evidence demonstrating that depression scores can be successfully reduced by at least 50% across the defined patient population within a twelve-month timeframe nor was any evidence provided supporting this requirement of 50%. For example, would the measure better capture clinical care and patient outcomes if it measured a minimal clinically significant difference in the depression score. It is important that the data demonstrate that practices can implement structures or processes that lead to improved outcomes and the measure results in rates that truly reflect the quality of care delivered by a practice rather than differences in patient mix or other factors outside of the practice's control. We also seek clarification on whether this measure is intended to be captured as an electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing and missing feasibility scorecard for NQF #710e would also apply to this measure. The AMA requests that the gaps in evidence and clarification on whether the measure is intended to be an eCQM be addressed prior to continued endorsement of this measure. We appreciate the Committee's consideration of our comments.

Appendix G: Post-Evaluation Comments

NQF #1884 Depression Response at Six Months- Progress Towards Remission (Endorsed)

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8120 (Submitted: 07/21/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Hello, Based on the committee discussion that occurred during the measure review meeting, there are three areas that we wanted to follow up on to ensure that the committee has clarity on the measures as they consider issues related to validity. Data Element Validity: Data element validity demonstrates that there is agreement with an authoritative source of the same information. The data elements for these measures are contained in structured fields extracted directly from the EHR, not abstracted, and agreement with the source (medical record) is high. However, because extraction is occurring, MNCM performs patient level data element audits against the source medical record to demonstrate that extraction programs are working correctly. Critical data element audit against the medical record demonstrated 100% agreement with diagnosis of depression or dysthymia, 100% agreement with exclusions, 95% agreement with assessment date of PHQ-9, and 94% agreement with the PHQ-9 score. Risk Adjustment: MNCM provided additional information to the committee about the fit of the risk adjustment model, in response to concerns raised by NQF staff. Missing Data: This criteria ensures that missing data does not bias the data results. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Several committee members suggested that patients who are not assessed with a follow-up PHQ-9/PHQ-9M should be removed from the denominator, however this would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up. Thank you for your consideration!

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8182 (Submitted: 09/07/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Greetings, MN Community Measurement (MNCM) would like to clarify an issue that was raised in the Overarching Themes in the draft Behavioral and Substance Use Spring 2022 Cycle report. It was noted in the report that the committee had concerns about the lack of telehealth services included in the specifications of several measures. Telehealth services are included in the use and specification of the measures we steward (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712) and have been included as part of the denominator definition for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement period. *For this measure, an encounter includes but is not limited to any of the following: office visit, psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. We look forward to addressing any additional questions or concerns the committee may have during the post-comment webinar. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer MN Community Measurement

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 8217 (Submitted: 09/08/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

COMMENT

The American Medical Association (AMA) continues to have concerns with the insufficient evidence demonstrating that depression scores can be successfully reduced by at least 50% across the defined patient population within a six-month timeframe nor was any evidence provided supporting this requirement of 50%. We also do not see any discussion of our request for clarification on whether this measure has met all of the requirements for electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing for National Quality Forum (NQF) #710e would also apply to this measure. In addition, we agree with the Standing Committee's concerns over the omission of telehealth services and inclusion of patients lost to follow-up in this measure as there is significant potential for the quality of care to be misrepresented. The AMA continues to question whether this measure meets the NQF measure evaluation criteria and as a result, believes that these concerns must be addressed before endorsement is continued.

DEVELOPER RESPONSE

Thank you for your comments and interest in measures that strive to improve health outcomes for patients with major depression or dysthymia. The response measures are considered as an interim outcome, or progress towards the ultimate outcome of the remission of symptoms. It is not an unreasonable expectation to have symptoms reduced at six months which has a +60-day window extending the assessment timeframe out to eight months. The acute treatment phase of depression is 6 to 12 weeks, so an assessment of symptoms during the measure's timeframe is well into the continuation phase of treatment. It is not unreasonable to have an interim goal that is reduced by 50% or greater during this timeframe. Soares et al in their 2014 study of the duration of current episode and treatment with desvenlafaxine or placebo used as one of their study outcomes defined a response of greater than or equal to 50% decrease in SDS and HDRS-17 total score. Howland et al in the evaluation of the effectiveness of duloxetine also used a measure of greater than or equal to 50% decrease in the HDRS-17 total score. [Functional Recovery in Major Depressive Disorder: Focus on Early Optimized Treatment PSYCHIATRIST.COM Sept 2016] In terms of e-CQM development, we agree. This measure has the same denominator of patients, using the same tool to measure outcomes. MNCM had early discussions with CMS about including the response measures as well, but at that time the Measure Authoring Tool could not handle the mathematical equation to calculate the measure, so the response measures were not included in the e-CQM program. This measure is not an e-CQM, but it is a digital quality measure. All components are captured from discrete data elements in the electronic record and MNCM has been capturing this information in a digital format via EHR extraction for 10+ years. Regarding telehealth services, they are included in the measure and have been for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--

depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. Lastly, in terms of missing data, all patients who are eligible for the measure (have the diagnosis of major depression or dysthymia and an elevated PHQ-9 score) remain in the denominator of the measure. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Removing patients who are not assessed with a follow-up PHQ-9/PHQ-9M from the denominator would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Ms. Stacy Miller, Mental Health Outcomes

Comment ID#: 8248 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

Depression and the PHQ-9 is a strong combination for PRO due to the proportion of impacted patients and the PHQ-9's brief nature, ability to detect depression improvement across care settings, and presence in the public domain. Though these strengths remain, the BH quality measure landscape has changed since this set was first endorsed in 2011. Many payers and The Joint Commission now require PRO measurement and CMS is investigating PRO adoption for IPFQR. We believe re-endorsement of the set must include consideration of its place in the modern landscape. Alignment of PRO tools across organizations is more beneficial for providers, facilities, and patients alike. When different organizations endorse tools that result in disrupted continuity across and within care settings, facilities and providers have higher burden, providers must become intimately familiar with multiple tools and results, and patients and providers lose continuity of measurement. Therefore, re-endorsement of the set is also endorsement of the PHQ-9 as the NQF's tool of choice in measuring outcomes for depressed patients. This may be appropriate, however it may also be that TJC's approach of allowing facilities to choose a tool meeting specified criteria is a path toward PRO that allows facilities and providers to align tools across organizations.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

NQF thanks you for your comment. We appreciate your concern that NQF endorsement of a measure may appear to endorse a certain tool due to its use in recommended measures but want to assure you that NQF does not endorse measurement tools. Any measure developer may submit quality measures for the Standing Committee to review against the measure evaluation criteria and no preference is given to measures using specific tools.

NQF COMMITTEE RESPONSE

N/A

Dr. Tim Hernandez

Comment ID#: 8257 (Submitted: 06/07/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

My organization, Entira Family Clinics, has been using the suite of depression measures for over 15 years. We strongly feel that our success in developing our entire care coordination system is in large part attributed to the infrastructure that we needed to build to be successful with this measure. The measure pushes us to have a robust registry tool to successfully follow these patients. In addition, we must do outreach in a population that eschews follow up. Complete remission can be a challenge. Having a measure that tracks progress towards the overall goal of complete remission is an excellent incentive to clinicians who deal with difficult to manage populations. My hope is that NQF continues to support the good work that this measure has allowed us to do.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Andrew Lyzenga, American Psychiatric Association; Submitted by Andrew Lyzenga

Comment ID#: 8253 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: Member Does NOT Support

COMMENT

The American Psychiatric Association appreciates the opportunity to provide feedback on these measures being considered by the Behavioral Health and Substance Use Standing Committee. Gathering information through clinician or patient-completed screening and assessment tools is a critical part of quality care for patients with behavioral health conditions, as is measurement and tracking of outcomes over time. These activities are a core aspect of measurement-based care (MBC), which has been shown to be effective in improving outcomes and patient and provider satisfaction in both primary and specialty care. We would urge the developers to ensure that telehealth visits are included in these measures; we are encouraged that comments submitted by MN Community Measurement after the Standing Committee's initial evaluation suggest that telehealth services are indeed included in the measure specifications. Given mental health workforce shortages and maldistribution of providers, psychiatric consultations through telehealth have become an integral part of clinical practice, especially for communities that lack local expertise. We have some concern that there remains a lack of widespread and standardized infrastructure for collecting and reporting data on measures, and that there may be associated challenges with implementing these measures at a national level. However, we are hopeful that continued emphasis on measurement and data collection will spur allocation of additional resources and development of improved infrastructure in this area.

DEVELOPER RESPONSE

Thank you for your comments and support for the depression measures. Thank you also for reviewing the response comment regarding telehealth services, which are included in the denominators for the depression measures. We agree that the provision of telehealth services is important to the overall ability to deliver mental health services. It appears that there is one positive related to the covid-19 pandemic: the rapid expansion of telehealth. We have received feedback from our behavioral health providers in MN that telehealth has allowed them to provide more services and expand their reach in the rural settings where distance may have prohibited the delivery of services. We share your concern (and hope!) about a lack of widespread and standardized infrastructure for collecting and reporting data on measures on a national level. We have witnessed a widespread adoption of the PHQ-9/ PHQ-9M as a tool for assessing and monitoring depression symptoms, not only in Minnesota, but nationally as well, although perhaps more in the primary care space nationally than specialty care. In the 10+ years that MN has been reporting this measure, we have seen a significant uptake in use of the tool by behavioral specialists and many view the common tool as a great communication mechanism between primary and specialty care. We also understand your concern surrounding infrastructure as well

but believe that there is hope on the horizon. Many EHR vendors have incorporated the PHQ-9 into their system, we are most experienced with Epic, but several other vendors list having the PHQ-9 screening tool available within their system- GE Centricity, Kareo, athenaOne, TherapyNotes, RXNT, NextGen and TheraNest to name a few. We are hopeful that continued emphasis on measurement and data collection for patients with depression will focus attention and resources in this area and lead to improved health outcomes.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF PROPOSED COMMITTEE RESPONSE

N/A

NQF #1885 Depression Response at Twelve Months- Progress Towards Remission (Not Recommended)

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8184 (Submitted: 09/07/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Greetings, MN Community Measurement (MNCM) would like to clarify an issue that was raised in the Overarching Themes in the draft Behavioral and Substance Use Spring 2022 Cycle report. It was noted in the report that the committee had concerns about the lack of telehealth services included in the specifications of several measures. Telehealth services are included in the use and specification of the measures we steward (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712) and have been included as part of the denominator definition for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement period. *For this measure, an encounter includes but is not limited to any of the following: office visit, psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. We look forward to addressing any additional questions or concerns the committee may have during the post-comment webinar. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer MN Community Measurement

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8121 (Submitted: 07/21/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Hello, Based on the committee discussion that occurred during the measure review meeting, there are three areas that we wanted to follow up on to ensure that the committee has clarity on the measures as they consider issues related to validity. Data Element Validity: Data element validity demonstrates that there is agreement with an authoritative source of the same information. The data elements for these measures are contained in structured fields extracted directly from the EHR, not abstracted, and agreement with the source (medical record) is high. However, because extraction is occurring, MNCM performs patient level data element audits against the source medical record to demonstrate that extraction programs are working correctly. Critical data element audit against the medical record demonstrated 100% agreement with diagnosis of depression or dysthymia, 100% agreement with exclusions, 95% agreement with assessment date of PHQ-9, and 94% agreement with the PHQ-9 score. Risk Adjustment: MNCM provided additional information to the committee about the fit of the risk adjustment model, in response to concerns raised by NQF staff. Missing Data: This criteria ensures that missing data does not bias the data results. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Several committee members suggested that patients who are not assessed with a follow-up PHQ-9/PHQ-9M should be removed from the denominator, however this would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up. Thank you for your consideration!

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 8218 (Submitted: 09/08/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

COMMENT

The American Medical Association (AMA) continues to have concerns with the insufficient evidence demonstrating that depression scores can be successfully reduced by at least 50% across the defined patient population within a twelve-month timeframe nor was any evidence provided supporting this requirement of 50%. We also do not see any discussion of our request for clarification on whether this measure has met all of the requirements for electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing for National Quality Forum (NQF) #710e would also apply to this measure. In addition, we agree with the Standing Committee's concerns over the omission of telehealth services and inclusion of patients lost to follow-up in this measure as there is significant potential for the quality of care to be misrepresented. The AMA continues to question whether this measure meets the NQF measure evaluation criteria and as a result, believes that these concerns must be addressed before endorsement is continued.

DEVELOPER RESPONSE

Thank you for your comments and interest in measures that strive to improve health outcomes for patients with major depression or dysthymia. The response measures are considered as an interim outcome, or progress towards the ultimate outcome of the remission of symptoms. It is a reasonable expectation to have symptoms reduced at 12 months with the +60-day window extending the assessment timeframe out to fourteen months. The development workgroup members supported a stepwise approach to depression treatment in that patients not responding successfully to treatment need a re-evaluation and a change to their plan of care. This was grounded in the STAR*D study Sequenced Treatment Alternatives to Relieve Depression which noted that "measurement-based care—that is, using brief, easy-to-administer instruments to monitor depression severity and side effects, following an evidence-based treatment algorithm,

making decisions at key time points, and having remission as a goal of treatment—is a feasible strategy that can be adapted in real-world practice settings—both psychiatric and primary care settings." [Gaynes, B.N et al What did STAR*D Teach Us? Results From a Large-Scale, Practical, Clinical Trial for Patients with Depression https://doi.org/10.1176/ps.2009.60.11.1439] The acute treatment phase of depression is 6 to 12 weeks, so an assessment of symptoms during the measure's timeframe is well into the continuation phase of treatment. It is not unreasonable to have an interim goal that is reduced by 50% or greater during this timeframe. Soares et al in their 2014 study of the duration of current episode and treatment with desvenlafaxine or placebo used as one of their study outcomes defined a response of greater than or equal to 50% decrease in SDS and HDRS-17 total score. Howland et al in the evaluation of the effectiveness of duloxetine also used a measure of greater than or equal to 50% decrease in the HDRS-17 total score. [Functional Recovery in Major Depressive Disorder: Focus on Early Optimized Treatment PSYCHIATRIST.COM Sept 2016] In terms of e-CQM development, we agree. This measure has the same denominator of patients, using the same tool to measure outcomes. MNCM had early discussions with CMS about including the response measures as well, but at that time the Measure Authoring Tool could not handle the mathematical equation to calculate the measure, so the response measures were not included in the e-CQM program. This measure is not an e-CQM, but it is a digital quality measure. All components are captured from discrete data elements in the electronic record and MNCM has been capturing this information in a digital format via EHR extraction for 10+ years. Regarding telehealth services, they are included in the measure and have been for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. Lastly, in terms of missing data, all patients who are eligible for the measure (have the diagnosis of major depression or dysthymia and an elevated PHQ-9 score) remain in the denominator of the measure. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Removing patients who are not assessed with a follow-up PHQ-9/PHQ-9M from the denominator would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

PROPOSED NQF COMMITTEE RESPONSE

N/A

Ms. Stacy Miller, Mental Health Outcomes

Comment ID#: 8249 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

Depression and the PHQ-9 is a strong combination for PRO due to the proportion of impacted patients and the PHQ-9's brief nature, ability to detect depression improvement across care settings, and presence in the public domain. Though these strengths remain, the BH quality measure landscape has changed since this set was first endorsed in 2011. Many payers and The Joint Commission now require PRO measurement and CMS is investigating PRO adoption for IPFQR. We believe re-endorsement of the set must include consideration of its place in the modern landscape. Alignment of PRO tools across organizations is more beneficial for providers, facilities, and patients alike. When different organizations endorse tools that result in disrupted continuity across and within care settings, facilities and providers have higher burden, providers must become intimately familiar with multiple tools and results, and patients and providers lose continuity of measurement. Therefore, re-endorsement of the set is also endorsement of the PHQ-9 as the NQF's tool of choice in measuring outcomes for depressed patients. This may be appropriate, however it may also be that TJC's approach of allowing facilities to choose a tool meeting specified criteria is a path toward PRO that allows facilities and providers to align tools across organizations.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

NQF thanks you for your comment. We appreciate your concern that NQF endorsement of a measure may appear to endorse a certain tool due to its use in recommended measures but want to assure you that NQF does not endorse measurement tools. Any measure developer may submit quality measures for the Standing Committee to review against the measure evaluation criteria and no preference is given to measures using specific tools.

NQF COMMITTEE RESPONSE

N/A

Dr. Tim Hernandez

Comment ID#: 8256 (Submitted: 06/07/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

My organization, Entira Family Clinics, has been using the suite of depression measures for over 15 years. We strongly feel that our success in developing our entire care coordination system is in large part attributed to the infrastructure that we needed to build to be successful with this measure. The measure pushes us to have a robust registry tool to successfully follow these patients. In addition, we must do outreach in a population that eschews follow up. Complete remission can be a challenge. Having a measure that tracks progress towards the overall goal of complete remission is an incentive to clinicians who deal with difficult to manage populations. My hope is that NQF continues to support the good work that this measure has allowed. We need measures, as well as the support of payers to help us manage these measures, to push the excellent work that we have done in the state of Minnesota.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Andrew Lyzenga, American Psychiatric Association; Submitted by Andrew Lyzenga

Comment ID#: 8253 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: Member Does NOT Support

COMMENT

The American Psychiatric Association appreciates the opportunity to provide feedback on these measures being considered by the Behavioral Health and Substance Use Standing Committee. Gathering information through clinician or patient-completed screening and assessment tools is a critical part of quality care for patients with behavioral health conditions, as is measurement and tracking of outcomes over time. These activities are a core aspect of measurement-based care (MBC), which has been shown to be effective in improving outcomes and patient and provider satisfaction in both primary and specialty care. We would urge the developers to ensure that

telehealth visits are included in these measures; we are encouraged that comments submitted by MN Community Measurement after the Standing Committee's initial evaluation suggest that telehealth services are indeed included in the measure specifications. Given mental health workforce shortages and maldistribution of providers, psychiatric consultations through telehealth have become an integral part of clinical practice, especially for communities that lack local expertise. We have some concern that there remains a lack of widespread and standardized infrastructure for collecting and reporting data on measures, and that there may be associated challenges with implementing these measures at a national level. However, we are hopeful that continued emphasis on measurement and data collection will spur allocation of additional resources and development of improved infrastructure in this area.

DEVELOPER RESPONSE

Thank you for your comments and support for the depression measures. Thank you also for reviewing the response comment regarding telehealth services, which are included in the denominators for the depression measures. We agree that the provision of telehealth services is important to the overall ability to deliver mental health services. It appears that there is one positive related to the covid-19 pandemic: the rapid expansion of telehealth. We have received feedback from our behavioral health providers in MN that telehealth has allowed them to provide more services and expand their reach in the rural settings where distance may have prohibited the delivery of services. We share your concern (and hope!) about a lack of widespread and standardized infrastructure for collecting and reporting data on measures on a national level. We have witnessed a widespread adoption of the PHQ-9/ PHQ-9M as a tool for assessing and monitoring depression symptoms, not only in Minnesota, but nationally as well, although perhaps more in the primary care space nationally than specialty care. In the 10+ years that MN has been reporting this measure, we have seen a significant uptake in use of the tool by behavioral specialists and many view the common tool as a great communication mechanism between primary and specialty care. We also understand your concern surrounding infrastructure as well but believe that there is hope on the horizon. Many EHR vendors have incorporated the PHQ-9 into their system, we are most experienced with Epic, but several other vendors list having the PHQ-9 screening tool available within their system- GE Centricity, Kareo, athenaOne, TherapyNotes, RXNT, NextGen and TheraNest to name a few. We are hopeful that continued emphasis on measurement and data collection for patients with depression will focus attention and resources in this area and lead to improved health outcomes.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

NQF #0711 Depression Remission at Six Months (Not Recommended)

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8118 (Submitted: 07/21/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Hello, Based on the committee discussion that occurred during the measure review meeting, there are three areas that we wanted to follow up on to ensure that the committee has clarity on the measures as they consider issues related to validity. Data Element Validity: Data element validity demonstrates that there is agreement with an authoritative source of the same information. The data elements for these measures are contained in structured fields extracted directly from the EHR, not abstracted, and agreement with the source (medical record) is high. However, because extraction is occurring, MNCM performs patient level data element audits against the source medical record to demonstrate that extraction programs are working correctly. Critical data element audit against the medical record demonstrated 100% agreement with diagnosis of depression or dysthymia, 100% agreement with exclusions, 95% agreement with assessment date of PHQ-9, and 94% agreement with the PHQ-9 score. Risk Adjustment: MNCM provided additional information to the committee about the fit of the risk adjustment model, in response to concerns raised by NQF staff. Missing Data: This criteria ensures that missing data does not bias the data results. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Several committee members suggested that patients who are not assessed with a follow-up PHQ-9/PHQ-9M should be removed from the denominator, however this would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up. Thank you for your consideration!

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8180 (Submitted: 09/07/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Greetings, MN Community Measurement (MNCM) would like to clarify an issue that was raised in the Overarching Themes in the draft Behavioral and Substance Use Spring 2022 Cycle report. It was noted in the report that the committee had concerns about the lack of telehealth services included in the specifications of several measures. Telehealth services are included in the use and specification of the measures we steward (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712) and have been included as part of the denominator definition for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement period. *For this measure, an encounter includes but is not limited to any of the following: office visit, psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. We look forward to addressing any additional questions or concerns the committee may have during the post-comment webinar. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer MN Community Measurement

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 8216 (Submitted: 09/08/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

COMMENT

The American Medical Association (AMA) continues to have concerns with the insufficient evidence demonstrating that remission can be successfully achieved across the defined patient population within a six-month timeframe. We also do not see any discussion of our request for clarification on whether this measure has met all of the requirements for electronic clinical quality measures (eCQMs) since the complimentary measure (710e Depression Remission at Twelve Months), which is an eCQM, uses the same data and is specified similarly. It would seem counterintuitive to have related measures endorsed that leverage what appear to be the same data, yet are endorsed with different data sources and specifications. If it is intended to be an eCQM, our concerns on the inadequate testing for National Quality Forum (NQF) #710e also apply to this measure. In addition, we agree with the Standing Committee's concerns over the omission of telehealth services and inclusion of patients lost to follow-up in this measure as there is significant potential for the quality of care to be misrepresented. The AMA continues to question whether this measure meets the NQF measure evaluation criteria and as a result, believes that these concerns must be addressed before endorsement is continued.

DEVELOPER RESPONSE

Thank you for your comments and interest in measures that strive to improve health outcomes for patients with major depression or dysthymia. This patient-centric outcome measure is seeking to improve symptoms of depression significantly; remission defined as a PHQ-9 or PHQ-9M score of less than 5 which is defined by the tool's cut points as mild or no symptoms. The Depression Remission at Six Months measure seeks an assessment of outcomes at six months +60 days which has a window extending the assessment timeframe out to eight months. The acute treatment phase of depression is 6 to 12 weeks, so an assessment of symptoms during the measure's timeframe is well into the continuation phase of treatment. It is not unreasonable to strive for this patient centric outcome of remission of symptoms during this timeframe. In terms of e-CQM development, we agree. This measure, Depression Remission at Six Months was simultaneously developed as an e-CQM along with the companion measure (#0710-e Depression Remission at 12 months). This six-month measure not adopted into the e-CQM program, however it is a digital quality measure. All components are captured from discrete data elements in the electronic record and MNCM has been capturing this information in a digital format via EHR extraction for 10+ years. Regarding telehealth services, they are included in the measure and have been for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. Lastly, in terms of missing data, all patients who are eligible for the measure (have the diagnosis of major depression or dysthymia and an elevated PHQ-9 score) remain in the denominator of the measure. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up

assessment is not missing data, rather represents a gap in care. Removing patients who are not assessed with a follow-up PHQ-9/PHQ-9M from the denominator would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

PROPOSED NQF COMMITTEE RESPONSE

N/A

Ms. Stacy Miller, Mental Health Outcomes

Comment ID#: 8246 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

Depression and the PHQ-9 is a strong combination for PRO due to the proportion of impacted patients and the PHQ-9's brief nature, ability to detect depression improvement across care settings, and presence in the public domain. Though these strengths remain, the BH quality measure landscape has changed since this set was first endorsed in 2011. Many payers and The Joint Commission now require PRO measurement and CMS is investigating PRO adoption for IPFQR. We believe re-endorsement of the set must include consideration of its place in the modern landscape. Alignment of PRO tools across organizations is more beneficial for providers, facilities, and patients alike. When different organizations endorse tools that result in disrupted continuity across and within care settings, facilities and providers have higher burden, providers must become intimately familiar with multiple tools and results, and patients and providers lose continuity of measurement. Therefore, re-endorsement of the set is also endorsement of the PHQ-9 as the NQF's tool of choice in measuring outcomes for depressed patients. This may be appropriate, however it may also be that TJC's approach of allowing facilities to choose a tool meeting specified criteria is a path toward PRO that allows facilities and providers to align tools across organizations.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

NQF thanks you for your comment. We appreciate your concern that NQF endorsement of a measure may appear to endorse a certain tool due to its use in recommended measures but want

to assure you that NQF does not endorse measurement tools. Any measure developer may submit quality measures for the Standing Committee to review against the measure evaluation criteria and no preference is given to measures using specific tools.

NQF COMMITTEE RESPONSE

N/A

Dr. Tim Hernandez

Comment ID#: 8260 (Submitted: 06/07/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

My organization, Entira Family Clinics, has been using the suite of depression measures for over 15 years. We strongly feel that our success in developing our entire care coordination system is in large part attributed to the infrastructure that we needed to build to be successful with this measure. The measure pushes us to have a robust registry tool to successfully follow these patients. In addition, we must do outreach in a population that eschews follow up. Our clinicians are challenged by the timing of the measure to "treat to target" in order to get this population at their stated goal, as measured by a PROM. My organization has been amongst the top performing primary clinics in achieving high rates of depression remission in large part due to the laser focus that this measure dictates. My hope is that NQF will continue to endorse this measure as well as other measures in the depression remission/response suite.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Andrew Lyzenga, American Psychiatric Association; Submitted by Andrew Lyzenga

Comment ID#: 8253 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: Member Does NOT Support

COMMENT

The American Psychiatric Association appreciates the opportunity to provide feedback on these measures being considered by the Behavioral Health and Substance Use Standing Committee. Gathering information through clinician or patient-completed screening and assessment tools is a critical part of quality care for patients with behavioral health conditions, as is measurement and tracking of outcomes over time. These activities are a core aspect of measurement-based care (MBC), which has been shown to be effective in improving outcomes and patient and provider satisfaction in both primary and specialty care. We would urge the developers to ensure that telehealth visits are included in these measures; we are encouraged that comments submitted by MN Community Measurement after the Standing Committee's initial evaluation suggest that telehealth services are indeed included in the measure specifications. Given mental health workforce shortages and maldistribution of providers, psychiatric consultations through telehealth have become an integral part of clinical practice, especially for communities that lack local expertise. We have some concern that there remains a lack of widespread and standardized infrastructure for collecting and reporting data on measures, and that there may be associated challenges with implementing these measures at a national level. However, we are hopeful that continued emphasis on measurement and data collection will spur allocation of additional resources and development of improved infrastructure in this area.

DEVELOPER RESPONSE

Thank you for your comments and support for the depression measures. Thank you also for reviewing the response comment regarding telehealth services, which are included in the denominators for the depression measures. We agree that the provision of telehealth services is important to the overall ability to deliver mental health services. It appears that there is one positive related to the covid-19 pandemic: the rapid expansion of telehealth. We have received feedback from our behavioral health providers in MN that telehealth has allowed them to provide more services and expand their reach in the rural settings where distance may have prohibited the delivery of services. We share your concern (and hope!) about a lack of widespread and standardized infrastructure for collecting and reporting data on measures on a national level. We have witnessed a widespread adoption of the PHQ-9/ PHQ-9M as a tool for assessing and monitoring depression symptoms, not only in Minnesota, but nationally as well, although perhaps more in the primary care space nationally than specialty care. In the 10+ years that MN has been reporting this measure, we have seen a significant uptake in use of the tool by behavioral specialists and many view the common tool as a great communication mechanism between primary and specialty care. We also understand your concern surrounding infrastructure as well but believe that there is hope on the horizon. Many EHR vendors have incorporated the PHQ-9 into their system, we are most experienced with Epic, but several other vendors list having the PHQ-9 screening tool available within their system- GE Centricity, Kareo, athenaOne, TherapyNotes, RXNT, NextGen and TheraNest to name a few. We are hopeful that continued emphasis on measurement and data collection for patients with depression will focus attention and resources in this area and lead to improved health outcomes.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

NQF #3312 Continuity of Care After Medically Managed Withdrawal from Alcohol and/or Drugs (Recommended)

Dr. Uddin and Aaron Mchone, UnityPoint Health ; Submitted by Stephanie Collingwood

Comment ID#: 8147 (Submitted: 09/01/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

UnityPoint Health supports NQF measures 3312 and 3313. In general, UnityPoint Health feels <7 and <14 day follow ups after discharges from a medically managed withdrawal episode as well as <28 day follow ups after new antipsychotic prescriptions is a best practice that we support. However, we do have concerns around the age range and payor population proposed within the measures. We would fully support future broadening of these measures to a larger patient population(age group) and payor mix. As currently proposed, these measures are limited to Medicaid patients only and within certain age ranges. For example, it's clinically appropriate for a 65-year-old discharging from medical managed withdrawal and a 17-year-old on a new antipsychotic to also meet these measurement requirements. Additionally, we find a significant number of patients who qualify for these metrics are dual eligible (Medicaid and Medicare). In our experience, CMS and Medicaid MCOs have historically been unable to share claims data effectively. As such we are concerned that data for this metric would be inaccurate as different claims would go to different health plans. This is particularly true for patients discharging from a medically managed withdrawal facility which tend to be primarily reimbursed by Medicaid with the follow up being reimbursed through Medicare Part B. Overall, we are concern that claims-based data may not be as accurate for these measures due to a relatively high volume of dual eligibility patients and thus the outcomes of these measures may not reveal a complete view of the patients who receive, or should receive, this best practice, standard of care.

DEVELOPER RESPONSE

In response to your concerns about the age range of the measures in question: • For NQF 3312, the upper limit of 64 years was chosen after careful consideration of evidence from the literature, input from experts, feasibility of data collection, and findings from measure testing. CMS will continue to review the relevant research and reconsider the age range should new evidence emerge. • For NQF

3313, the current specifications currently do not set an upper limit for the age of individuals eligible for inclusion in the measure; only pediatric cases (i.e., those under age 18) are excluded from assessment. Like NQF 3312, CMS will reassess the age of those included in the measure's population, should new relevant research emerge. To your feedback about the accuracy of claims-based data for individuals eligible for both Medicare and Medicaid (i.e., dually enrolled participants): • CMS and its measurement development contractor completed comprehensive reevaluation of the NQF 3312 and NQF 3313 technical specifications in advance of NQF endorsement review. As part of this effort, CMS reviewed data sources to ensure the most accurate and complete data were used for measure calculation and testing. The primary data used for testing in the NQF 3312 and NQF 3313 submissions were Transformed Medicaid Statistical Information System (T-MSIS) Analytical Files (TAF); for participants dually enrolled in both Medicare and Medicaid, Medicare Parts A, B, C, and D claims data were also used. Using your example, an individual who had been discharged from a facility for medically managed withdrawal whose care was reimbursed using both Medicare and Medicaid would be captured in the NQF 3312 denominator population

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Ms. Stacy Miller, Mental Health Outcomes

Comment ID#: 8236 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

Mental Health Outcomes recognizes the importance of continuity of care for patients discharging from a medically managed withdrawal episode. With behavioral health challenges impacting greater numbers following the COVID-19 pandemic, it is important to consider the applications of measures that focus on the quality of inpatient psychiatric services. However, we are concerned with the exclusion of telemedicine codes in the current specifications for NQF#3312. According to the Morbidity and Mortality Weekly Report from CDC, telemedicine services have nearly doubled since the start of the pandemic. As smartphone ownership continues to increase among low-income populations, telehealth further increases patient access to care services. We are concerned that the measure in its current state is not consistent with the World Health Organization's recommendations to apply approaches that build supportive services for the future.

DEVELOPER RESPONSE

Thank you for your comment about Continuity of Care After Medically Managed Withdrawal from Alcohol and/or Drugs (NQF 3312). Testing of the technical specifications for NQF 3312 endorsement maintenance used Medicaid claims data from January 1, 2018, through December 15, 2018 (allowing for 7- and 14-day follow-up after the discharge); thus, the measure was assessed, quantitatively, using data for services provided prior to the COVID-19 public health emergency. In the future, when more recent Medicaid administrative claims data are available, CMS will consider retesting and updating of the technical specifications to capture follow up for those seeking medically managed care using a telemedicine CPT code(s).

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

NQF #3313 Follow-Up Care for Adult Medicaid Beneficiaries Who are Newly Prescribed an Antipsychotic Medication (Not Recommended)

Dr. Uddin and Aaron Mchone, UnityPoint Health; Submitted by Stephanie Collingwood

Comment ID#: 8148 (Submitted: 09/01/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

UnityPoint Health supports NQF measures 3312 and 3313. In general, UnityPoint Health feels <7 and <14 day follow ups after discharges from a medically managed withdrawal episode as well as <28 day follow ups after new antipsychotic prescriptions is a best practice that we support. However, we do have concerns around the age range and payor population proposed within the measures. We would fully support future broadening of these measures to a larger patient population(age group) and payor mix. As currently proposed, these measures are limited to Medicaid patients only and within certain age ranges. For example, it's clinically appropriate for a 65-year-old discharging from medical managed withdrawal and a 17-year-old on a new antipsychotic to also meet these measurement requirements. Additionally, we find a significant number of patients who qualify for these metrics are dual eligible (Medicaid and Medicare). In our experience, CMS and Medicaid MCOs have historically been unable to share claims data effectively. As such we are concerned that data for this metric would be inaccurate as different claims would go to different health plans. This is particularly true for patients discharging from a medically managed withdrawal facility which tend to be primarily reimbursed by Medicaid with the follow up

being reimbursed through Medicare Part B. Overall, we are concern that claims-based data may not be as accurate for these measures due to a relatively high volume of dual eligibility patients and thus the outcomes of these measures may not reveal a complete view of the patients who receive, or should receive, this best practice, standard of care.

DEVELOPER RESPONSE

Thank you for your support of continued endorsement for NQF 3312 and 3313. In response to your concerns about the age range of the measures in question: • For NQF 3312, the upper limit of 64 years was chosen after careful consideration of evidence from the literature, input from experts, feasibility of data collection, and findings from measure testing. CMS will continue to review the relevant research and reconsider the age range should new evidence emerge. • For NQF 3313, the current specifications currently do not set an upper limit for the age of individuals eligible for inclusion in the measure; only pediatric cases (i.e., those under age 18) are excluded from assessment. Like NQF 3312, CMS will reassess the age of those included in the measure's population, should new relevant research emerge. To your feedback about the accuracy of claims-based data for individuals eligible for both Medicare and Medicaid (i.e., dually enrolled participants): • CMS and its measurement development contractor completed comprehensive reevaluation of the NQF 3312 and NQF 3313 technical specifications in advance of NQF endorsement review. As part of this effort, CMS reviewed data sources to ensure the most accurate and complete data were used for measure calculation and testing. The primary data used for testing in the NQF 3312 and NQF 3313 submissions were Transformed Medicaid Statistical Information System (T-MSIS) Analytical Files (TAF); for participants dually enrolled in both Medicare and Medicaid, Medicare Parts A, B, C, and D claims data were also used. Using your example, an individual who had been discharged from a facility for medically managed withdrawal whose care was reimbursed using both Medicare and Medicaid would be captured in the NQF 3312 denominator population.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

NQF #0710e Depression Remission at Twelve Months (Not Recommended)

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8181 (Submitted: 09/07/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Greetings, MN Community Measurement (MNCM) would like to clarify an issue that was raised in the Overarching Themes in the draft Behavioral and Substance Use Spring 2022 Cycle report. It was noted in the report that the committee had concerns about the lack of telehealth services included in the specifications of several measures. Telehealth services are included in the use and specification of the measures we steward (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712) and have been included as part of the denominator definition for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement period. *For this measure, an encounter includes but is not limited to any of the following: office visit, psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. We look forward to addressing any additional questions or concerns the committee may have during the post-comment webinar. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer MN Community Measurement

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8119 (Submitted: 07/21/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Hello, Based on the committee discussion that occurred during the measure review meeting, there are three areas that we wanted to follow up on to ensure that the committee has clarity on the measures as they consider issues related to validity. Data Element Validity: Data element validity demonstrates that there is agreement with an authoritative source of the same information. The data elements for these measures are contained in structured fields extracted directly from the

EHR, not abstracted, and agreement with the source (medical record) is high. However, because extraction is occurring, MNCM performs patient level data element audits against the source medical record to demonstrate that extraction programs are working correctly. Critical data element audit against the medical record demonstrated 100% agreement with diagnosis of depression or dysthymia, 100% agreement with exclusions, 95% agreement with assessment date of PHQ-9, and 94% agreement with the PHQ-9 score. Risk Adjustment: MNCM provided additional information to the committee about the fit of the risk adjustment model, in response to concerns raised by NQF staff. Missing Data: This criteria ensures that missing data does not bias the data results. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Several committee members suggested that patients who are not assessed with a follow-up PHQ-9/PHQ-9M should be removed from the denominator, however this would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up. Thank you for your consideration!

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Koryn Y. Rubin, MHA, American Medical Association

Comment ID#: 8215 (Submitted: 09/08/2022)

Council / Public: HPR

Level of Support: Member Does NOT Support

COMMENT

The American Medical Association (AMA) continues to have concerns with the insufficient evidence demonstrating that remission can be successfully achieved across the defined patient population within a twelve-month timeframe. We also do not see any discussion of our request for clarification on whether this measure has met all of the requirements for electronic clinical quality measures (eCQMs) since the National Quality Forum (NQF) measure evaluation criteria require patient-encounter-level validity testing. In addition, we agree with the Standing Committee's concerns over

the omission of telehealth services and inclusion of patients lost to follow-up in this measure as there is significant potential for the quality of care to be misrepresented. The AMA continues to question whether this measure meets the NQF measure evaluation criteria and as a result, believes that these concerns must be addressed before endorsement is continued.

DEVELOPER RESPONSE

Thank you for your comments and interest in measures that strive to improve health outcomes for patients with major depression or dysthymia. This patient-centric outcome measure is seeking to improve symptoms of depression significantly; remission defined as a PHQ-9 or PHQ-9M score of less than 5 which is defined by the tool's cut points as mild or no symptoms. The Depression Remission at Twelve Months measure seeks an assessment of outcomes at twelve months +60 days which has a window extending the assessment timeframe out to fourteen months. The acute treatment phase of depression is 6 to 12 weeks, so an assessment of symptoms (outcome) during the measure's timeframe is well beyond the continuation phase of treatment. It is not unreasonable to strive for this patient centric outcome of remission of symptoms during this timeframe. Patient level, and in fact contact level (each PHQ-9/PHQ-9M) validation did occur. In addition to the NQF feasibility scorecard, the HQMF specifications (Measure Authoring Tool) and BONNIE testing results, individual data element (patient, encounter and contact level) validity results were submitted in question 2b.02 and 2b.03 of the application. In response to questions from the committee, additional statistics about the data element validation were provided: Data element validity demonstrates that there is agreement with an authoritative source of the same information. The data elements for these measures are contained in structured fields extracted directly from the EHR, not abstracted, and agreement with the source (medical record) is high. However, because extraction is occurring, MNCM performs patient level data element audits against the source medical record to demonstrate that extraction programs are working correctly. Critical data element audit against the medical record demonstrated 100% agreement with diagnosis of depression or dysthymia, 100% agreement with exclusions, 95% agreement with assessment date of PHQ-9, and 94% agreement with the PHQ-9 score. Regarding telehealth services, they are included in the measure and have been for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. Lastly, in terms of missing data, all patients who are eligible for the measure (have the diagnosis of major depression or dysthymia and an elevated PHQ-9 score) remain in the denominator of the measure. The measure construct purposely includes patients without a follow up in denominator in order to avoid bias in the measure. (e.g., selective tool administration to only patients who are doing well). It is important to note that lack of a follow-up assessment is not missing data, rather represents a gap in care. Removing patients who are not assessed with a follow-up PHQ-9/PHQ-9M from the denominator would introduce bias into the measure that currently does not exist. The current construct of keeping eligible patients in the denominator promotes inclusiveness and accountability for follow-up.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

PROPOSED NQF COMMITTEE RESPONSE

N/A

Ms. Stacy Miller, Mental Health Outcomes

Comment ID#: 8247 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

Depression and the PHQ-9 is a strong combination for PRO due to the proportion of impacted patients and the PHQ-9's brief nature, ability to detect depression improvement across care settings, and presence in the public domain. Though these strengths remain, the BH quality measure landscape has changed since this set was first endorsed in 2011. Many payers and The Joint Commission now require PRO measurement and CMS is investigating PRO adoption for IPFQR. We believe re-endorsement of the set must include consideration of its place in the modern landscape. Alignment of PRO tools across organizations is more beneficial for providers, facilities, and patients alike. When different organizations endorse tools that result in disrupted continuity across and within care settings, facilities and providers have higher burden, providers must become intimately familiar with multiple tools and results, and patients and providers lose continuity of measurement. Therefore, re-endorsement of the set is also endorsement of the PHQ-9 as the NQF's tool of choice in measuring outcomes for depressed patients. This may be appropriate, however it may also be that TJC's approach of allowing facilities to choose a tool meeting specified criteria is a path toward PRO that allows facilities and providers to align tools across organizations.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

NQF thanks you for your comment. We appreciate your concern that NQF endorsement of a measure may appear to endorse a certain tool due to its use in recommended measures but want to assure you that NQF does not endorse measurement tools. Any measure developer may submit quality measures for the Standing Committee to review against the measure evaluation criteria and no preference is given to measures using specific tools.

NQF COMMITTEE RESPONSE

N/A

Dr. Tim Hernandez

Comment ID#: 8259 (Submitted: 06/07/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

My organization, Entira Family Clinics, has been using the suite of depression measures for over 15 years. We strongly feel that our success in developing our entire care coordination system is in large part attributed to the infrastructure that we needed to build to be successful with this measure. The measure pushes us to have a robust registry tool to successfully follow these patients. In addition, we must do outreach in a population that eschews follow up. The 12 month time period keeps us focused on what is clearly a chronic disease. Our clinicians are challenged by the timing of the measure to "treat to target" in order to get this population at their stated goal, as measured by a PROM. My organization has been amongst the top performing primary clinics in achieving high rates of depression remission in large part due to the laser focus that this measure dictates. My hope is that NQF will continue to endorse this measure as well as other measures in the depression remission/response suite.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Andrew Lyzenga, American Psychiatric Association; Submitted by Andrew Lyzenga

Comment ID#: 8253 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: Member Does NOT Support

COMMENT

The American Psychiatric Association appreciates the opportunity to provide feedback on these measures being considered by the Behavioral Health and Substance Use Standing Committee.

Gathering information through clinician or patient-completed screening and assessment tools is a critical part of quality care for patients with behavioral health conditions, as is measurement and tracking of outcomes over time. These activities are a core aspect of measurement-based care (MBC), which has been shown to be effective in improving outcomes and patient and provider satisfaction in both primary and specialty care. We would urge the developers to ensure that telehealth visits are included in these measures; we are encouraged that comments submitted by MN Community Measurement after the Standing Committee's initial evaluation suggest that telehealth services are indeed included in the measure specifications. Given mental health workforce shortages and maldistribution of providers, psychiatric consultations through telehealth have become an integral part of clinical practice, especially for communities that lack local expertise. We have some concern that there remains a lack of widespread and standardized infrastructure for collecting and reporting data on measures, and that there may be associated challenges with implementing these measures at a national level. However, we are hopeful that continued emphasis on measurement and data collection will spur allocation of additional resources and development of improved infrastructure in this area.

DEVELOPER RESPONSE

Thank you for your comments and support for the depression measures. Thank you also for reviewing the response comment regarding telehealth services, which are included in the denominators for the depression measures. We agree that the provision of telehealth services is important to the overall ability to deliver mental health services. It appears that there is one positive related to the covid-19 pandemic: the rapid expansion of telehealth. We have received feedback from our behavioral health providers in MN that telehealth has allowed them to provide more services and expand their reach in the rural settings where distance may have prohibited the delivery of services. We share your concern (and hope!) about a lack of widespread and standardized infrastructure for collecting and reporting data on measures on a national level. We have witnessed a widespread adoption of the PHQ-9/ PHQ-9M as a tool for assessing and monitoring depression symptoms, not only in Minnesota, but nationally as well, although perhaps more in the primary care space nationally than specialty care. In the 10+ years that MN has been reporting this measure, we have seen a significant uptake in use of the tool by behavioral specialists and many view the common tool as a great communication mechanism between primary and specialty care. We also understand your concern surrounding infrastructure as well but believe that there is hope on the horizon. Many EHR vendors have incorporated the PHQ-9 into their system, we are most experienced with Epic, but several other vendors list having the PHQ-9 screening tool available within their system- GE Centricity, Kareo, athenaOne, TherapyNotes, RXNT, NextGen and TheraNest to name a few. We are hopeful that continued emphasis on measurement and data collection for patients with depression will focus attention and resources in this area and lead to improved health outcomes.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

NQF #0712 Depression Assessment with PHQ-9/ PHQ-9M (Not Recommended)

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8117 (Submitted: 07/21/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Hello, Based on the preliminary worksheet feedback and committee discussion during the measure review meeting, we would like to provide the information that was shared verbally during the measure introduction to address concerns related to evidence that administering a PHQ-9/PHQ-9M assessment tool is unrelated to the outcomes of the improvement in depression symptoms. Simply administering a PHQ-9 tool itself in isolation will not improve outcomes. Administering the PHQ-9 is like taking a blood pressure; you need to do something with the information to affect the outcome of hypertension. Depression is now being considered the sixth vital sign by many and assessing patients is critical to identifying depression and improving outcomes. We examined a set of over 26,000 patients to determine if the actual frequency of assessments was related to the outcomes of response and remission. Patients who are assessed more frequently, for example those with four to twelve PHQ-9's during the assessment period were three times more likely to achieve remission or response outcomes at twelve months as compared to patients who were assessed only one to three times. Odds ratio 2.79 for remission and 3.36 for response. In other words, patients with 3 or less assessments had a 6.3% rate of remission compared to 15.8% for those assessed more frequently. Patients with 3 or less assessments had a 10.5% rate of response compared to 28.3% for those assessed more frequently. Thank you for you consideration!

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Collette M. Cole, BSN, RN, CPHQ, Minnesota Community Measurement

Comment ID#: 8179 (Submitted: 09/07/2022)

Council / Public: PCH

Level of Support: N/A

COMMENT

Greetings, MN Community Measurement (MNCM) would like to clarify an issue that was raised in the Overarching Themes in the draft Behavioral and Substance Use Spring 2022 Cycle report. It was noted in the report that the committee had concerns about the lack of telehealth services included in the specifications of several measures. Telehealth services are included in the use and specification of the measures we steward (NQF #0710e, NQF #0711, NQF #1884, NQF #1885, and NQF #0712) and have been included as part of the denominator definition for several years. Currently, the denominator encounter event is defined as "Patients with an encounter* coded with Major Depression or Dysthymia (Major Depression or Dysthymia Value Set) during the specific measurement period. *For this measure, an encounter includes but is not limited to any of the following: office visit, psychiatry, or psychotherapy visit, telephone, or online encounter." https://helpdesk.mncm.org/helpdesk/KB/View/24186732-data-collection-technical-guide--depression-care. Specifications were updated in this manner to clarify the previous expression "face to face visit or contact" in which the intent of contact was any contact with the patient in which a diagnosis was made. We look forward to addressing any additional questions or concerns the committee may have during the post-comment webinar. Sincerely, Collette Cole, RN BSN CPHQ Clinical Measure Developer MN Community Measurement

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee.

NQF COMMITTEE RESPONSE

N/A

Ms. Stacy Miller, Mental Health Outcomes

Comment ID#: 8245 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

Depression and the PHQ-9 is a strong combination for PRO due to the proportion of impacted patients and the PHQ-9's brief nature, ability to detect depression improvement across care settings, and presence in the public domain. Though these strengths remain, the BH quality measure landscape has changed since this set was first endorsed in 2011. Many payers and The Joint Commission now require PRO measurement and CMS is investigating PRO adoption for IPFQR. We believe re-endorsement of the set must include consideration of its place in the modern landscape. Alignment of PRO tools across organizations is more beneficial for providers, facilities, and patients alike. When different organizations endorse tools that result in disrupted continuity across and within care settings, facilities and providers have higher burden, providers must become intimately familiar with multiple tools and results, and patients and providers lose continuity of measurement. Therefore, re-endorsement of the set is also endorsement of the PHQ-9 as the NQF's tool of choice in measuring outcomes for depressed patients. This may be appropriate, however it may also be that TJC's approach of allowing facilities to choose a tool meeting specified criteria is a path toward PRO that allows facilities and providers to align tools across organizations.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

NQF thanks you for your comment. We appreciate your concern that NQF endorsement of a measure may appear to endorse a certain tool due to its use in recommended measures but want to assure you that NQF does not endorse measurement tools. Any measure developer may submit quality measures for the Standing Committee to review against the measure evaluation criteria and no preference is given to measures using specific tools.

NQF COMMITTEE RESPONSE

N/A

Dr. Tim Hernandez

Comment ID#: 8258 (Submitted: 06/07/2022)

Council / Public: Public

Level of Support: N/A

COMMENT

My organization, Entira Family Clinics, has been using the suite of depression measures for over 15 years. We were able to give feedback to Minnesota Community Measurement when they began introducing the PHQ 9 as a diagnostic tool and we have used it since the inception. It is simple and

reproducible. As a PROM it is clearly patient centered. We have built it into our systems, i.e. EHR, to hardware the use of the tool. It does not take patients long to fill out and lends itself to a variety of implementation styles. We strongly endorse the continued use of this tool with the hope that NQF supports it.

DEVELOPER RESPONSE

N/A

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

Andrew Lyzenga, American Psychiatric Association; Submitted by Andrew Lyzenga

Comment ID#: 8253 (Submitted: 09/13/2022)

Council / Public: Public

Level of Support: Member Does NOT Support

COMMENT

The American Psychiatric Association appreciates the opportunity to provide feedback on these measures being considered by the Behavioral Health and Substance Use Standing Committee. Gathering information through clinician or patient-completed screening and assessment tools is a critical part of quality care for patients with behavioral health conditions, as is measurement and tracking of outcomes over time. These activities are a core aspect of measurement-based care (MBC), which has been shown to be effective in improving outcomes and patient and provider satisfaction in both primary and specialty care. We would urge the developers to ensure that telehealth visits are included in these measures; we are encouraged that comments submitted by MN Community Measurement after the Standing Committee's initial evaluation suggest that telehealth services are indeed included in the measure specifications. Given mental health workforce shortages and maldistribution of providers, psychiatric consultations through telehealth have become an integral part of clinical practice, especially for communities that lack local expertise. We have some concern that there remains a lack of widespread and standardized infrastructure for collecting and reporting data on measures, and that there may be associated challenges with implementing these measures at a national level. However, we are hopeful that continued emphasis on measurement and data collection will spur allocation of additional resources and development of improved infrastructure in this area.

DEVELOPER RESPONSE

Thank you for your comments and support for the depression measures. Thank you also for reviewing the response comment regarding telehealth services, which are included in the denominators for the depression measures. We agree that the provision of telehealth services is important to the overall ability to deliver mental health services. It appears that there is one positive related to the covid-19 pandemic: the rapid expansion of telehealth. We have received feedback from our behavioral health providers in MN that telehealth has allowed them to provide more services and expand their reach in the rural settings where distance may have prohibited the delivery of services. We share your concern (and hope!) about a lack of widespread and standardized infrastructure for collecting and reporting data on measures on a national level. We have witnessed a widespread adoption of the PHQ-9/ PHQ-9M as a tool for assessing and monitoring depression symptoms, not only in Minnesota, but nationally as well, although perhaps more in the primary care space nationally than specialty care. In the 10+ years that MN has been reporting this measure, we have seen a significant uptake in use of the tool by behavioral specialists and many view the common tool as a great communication mechanism between primary and specialty care. We also understand your concern surrounding infrastructure as well but believe that there is hope on the horizon. Many EHR vendors have incorporated the PHQ-9 into their system, we are most experienced with Epic, but several other vendors list having the PHQ-9 screening tool available within their system- GE Centricity, Kareo, athenaOne, TherapyNotes, RXNT, NextGen and TheraNest to name a few. We are hopeful that continued emphasis on measurement and data collection for patients with depression will focus attention and resources in this area and lead to improved health outcomes.

NQF RESPONSE

Thank you for your comment. It has been shared with the Standing Committee and developer.

NQF COMMITTEE RESPONSE

N/A

National Quality Forum
1099 14th Street NW, Suite 500
Washington, DC 20005
https://www.qualityforum.org