Measure Worksheet (MEW-PA-New)

Partnership for Quality Measurement

Click here for Pre-Evaluation Public Comments

Click here for Measure Specifications

Content

Brief Measure Information

CBE #: 3747

Corresponding Measures: N/A

Measure Title: Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization

Measure Steward: New York State Office of Mental Health

sp.02. Brief Description of Measure: The percentage of discharges for members 6 years of age and older who were hospitalized for treatment of selected mental illness or intentional self-harm diagnoses and who had at least five follow-up community-based mental health care visits in the 90 days after discharge.

1b.01. Developer Rationale: The HEDIS measure Follow-Up After Hospitalization for Mental Illness captures one outpatient follow-up visit within seven and thirty days after psychiatric discharge. We think that it is important to also measure engagement in community-based mental health care over a longer time period post-discharge. People being discharged from a psychiatric hospitalization are in an acute phase of their mental health condition and need to receive more than one visit for adequate treatment.

This measure will improve the quality of care by encouraging both linkage to community-based mental health care services after an inpatient psychiatric discharge and engagement with community-based mental health services. Follow-up care after an inpatient psychiatric discharge has been associated with reduced inpatient readmissions, increased medication utilization, increased outpatient encounters, and increased functioning.

sp.12. Numerator Statement: To meet the numerator criteria for this measure, discharges must receive five or more follow-up

Content
visits with a community-based mental health care provider within 90 days after discharge for inpatient treatment of select mental health or intentional self-harm diagnoses. Follow-up visits that occur on the date of discharge are not included. **sp.14. Denominator Statement:** The eligible population for this measure is acute inpatient discharges ages 6-64 principally hospitalized for select mental illnesses or intentional self-harm and enrolled in Medicaid on the date of discharge through 90 days after discharge. Discharges with acute direct transfers, acute readmissions with total length of stay of 42 or more days, non-acute direct transfers, and non-acute readmissions are excluded. If members have more than one discharge during the measurement year, all discharges during the measurement year will be included. **sp.16. Denominator Exclusions:** In addition to the discharges with acute direct transfers, certain acute readmissions, non-acute direct transfers, and non-acute readmissions detailed above, discharges who are dually enrolled in Medicare and Medicaid and discharges in hospice or using hospice services anytime during the measurement year are excluded.
Measure Type: Process **sp.28. Data Source:** Claims **sp.07. Level of Analysis:** Health Plan
IF Endorsement Maintenance—Original Endorsement Date: N/A New Measure **Most Recent Endorsement Date:** N/A New Measure
IF this measure is included in a composite, Composite#/title: N/A **IF this measure is paired/grouped, CBE#/title:** N/A **sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?** N/A
Staff Assessment: New Measure

Content
Criterion 1: Importance to Measure and Report
1a. Evidence. The evidence requirements for a ***structure, process, or intermediate outcome*** measure are that it is based on a systematic review (SR) and grading of the body of empirical evidence in which the specific focus of the evidence matches what is being measured. For measures derived from a patient report, the evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.
The developer provides the following description for this measure: • This is a new process measure at the health plan level that assesses member discharges (%) who underwent hospitalization treatment of 1) selected mental illness or 2) intentional self-harm diagnoses and who completed at least five follow-up community-based mental health care visits. • The developer provides a logic model that depicts the flow from hospitalization (for mental illness or self-ham diagnosis) to short term outcomes (e.g., discharge planning, connection to provider), to intermediate outcomes (e.g., 5+ visits to a community-based mental health care facility, symptom management, medication management), and long term outcomes (e.g., medication adherence, continued treatment, improved functioning). • Benefits described included improvement of care via linkage to community-based mental health care services, noting need for follow-up care (beyond a single visit) after discharge for best patient results.
The developer provides the following evidence for this measure: • SR of the evidence specific to this measure? ☒ Yes ☐ No • Quality, Quantity, and Consistency of evidence provided? ☒ Yes ☐ No • Evidence graded? ☒ Yes ☐ No

Summary:
• The developer provided evidence from the American Psychiatric Association (APA) Practice Guideline for The Treatment of Patients with Schizophrenia supporting a holistic approach to treatment, combining medication, psychosocial interventions, and support services to address the various needs of individuals with Schizophrenia.
  ○ The developer noted the guidelines are either recommended or suggested. The guidelines use a numbering system to indicate the level of confidence in the recommendations. A "1" signifies a strong recommendation with clear benefits outweighing harms. Conversely, a "2" indicates greater uncertainty, where assessing the balance between benefits and harms is more difficult or the benefits and harms are less clear.
  ○ The developer notes that three ratings are assigned to the evidence: high, moderate, and low (denoted by the letters A, B, and C, respectively). The ratings reflect the level of confidence that the evidence for a guideline statement reflects a true effect based on consistency of findings across studies, directness of the effect on a specific

Content

health outcome, precision of the estimate of effect, and risk of bias in available studies.

- The developer provided evidence from the Practice Guideline for the Treatment of Patients with Bipolar Disorder supporting the importance of psychiatric management and maintenance treatment for the care of adult patients with bipolar disorder.
  - The developer notes each reference cited in the guideline was assigned a letter corresponding to the nature of the supporting evidence.
    - [A] Randomized clinical trial, [B] Clinical trial, [C] Cohort or longitudinal study, [D] Control study, [E] Review with secondary data analysis, [F] Review, and [G] Other.
  - The developer notes each recommendation is identified as falling into one of three categories of endorsement: [I] Recommended with substantial clinical confidence, [II] Recommended with moderate clinical confidence, and [III] May be recommended on the basis of individual circumstances.
- The developer provided evidence from the Practice Guideline for the Treatment of Patients with Major Depressive Disorder supporting a comprehensive approach that addresses functional impairment and quality of life, monitor the patient's psychiatric status, tailors treatment to the acute phase, and continually evaluate treatment responses.
  - The developer notes each reference cited in the guideline was assigned a letter corresponding to the nature of the supporting evidence.
    - [A] Randomized double-blind clinical trial, [A−] Randomized clinical trial, [B] Clinical trial, [C] Cohort or longitudinal study, [D] Case-control study, [E] Review with secondary data analysis, [F] Review, and [G] Other.
  - The developer notes each recommendation is identified as falling into one of three categories of endorsement: [I] Recommended with substantial clinical confidence, [II] Recommended with moderate clinical confidence, and [III] May be recommended on the basis of individual circumstances.
- The developer provided evidence from the Substance Abuse and Mental Health Services Administration's (SAMHSA) publication on Treatment for Suicidal Ideation, Self-Harm, and Suicide Attempts Among Youth that highlights practices and programs used to treat suicidal thoughts and behaviors.
  - The developer notes trained reviewers gave each study a rating for causal impact: high support of causal evidence, moderate support of causal evidence, and low support of causal evidence.
  - The reviewers gave each program a rating based on the number of studies with strong, moderate, or emerging support of causal impact.

Exception to evidence

- NA

Questions for the Standing Committee:

- *How strong is the evidence for this relationship?*

Content

- *Are the selected mental illnesses sufficient to address the greatest need for improvement in care?*
- *Is the evidence directly applicable to the process of care being measured?*

Guidance From the Evidence Algorithm

Process measure based on systematic review (Box 3) -> QQC presented (Box 4) -> Quantity: moderate; Quality: moderate; Consistency: high (Box 5) -> Moderate (Box 5b) -> Moderate.

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and Disparities

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The evidence provided is inclusive of several years and diagnoses, indicating variability in scores (SD, minimums, maximums) but does not report if differences are statistically significant. For example, the mean for children (aged 6-20 years) was 51.2, whereas the mean for adults (aged 21-64 years) was 43.6. SDs varied (Children = 20.8; Adult = 11.5) with minimums equaling 0.0 and 11.3, respectively). However, it is not clear if age scores significantly differ.

Disparities

- Developer provided scores derived from analysis of NYS Medicaid discharge data (2018-2021). Collapsing across age, ethnicity, race, sex, and primary diagnosis. Means across years were 46.1,46.4,45.5, and 46.9 respectively. Standard deviations (SD) ranged between 12.0-13.1. Minimums were all below 20. Maximums were all above 61. Thus, great variability was apparent.
- Developer also provided annual data stratified by age, ethnicity, race, sex, and primary diagnosis for 2018-2021 NYS Medicaid Discharge data.
- For example, 2018 age results: mean for children (aged 6-20 years) was 51.2, whereas mean for adults (aged 21-64 years) was 43.6. SDs varied (Children = 20.8; Adult = 11.5) with minimums equaling 0.0 and 11.3, respectively). Further, a maximum for children of 69.2 was reported (compared to 66.4 for adults).

Questions for the Standing Committee:

- *Is there a gap in care that warrants a national performance measure?*

Preliminary rating for opportunity for improvement:
☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Criteria 2: Scientific Acceptability of Measure Properties

Content
Complex measure evaluated by the Scientific Methods Panel (SMP)? ☐ Yes ☒ No
Evaluators: Battelle Staff
2a. Reliability: Specifications and Testing
2a1. Specifications require the measure, as specified, to produce consistent (i.e., reliable) and credible (i.e., valid) results about the quality of care when implemented.
2a2. Reliability testing demonstrates whether the measure data elements are repeatable and producing the same results a high proportion of the time when assessed in the same population in the same time period, and/or whether the measure score is precise enough to distinguish differences in performance across providers.
Specifications: • Measure specifications are clear and precise.
Reliability Testing: • Reliability testing conducted at the Accountable Entity Level: o The developer conducted a signal to noise analysis using a beta-binomial model for thirty-one Medicaid managed plans and a Medicaid fee-for-service that operates in New York State for a total of thirty-two entities. o All Calendar Year 2018 discharges, 49,873 in total, found in NYS Medicaid administrative data were included in the analysis. o Patient counts and percentages were broken down by age, sex, ethnicity, race and primary diagnosis at discharge. o Information on social risk factors was not available for the data used in the analysis. o The developer ran analysis using a SAS Macro program created by John Adams, who developed this methodology for the National Committee for Quality Assurance. o Reporting entities were divided into terciles based on the denominator size. o Ten entities with smallest denominators were in the first tercile, next 11 entities were in second tercile and 11 largest entities based on denominator size were in third tercile. o Mean Signal-to-Noise Reliability, standard error (SE) and 95% confidence interval (CI) were calculated for all discharges combined and each tercile. o Minimum, maximum and percentiles (10, 25, 50, 75, 90) of the Signal-to-Noise Reliability were given for all discharges combined and each tercile. o The mean reliability for all eligible discharges was 0.946 with a 95% CI of 0.924 to 0.967, which the developer notes is considered high (greater than 0.9). The second and third terciles had mean reliability of 0.969 and 0.992, respectively, which are also considered high. The first tercile had a mean reliability of 0.870 which is considered moderate.

Content

Questions for the Standing Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are the measure specifications adequate)?*

Guidance From the Reliability Algorithm

Measure specifications precise, unambiguous, and complete (Box 1) -> Empirical reliability testing conducted with the measure as specified (Box 2) -> Reliability testing conducted with computed measure scores (Box 4) -> Method appropriate for assessing variability (signal-to-noise analysis) (Box 5) -> High certainty or confidence that the performance scores are reliable (Box 6a) -> High

Preliminary rating for reliability: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity: Validity Testing; Exclusions; Risk Adjustment; Meaningful Differences; Comparability; Missing Data

2b2. Validity testing should demonstrate that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Validity Testing

- Accountable Entity Level
  - To test construct validity, the developer calculated the Pearson correlation coefficient between Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization and the NCQA HEDIS measure Follow-Up After Hospitalization for Mental Illness.
    - Hypothesized moderate positive correlation due to overlapping eligible population and similarities in measure (62% of HEDIS overlap)
    - Moderately positive correlation (r = 0.56, N = 50,234)
  - Empirical validity
    - The developer calculated Concordance Statistics (or C Statistics) between Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization and three outcomes: mental health inpatient readmissions, psychotropic medication adherence, and continued engagement in care at six months post discharge.
  - Concordance statistics between measure and 3 outcomes:
    - Mental health inpatient readmissions (C = 0.53)
    - Psychotropic medication adherence (C = ~0.64 for 3 medication types)
    - Continued engagement in care at 6 months post-discharge (C = 0.72, meets predictability threshold)

Content
○ The developer notes the measure was reasonably predictive of engagement in community based mental health care at six months post discharge but did not meet the 0.7 C Statistic threshold for predictability for acute service use or medication adherence. ○ Face validity ▪ The developer consulted a workgroup consisting of mental health clinicians and researchers. ▪ The developer notes the workgroup agreed the measure was clinically sound. ▪ The developer highlights the workgroup poll found 92% of respondents thought there was a clinical benefit to the measure.
Exclusions • This measure excludes: • Discharges with acute direct transfers • Certain acute readmissions • Non-acute direct transfers • All non-acute readmissions • Discharges who are dually enrolled in Medicare and Medicaid • Discharges in hospice or using hospice services any time during the measurement year • The developer states that these exclusions are common for Medicaid performance measures, including the NCQA HEDIS measure Follow-Up After Hospitalization for Mental Illness, and were not tested.
Risk Adjustment • Measure is stratified by risk category/subgroup ○ Age stratification: 6-20 years old, 21-64 years old ○ Total (non-stratified) rate also reported
Meaningful Differences • The developer calculated an interquartile range to determine if there were practically meaningful differences in performance measure scores between managed care plans. • T-test conducted to compare randomly selected plan above the 75th percentile performance and below the 25th percentile performance and was statistically significant ($p < 0.0001$) • Interquartile range 12 percentage points
Missing Data • The developer notes that it is not possible to know how many claims were not submitted in Calendar Year 2018 because they have no secondary data sources reporting the number of services provided. • The developer notes it is not expected that a large proportion of claims to be missing as they are tied to reimbursement.
Comparability • The measure only uses one set of specifications.

Content
Questions for the Standing Committee regarding validity: • *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?*
Guidance From the Validity Algorithm All potential threats to validity were empirically assessed (Box 1) -> Empirical validity testing was conducted using the measure as specified (Box 2) -> Validity testing was conducted of the accountable entities using multiple statistical methods (Box 5) -> Validity methods were described and appropriate for assessing conceptual and theoretically hypothesized relationships (Box 6) -> Based on the empirical validity testing results, there is moderate confidence that the accountable entity data are a valid indicator of quality (Box 7b) -> Moderate The highest possible rating is Moderate.
Preliminary rating for validity: ☐ High ☒ Moderate ☐ Low ☐ Insufficient
Criterion 3. Feasibility
3. Feasibility is the extent to which the specifications, including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement. • The developer indicates data elements are in defined fields in electronic claims. • The developer notes there are no efforts to develop an eCQM with this measure concept by NYS OMH. • The developer notes that HEDIS value sets are used to calculate this measure, therefore, to use HEDIS value sets, fees must be paid to NCQA. No fees will be collected by NYS OMH.
Questions for the Standing Committee: • *Is the data collection strategy ready to be put into operational use?*
Preliminary rating for feasibility: ☒ High ☐ Moderate ☐ Low ☐ Insufficient
Criterion 4: Use and Usability
4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)
4a. Use evaluates the extent to which audiences (e.g., consumers, purchasers, providers, and policymakers) use or could use performance results for both accountability and performance improvement activities.
4a1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If they are not in use at the time of initial endorsement, then a credible plan for implementation within the specified time

Content
frames is provided.
Current uses of the measure Publicly reported? ☐ Yes ☒ No Current use in an accountability program? ☐ Yes ☒ No ☐ UNCLEAR Planned use in an accountability program? ☒ Yes ☐ No ☐ N/A

Accountability program details
• The developer outlines planned use in one or more New York State accountability and quality improvement platforms that share performance data by region, county, plan, provider agencies, and/or provider networks.
• The developer notes the measure will be embedded within a robust statewide mental health quality improvement infrastructure including:
  ○ NYS Quality Assurance Reporting Requirements
  ○ NYS OMH Vital Signs Dashboard
  ○ Psychiatric Services and Clinical Knowledge Enhancement System (PSYCKES)
  ○ Statewide Quality Improvement Collaboratives (QIC)

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: (1) Those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; (2) Those being measured, and other users have been given an opportunity to provide feedback on the measure performance or implementation; and (3) This feedback has been considered when changes are incorporated into the measure.

Feedback on the measure provided by those being measured or others
• The developer reports that:
  ○ Blinded plan level measure results for 28 Medicaid Managed Care plans were presented during the Fall 2022 Behavioral Health Clinical Advisory Group meeting.
  ○ Four NYS Medicaid Managed Care plans were offered the opportunity to participate in a field test of this measure and statewide product line level measure data were provided.
  ○ Once the measure is implemented, NYS has an infrastructure in place for providing performance results and assistance with interpretation.
  ○ There are statewide provider webinars, office hours and a help desk to address questions on interpretation of measure performance and how to use data to drive change.
  ○ In a poll taken during the Fall 2022 meeting of the NYS Behavioral Health Clinical Advisory Group, 92% of respondents see benefit in measuring 3-month engagement, 69% find five visits reasonable, 15% find five visits

Content
insufficient and 16% find it excessive. ○ In the spring 2022 meeting of the NYS Behavioral Health Clinical Advisory group, respondents thought the numerator standard of five visits was reasonable and raised no objections.
Questions for the Standing Committee: • *How have (or can) the performance results be used to further the goal of high quality, efficient healthcare?* • *How has the measure been vetted in real-world settings by those being measured or others?*
Preliminary rating for Use: ☒ Pass ☐ No Pass
4b. Usability (4b1. Improvement; 4b2. Benefits of measure)
4b. Usability evaluates the extent to which audiences (e.g., consumers, purchasers, providers, and policymakers) use or could use performance results for both accountability and performance improvement activities.
4b1 Improvement. Progress toward achieving the goal of high quality, efficient healthcare for individuals or populations is demonstrated. **Improvement results** • The developer notes the measure is new and not yet implemented.
4b2. Benefits versus harms. The benefits of the performance measure in facilitating progress toward achieving high quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists). **Unexpected findings (positive or negative) during implementation** • The measure has not been implemented. **Potential harms** • None identified by the developer.
Additional Feedback: • N/A
Questions for the Standing Committee: • *How can the performance results be used to further the goal of high quality, efficient healthcare?* • *Do the benefits of the measure outweigh any potential unintended consequences?*

Content
Preliminary rating for Usability and Use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**
Criterion 5: Related and Competing Measures
Related Measures • CBE #0576 Follow-Up After Hospitalization for Mental Illness
Harmonization • The developer reports that this measure and #0576 use the same value sets, with the exception of the discharge diagnosis value set. • This measure has a broader discharge diagnosis value set than the HEDIS measure and includes diagnoses such as anxiety and eating disorders that the HEDIS measure does not include.

QUALITY MEASURE SUBMISSION FORM

Version: 1.0; Generated: 13 April 2023

Introduction

Thank you for your interest in submitting a measure to Battelle for possible endorsement.

What criteria are used to evaluate measures? Measures are evaluated on standardized criteria: importance to measure and report, scientific acceptability of measure properties, feasibility, usability and use, and related and competing measures. For your measure to be evaluated against these measure evaluation criteria, you must complete the measure submission form.

Why do I have to complete a form? Due to the volume and/or complexity of proposed measures, Battelle provides measure information to committee reviewers in a standardized format to facilitate their evaluation of whether the measure meets the measure evaluation criteria. This form allows the measure steward to present information demonstrating that the proposed measure meets endorsement criteria.

What is on the form? The information requested in this form is directly related to the measure evaluation criteria.

Can't I just submit our files for consideration? No. Measures must be submitted through the online form to be considered for the Spring 2023 cycle. Requested information should be entered directly into this form and as well as any necessary or required attachments.

Can I submit additional details and materials? Additional materials will be considered only as supplemental. Do NOT rely on material provided in an appendix to provide measure specifications or to demonstrate meeting the criteria. The core information needed to evaluate the measure should be provided in the appropriate submission form fields and required attachments. Please contact PQMsupport@battelle.org regarding questions about submitting supplemental materials.

What do I do first? If you have started a new submission by answering five qualifying questions, you may proceed to the “Previous Submission Information" tab to continue with your submission. The “Conditions” tab will list the conditions that must be met before your proposed measures may be considered and evaluated for suitability as endorsed voluntary consensus standards. You are asked to acknowledge reading and accepting the conditions.

Can I make changes to a form once I have submitted it? No. Once you submit your measure, you will NOT be able to return to this submission form to make further revisions. You will need to contact project staff.

What if I need additional help? Please contact the project staff at PQMsupport@battelle.org if you have questions regarding the information requested or submitting supplemental materials.

NOTE: All measure submissions should be 508-compliant. Refer to the Checklist for Developer 508 Guidelines (PDF) to ensure all guidelines apply to all parts of your submission, including all fields and attachments used within the measure submission form.

Please email us at PQMsupport@battelle.org if you experience technical difficulties using the online submission form.

Thank you for your interest in submitting measures to Battelle.

Previous Submission Information (1 – 4)

1) Select whether this measure was previously submitted to the prior consensus-based entity (the National Quality Forum [NQF]) and given an identifying number.

☐ Previously submitted to NQF
☒ New measure, never submitted.

2) Provide the measure number of the previously submitted measure.
An intent to submit for Spring 2023 was sent to NQF and this measure was given number 3747.

3) If the measure has an electronic clinical quality measure (eCQM) version, provide the measure number of the previously submitted measure.
N/A
4) If this eCQM has a registry version, provide the measure numbers of the previously submitted measure.
N/A

Conditions (1 - 2)

Several conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. If any of the conditions are not met, the measure will not be accepted for consideration.

A. A Measure Steward Agreement is signed or the steward is a government organization. (All non-government organizations must sign a Measure Steward Agreement.) For more information about completing a Measure Steward Agreement, please go to: Endorsement | Partnership for Quality Measurement (p4qm.org) and follow the instructions.

B. The measure owner/steward verifies there is an identified responsible entity and a process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every three years.

C. The intended use of the measure includes both accountability applications (including public reporting) and performance improvement to achieve high-quality, efficient healthcare.

D. The measure is fully specified and tested for reliability and validity.

E. The measure developer/steward attests that harmonization with related measures and issues with competing measures have been considered and addressed, as appropriate.

F. The requested measure submission information is complete and responsive to the questions so that all the information needed to evaluate all criteria is provided.

1) Check if either of the following apply.

☐ Proprietary measure or components (e.g., risk model, codes)
☒ Proprietary measure or components with fees
☐ None of the above

2) Check the box below to agree to the conditions listed above.

☒ I have read and accept the conditions as specified above

Specifications: Maintenance Update (spma.01 - spma.02)

spma.01) Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.

☒ No
☐ Yes

spma.02) Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

For example, specifications may have been updated based on suggestions from a previous measure endorsement review.

Measure Specifications (sp.01 - sp.32)

sp.01) Provide the measure title.

Measure titles should be concise yet convey who and what is being measured.
Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization

sp.02) Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).
The percentage of discharges for members 6 years of age and older who were hospitalized for treatment of selected mental illness or intentional self-harm diagnoses and who had at least five follow-up community-based mental health care visits in the 90 days after discharge.

sp.03) Provide a rationale for why this measure must be reported with other measures to appropriately interpret results.

This measure does not need to be reported with other measures to appropriately interpret results.

sp.04) Check all the clinical condition/topic areas that apply to your measure, below.

☒ Behavioral Health
☐ Behavioral Health: Alcohol, Substance Use/Abuse
☒ Behavioral Health: Anxiety
☐ Behavioral Health: Attention Deficit Hyperactivity Disorder (ADHD)
☒ Behavioral Health: Bipolar Disorder
☒ Behavioral Health: Depression
☐ Behavioral Health: Domestic Violence
☒ Behavioral Health: Other Serious Mental Illness
☒ Behavioral Health: Post-Traumatic Stress Disorder (PTSD)
☒ Behavioral Health: Schizophrenia
☒ Behavioral Health: Suicide
☐ Cancer
☐ Cancer: Bladder
☐ Cancer: Breast
☐ Cancer: Colorectal
☐ Cancer: Gynecologic
☐ Cancer: Hematologic

- ☐ Cancer: Liver
- ☐ Cancer: Lung, Esophageal
- ☐ Cancer: Prostate
- ☐ Cancer: Renal
- ☐ Cancer: Skin
- ☐ Cancer: Thyroid
- ☐ Cardiovascular
- ☐ Cardiovascular: Arrythmia
- ☐ Cardiovascular: Congestive Heart Failure
- ☐ Cardiovascular: Coronary Artery Disease
- ☐ Cardiovascular: Coronary Artery Disease (AMI)
- ☐ Cardiovascular: Coronary Artery Disease (PCI)
- ☐ Cardiovascular: Hyperlipidemia
- ☐ Cardiovascular: Hypertension
- ☐ Cardiovascular: Secondary Prevention
- ☐ Critical Care
- ☐ Critical Care: Assisted Ventilation
- ☐ Critical Care: Intensive Monitoring
- ☐ Dental
- ☐ Dental: Caries
- ☐ Dental: Tooth Loss
- ☐ Ears, Nose, Throat (ENT)
- ☐ Ears, Nose, Throat (ENT): Ear Infection
- ☐ Ears, Nose, Throat (ENT): Hearing
- ☐ Ears, Nose, Throat (ENT): Pharyngitis
- ☐ Ears, Nose, Throat (ENT): Tonsilitis
- ☐ Endocrine
- ☐ Endocrine: Calcium and Metabolic Bone Disorders
- ☐ Endocrine: Diabetes
- ☐ Endocrine: Female and Male Endocrine Disorders
- ☐ Endocrine: Hypothalamic-Pituitary Disorders
- ☐ Endocrine: Thyroid Disorders
- ☐ Eye Care
- ☐ Eye Care: Age-related macular degeneration (AMD)
- ☐ Eye Care: Cataracts
- ☐ Eye Care: Diabetic retinopathy
- ☐ Eye Care: Glaucoma
- ☐ Gastrointestinal (GI)
- ☐ Gastrointestinal (GI): Constipation
- ☐ Gastrointestinal (GI): Gall Bladder Disease

- ☐ Gastrointestinal (GI): Gastroenteritis
- ☐ Gastrointestinal (GI): Gastro-Esophageal Reflux Disease (GERD)
- ☐ Gastrointestinal (GI): Hemorrhoids
- ☐ Gastrointestinal (GI): Hernia
- ☐ Gastrointestinal (GI): Inflammatory Bowel Disease
- ☐ Gastrointestinal (GI): Irritable Bowel Syndrome
- ☐ Gastrointestinal (GI): Peptic Ulcer
- ☐ Genitourinary (GU)
- ☐ Genitourinary (GU): Benign Prostatic Hyperplasia
- ☐ Genitourinary (GU): Erectile Dysfunction/Premature Ejaculation
- ☐ Genitourinary (GU): Incontinence/pelvic floor disorders
- ☐ Genitourinary (GU): Prostatitis
- ☐ Genitourinary (GU): Urinary Tract Injection (UTI)
- ☐ Gynecology (GYN)
- ☐ Gynecology (GYN): Abnormal bleeding
- ☐ Gynecology (GYN): Endometriosis
- ☐ Gynecology (GYN): Infections
- ☐ Gynecology (GYN): Menopause
- ☐ Gynecology (GYN): Pelvic Pain
- ☐ Gynecology (GYN): Uterine fibroids
- ☐ Infectious Diseases (ID)
- ☐ Infectious Diseases (ID): HIV/AIDS
- ☐ Infectious Diseases (ID): Influenza
- ☐ Infectious Diseases (ID): Lyme Disease
- ☐ Infectious Diseases (ID): Meningococcal Disease
- ☐ Infectious Diseases (ID): Pneumonia and respiratory infections
- ☐ Infectious Diseases (ID): Sepsis
- ☐ Infectious Diseases (ID): Sexually Transmitted
- ☐ Infectious Diseases (ID): Tuberculosis
- ☐ Liver
- ☐ Liver: Viral Hepatitis
- ☐ Musculoskeletal
- ☐ Musculoskeletal: Falls and Traumatic Injury
- ☐ Musculoskeletal: Gout
- ☐ Musculoskeletal: Joint Surgery
- ☐ Musculoskeletal: Low Back Pain
- ☐ Musculoskeletal: Osteoarthritis
- ☐ Musculoskeletal: Osteoporosis
- ☐ Musculoskeletal: Rheumatoid Arthritis
- ☐ Neurology

- ☐ Neurology: Alzheimer's Disease
- ☐ Neurology: Autism
- ☐ Neurology: Brain Injury
- ☐ Neurology: Epilepsy
- ☐ Neurology: Migraine
- ☐ Neurology: Parkinson's Disease
- ☐ Neurology: Spinal Cord Injury
- ☐ Neurology: Stroke/Transient Ischemic Attack (TIA)
- ☐ Other (please specify here:)
- ☐ Palliative Care and End-of-Life Care
- ☐ Palliative Care and End-of-Life Care: Advanced Directives
- ☐ Palliative Care and End-of-Life Care: Amyotrophic Lateral Sclerosis (ALS)
- ☐ Palliative Care and End-of-Life Care: Hospice Management
- ☐ Palliative Care and End-of-Life Care: Inappropriate use of acute care services
- ☐ Palliative Care and End-of-Life Care: Pain Management
- ☐ Perinatal Health
- ☐ Perinatal Health: Labor and Delivery
- ☐ Perinatal Health: Newborn Care
- ☐ Perinatal Health: Post-Partum Care
- ☐ Perinatal Health: Preconception Care
- ☐ Perinatal Health: Prenatal Care
- ☐ Renal
- ☐ Renal: Acute Kidney Injury
- ☐ Renal: Chronic Kidney Disease (CKD)
- ☐ Renal: End Stage Renal Disease (ESRD)
- ☐ Renal: Infections
- ☐ Reproductive Health
- ☐ Reproductive Health: Family planning and contraception
- ☐ Reproductive Health: Infertility
- ☐ Reproductive Health: Male reproductive health
- ☐ Respiratory
- ☐ Respiratory: Acute Bronchitis
- ☐ Respiratory: Allergy
- ☐ Respiratory: Asthma
- ☐ Respiratory: Chronic Obstructive Pulmonary Disease (COPD)
- ☐ Respiratory: Dyspnea
- ☐ Respiratory: Pneumonia
- ☐ Respiratory: Sleep Apnea
- ☐ Surgery
- ☐ Surgery: Cardiac Surgery

☐ Surgery: Colorectal
☐ Surgery: Neurosurgery / Spinal
☐ Surgery: Orthopedic
☐ Surgery: Orthopedic Hip/Pelvic Fractures
☐ Surgery: Pediatric
☐ Surgery: Perioperative and Anesthesia
☐ Surgery: Plastic
☐ Surgery: Thoracic Surgery
☐ Surgery: Trauma
☐ Surgery: Vascular Surgery

sp.05) Check all the non-condition specific measure domain areas that apply to your measure, below.

☒ Access to Care
☒ Care Coordination
☐ Care Coordination: Readmissions
☒ Care Coordination: Transitions of Care
☐ Disparities Sensitive
☐ Health and Functional Status
☐ Health and Functional Status: Change
☐ Health and Functional Status: Nutrition
☐ Health and Functional Status: Obesity
☐ Health and Functional Status: Physical Activity
☐ Health and Functional Status: Quality of Life
☐ Health and Functional Status: Total Health
☐ Immunization
☐ Other (please specify here:)
☐ Person-and Family-Centered Care: Person-and Family-Centered Care
☐ Person-and Family-Centered Care: Workforce
☐ Primary Prevention
☐ Primary Prevention: Nutrition
☐ Primary Prevention: Tobacco Use
☐ Safety
☐ Safety: Complications
☐ Safety: Healthcare Associated Infections
☐ Safety: Medication
☐ Safety: Overuse
☐ Screening

sp.06) Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.

☒ Adults (Age >= 18)
☒ Children (Age < 18)
☐ Elderly (Age >= 65)
☐ Populations at Risk: Dual eligible beneficiaries of Medicare and Medicaid
☐ Populations at Risk: Individuals with multiple chronic conditions
☐ Populations at Risk: Veterans
☐ Women

sp.07) Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.

☐ Accountable Care Organization
☐ Clinician: Group/Practice
☐ Clinician: Individual
☐ Facility
☒ Health Plan
☐ Integrated Delivery System
☐ Other (please specify here:)
☐ Population: Community, County or City
☐ Population: Regional and State

sp.08) Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.
☐ Ambulatory Care
☒ Behavioral Health
☐ Home Care
☐ Inpatient/Hospital
☐ Other (please specify here:)
☐ Outpatient Services
☒ Post-Acute Care

sp.09) Provide a Uniform Resource Locator (URL) link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.

Do not enter a URL linking to a home page or to general information. If no URL is available, indicate “none available".

None available.

sp.10) Indicate whether Health Quality Measure Format (HQMF) specifications are attached.

Attach the zipped output from the measure authoring tool (MAT) for eCQMs - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications). HQMF specifications are attached.

☒ HQMF specifications are NOT attached (Please explain).
This measure is not an eCQM.

sp.11) Attach the simulated testing attachment.

All eCQMs require a simulated testing attachment to confirm that the HTML output from Bonnie testing (or testing of some other simulated data set) includes 100% coverage of measured patient population testing, with pass/fail test cases for each sub-population. This can be submitted in the form of a screenshot.

☐ Testing is attached
☒ Testing is NOT attached (please explain)
This measure is not an eCQM.

sp.12) Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, contact staff at PQMsupport@battelle.org. Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

☒ Available in attached Excel or csv file
☐ No data dictionary/code table – all information provided in the submission form

For the question below: state the outcome/process being measured. Calculations of the risk-adjusted outcome measures should be described in sp.22.

sp.13) State the numerator.

Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).

DO NOT include the rationale for the measure.

To meet the numerator criteria for this measure, discharges must receive five or more follow-up visits with a community-based mental health care provider within 90 days after

discharge for inpatient treatment of select mental health or intentional self-harm diagnoses. Follow-up visits that occur on the date of discharge are not included.

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.14) Provide details needed to calculate the numerator.

All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

Five of any of the following within 90 days of discharge meet the criteria for community-based mental health follow-up visits.

- An outpatient visit (Visit Setting Unspecified Value Set) with (Outpatient POS Value Set) with a mental health provider.
- An outpatient visit (BH Outpatient Value Set) with a mental health provider.
- An intensive outpatient encounter or partial hospitalization (Visit Setting Unspecified Value Set) with (Partial Hospitalization POS Value Set).
- An intensive outpatient encounter or partial hospitalization (Partial Hospitalization or Intensive Outpatient Value Set).
- A community mental health center visit (Visit Setting Unspecified Value Set; BH Outpatient Value Set; Observation Value Set; Transitional Care Management Services Value Set) with (Community Mental Health Center POS Value Set).
- Electroconvulsive therapy (Electroconvulsive Therapy Value Set) with (Ambulatory Surgical Center POS Value Set; Community Mental Health Center POS Value Set; Outpatient POS Value Set; Partial Hospitalization POS Value Set).
- A telehealth visit: (Visit Setting Unspecified Value Set) with (Telehealth POS Value Set) with a mental health provider.
- An observation visit (Observation Value Set) with a mental health provider.
- Transitional care management services (Transitional Care Management Services Value Set), with a mental health provider.
- A visit in a behavioral healthcare setting (Behavioral Healthcare Setting Value Set).
- A telephone visit (Telephone Visits Value Set) with a mental health provider.
- Psychiatric collaborative care management (Psychiatric Collaborative Care Management Value Set).

Question Response: A new Excel file (3747_Engagement in Care Value Sets V2) was uploaded to sp. 12. This file has a separate tab for each value set listed in sp. 14. Please see the Excel file for the definition of the value sets. Each tab shows the codes

that are included in the value set.

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.15) State the denominator.

Brief, narrative description of the target population being measured.
The eligible population for this measure is acute inpatient discharges ages 6-64 principally hospitalized for select mental illnesses or intentional self-harm and enrolled in Medicaid on the date of discharge through 90 days after discharge. Discharges with acute direct transfers, acute readmissions with total length of stay of 42 or more days, non-acute direct transfers, and non-acute readmissions are excluded. If members have more than one discharge during the measurement year, all discharges during the measurement year will be included.

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.16) Provide details needed to calculate the denominator.

All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets.

Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

Product lines	Medicaid
Ages	6 – 64 years old as of the date of discharge. Report two age stratifications and a total rate: 1. 6-20 years old 2. 21-64 years old 3. Total
Continuous Enrollment Allowable gap	Date of discharge through 90 days after discharge No gaps in enrollment

Anchor date	None.
Benefits	Medical, Mental Health, and Substance Use.
Event/diagnosis	An acute inpatient discharge with a principal diagnosis of mental illness or intentional self-harm (Mental Health Discharge Diagnosis Value Set ("MH Discharge Diagnosis" tab in the Excel file attached to sp.12); Intentional Self-Harm Value Set) on the discharge claim on or between October 1 of the year prior to the measurement year and September 30 of the measurement year. To identify acute inpatient discharges: 1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set). 2. Exclude nonacute inpatient stays (Nonacute Inpatient Stay Value Set). 3. Identify the discharge date for the stay. The denominator for this measure is based on discharges, not on members. If members have more than one discharge, include all discharges on or between October 1 of the year prior to the measurement year and September 30 of the measurement year.

Acute readmission or direct transfers	Follow the steps below to exclude discharges with acute direct transfers, certain acute readmissions, non-acute direct transfers, and non-acute readmissions. Identify readmissions and direct transfers to an acute inpatient care setting during the 90-day follow-up period: 1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set). 2. Exclude nonacute inpatient stays (Nonacute Inpatient Stay Value Set). 3. Identify the admission date for the stay (the admission date must occur during the 90 day follow up period). If the admission date is the same date or one day later than the initial discharge, the admission is considered a direct transfer. 4. Identify the discharge date for the stay 5. For readmissions (admission dates 2-90 days after the initial discharge), calculate the length of stay in days (discharge date of the stay minus the admission date of the stay). Exclude both the initial discharge and direct transfer discharge if the last discharge occurs after September 30 of the measurement year. If the direct transfer to the acute inpatient care setting was for a principal diagnosis (use only the principal diagnosis on the discharge claim) of mental health disorder or intentional self-harm (Mental Health Diagnosis Value Set; Intentional Self-Harm Value Set), exclude the initial discharge. If the direct transfer to the acute inpatient care setting was for any other principal diagnosis (use only the principal diagnosis on the discharge claim) exclude both the original and the direct transfer discharge. Calculate the total length of stay in days for acute readmissions 2-90 days after the initial discharge. If this sum is 42 days or more, exclude the initial discharge.
Nonacute readmission or direct transfer	Exclude discharges followed by readmission or direct transfer to a nonacute inpatient care setting within the 90-day follow-up period, regardless of principal diagnosis for the readmission. To identify readmissions and direct transfers to a nonacute inpatient care setting: 1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set). 2. Confirm the stay was for nonacute care based on the presence of a nonacute code (Nonacute Inpatient Stay Value Set) on the claim. 3. Identify the admission date for the stay. These discharges are excluded from the measure because rehospitalization or direct transfer may prevent an outpatient follow-up visit from taking place.

Question Response: A new Excel file (3747_Engagement in Care Value Sets V2) was uploaded to sp.

12. This file has a separate tab for each value set listed in sp. 14. The tabs have the same name as the value set except where noted. Please see the Excel file for the definition of the value sets. Each tab shows the codes that are included in the value set.

sp.17) Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.
In addition to the discharges with acute direct transfers, certain acute readmissions, non-acute direct transfers, and non-acute readmissions detailed above, discharges who are dually enrolled in Medicare and Medicaid and discharges in hospice or using hospice services anytime during the measurement year are excluded.

sp.18) Provide details needed to calculate the denominator exclusions.

All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.
To identify Medicare and Medicaid dual enrollment, Medicare enrollment data should be queried. If the discharge was enrolled in Medicare any time between the date of discharge through 90 days post discharge, exclude the discharge from the denominator. Exclude discharges who use hospice services or elect to use a hospice benefit any time during the measurement year, regardless of when the services began. These discharges may be identified using the NCQA HEDIS Hospice Encounter Value Set or NCQA HEDIS Hospice Intervention Value Set.

sp.19) Provide all information required to stratify the measure results, if necessary.

Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.

Two age stratifications and a total rate should be reported. Age should be calculated at the date of discharge as date of discharge – date of birth:

- Stratification 1: 6-20 years old
- Stratification 2: 21-64 years old
- Total Rate

sp.20) Is this measure adjusted for socioeconomic status (SES)?

☐ Yes
☒ No

sp.21) Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.

☐ No risk adjustment or risk stratification
☐ Statistical risk model
☒ Stratification by risk category/subgroup (specify number of risk factors)
☐ Other approach to address risk factors (please specify here:)

Two age stratifications and a total rate will be reported.

- Stratification 1: 6-20 years old
- Stratification 2: 21-64 years old
- Total rate

sp.22) Select the most relevant type of score.

Attachment: If available, please provide a sample report.

☐ Categorical, e.g., yes/no
☐ Continuous variable, e.g. average
☐ Count
☐ Frequency Distribution
☐ Non-weighted score/composite/scale
☐ Other (please specify here:)
☒ Rate/proportion
☐ Ratio
☐ Weighted score/composite scale

sp.23) Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score.

☒ Better quality = Higher score
☐ Better quality = Lower score
☐ Better quality = Score within a defined interval
☐ Passing score defines better quality

sp.24) Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

Step One – calculate the eligible population as detailed below:

Product lines	Medicaid
Ages	6 – 64 years old as of the date of discharge. Report two age stratifications and a total rate: 1. 6-20 years old 2. 21-64 years old 3. Total
Continuous Enrollment Allowable gap	Date of discharge through 90 days after discharge No gaps in enrollment.
Anchor date	None.
Benefits	Medical, Mental Health, and Substance Use.
Event/diagnosis	An acute inpatient discharge with a principal diagnosis of mental illness or intentional self-harm (Mental Health Discharge Diagnosis Value Set ("MH Discharge Diagnosis" tab in the Excel file attached to sp.12); Intentional Self-Harm Value Set) on the discharge claim on or between October 1 of the year prior to the measurement year and September 30 of the measurement year. To identify acute inpatient discharges: 1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set). 2. Exclude nonacute inpatient stays (Nonacute Inpatient Stay Value Set). 3. Identify the discharge date for the stay. The denominator for this measure is based on discharges, not on members. If members have more than one discharge, include all discharges on or between October 1 of the year prior to the measurement year and September 30 of the measurement year.

Acute readmission or direct transfers	Follow the steps below to exclude discharges with acute direct transfers, certain acute readmissions, non-acute direct transfers, and non-acute readmissions. Identify readmissions and direct transfers to an acute inpatient care setting during the 90-day follow-up period: 1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set). 2. Exclude nonacute inpatient stays (Nonacute Inpatient Stay Value Set). 3. Identify the admission date for the stay (the admission date must occur during the 90 day follow-up period). If the admission date is the same date or one day later than the initial discharge, the admission is considered a direct transfer. 4. Identify the discharge date for the stay 5. For readmissions (admission dates 2-90 days after the initial discharge), calculate the length of stay in days (discharge date of the stay minus the admission date of the stay). 6. Exclude both the initial discharge and direct transfer discharge if the last discharge occurs after September 30 of the measurement year. 7. If the direct transfer to the acute inpatient care setting was for a principal diagnosis (use only the principal diagnosis on the discharge claim) of mental health disorder or intentional self-harm (Mental Health Diagnosis Value Set; Intentional Self-Harm Value Set), exclude the initial discharge. If the direct transfer to the acute inpatient care setting was for any other principal diagnosis (use only the principal diagnosis on the discharge claim) exclude both the original and the direct transfer discharge. Calculate the total length of stay in days for acute readmissions 2-90 days after the initial discharge. If this sum is 42 days or more, exclude the initial discharge.
Nonacute readmission or direct transfer	Exclude discharges followed by readmission or direct transfer to a nonacute inpatient care setting within the 90-day follow-up period, regardless of principal diagnosis for the readmission. To identify readmissions and direct transfers to a nonacute inpatient care setting: 1. Identify all acute and nonacute inpatient stays (Inpatient Stay Value Set). 2. Confirm the stay was for nonacute care based on the presence of a nonacute code (Nonacute Inpatient Stay Value Set) on the claim.

Product lines	Medicaid
Exclusions	3. Identify the admission date for the stay. These discharges are excluded from the measure because rehospitalization or direct transfer may prevent an outpatient follow-up visit from taking place. Medicare duals are excluded. Discharges in hospice or using hospice services anytime during the measurement year are excluded.

Step Two – include all discharges in the eligible population in the measure denominator.

Step Three – determine how many of the discharges in the denominator meet the numerator criteria as detailed below:

To be included in the numerator, a discharge must have received five or more follow-up visits with a community-based mental health care provider within 90 days after discharge. Do not include visits that occur on the date of discharge. Any of the following meet the criteria for a follow-up visit.

- An outpatient visit (Visit Setting Unspecified Value Set) with (Outpatient POS Value Set) with a mental health provider.
- An outpatient visit (BH Outpatient Value Set) with a mental health provider.
- An intensive outpatient encounter or partial hospitalization (Visit Setting Unspecified Value Set) with (Partial Hospitalization POS Value Set).
- An intensive outpatient encounter or partial hospitalization (Partial Hospitalization or Intensive Outpatient Value Set).
- A community mental health center visit (Visit Setting Unspecified Value Set; BH Outpatient Value Set; Observation Value Set; Transitional Care Management Services Value Set) with (Community Mental Health Center POS Value Set).
- Electroconvulsive therapy (Electroconvulsive Therapy Value Set) with (Ambulatory Surgical Center POS Value Set; Community Mental Health Center POS Value Set; Outpatient POS Value Set; Partial Hospitalization POS Value Set).
- A telehealth visit: (Visit Setting Unspecified Value Set) with (Telehealth POS Value Set) with a mental health provider.
- An observation visit (Observation Value Set) with a mental health provider.
- Transitional care management services (Transitional Care Management Services Value Set), with a mental health provider.
- A visit in a behavioral healthcare setting (Behavioral Healthcare Setting Value Set).

- A telephone visit (Telephone Visits Value Set) with a mental health provider.
- Psychiatric collaborative care management (Psychiatric Collaborative Care Management Value Set).

Question Response: A new Excel file (3747_Engagement in Care Value Sets V2) was uploaded to sp. 12. This file has a separate tab for each value set listed in sp. 14. The tabs have the same name as the value set except where noted. Please see the Excel file for the definition of the value sets. Each tab shows the codes that are included in the value set.

sp.25) Attach a copy of the instrument (e.g. survey, tool, questionnaire, scale) used as a data source for your measure, if available.

☐ Copy of instrument is attached.
☒ Copy of instrument is NOT attached (please explain).
An instrument was not used as a data source for this measure.

sp.26) Indicate the responder for your instrument.

☐ Patient
☐ Family or other caregiver
☐ Clinician
☐ Other (specify)
N/A

sp.27) If measure testing is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

Examples of samples used for testing:

• Testing may be conducted on a sample of the accountable entities (e.g., hospital, physician). The analytic unit specified for the particular measure (e.g., physician, hospital, home health agency) determines the sampling strategy for scientific acceptability testing.

• The sample should represent the variety of entities whose performance will be measured. The samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.

• The sample should include adequate numbers of units of measurement and adequate numbers of patients to answer the specific reliability or validity question with the chosen statistical method.

• When possible, units of measurement and patients within units should be randomly

selected.
Not applicable. Testing for this measure was not based on a sample.

sp.28) Identify whether and how proxy responses are allowed.
Not applicable.

sp.29) Survey/Patient-reported data.

Provide instructions for data collection and guidance on minimum response rate. Specify calculation of response rates to be reported with performance measure results.
Not applicable.

sp.30) Select only the data sources for which the measure is specified.

☐ Assessment Data
☒ Claims
☐ Electronic Health Data
☐ Electronic Health Records
☐ Instrument-Based Data
☐ Management Data
☐ Other (please specify here:)
☐ Paper Medical Records
☐ Registry Data

sp.31) Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.
Medicaid administrative data.

sp.32) Provide the data collection instrument.

☐ Available at measure-specific web page URL identified in sp.09
☐ Available in attached appendix in Question 1 of the Additional Section
☒ No data collection instrument provided

Importance to Measure and Report: Maintenance of Endorsement (1ma.01)

1ma.01) Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

☐ Yes
☐ No

Importance to Measure and Report: Evidence (Complete for Outcome Measures) (1a.01 - 1a.03)

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

Current Submission:

Updated evidence information here.

Previous (Year) Submission:

Evidence from the previous submission here.

1a.01) Provide a logic model.

Briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

1a.02) Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.

Describe how and from whom input was obtained.

1a.03) Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.

Importance to Measure and Report: Evidence (Complete for Process Measures) (1a.03 - 1a.16)

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:

Current Submission:

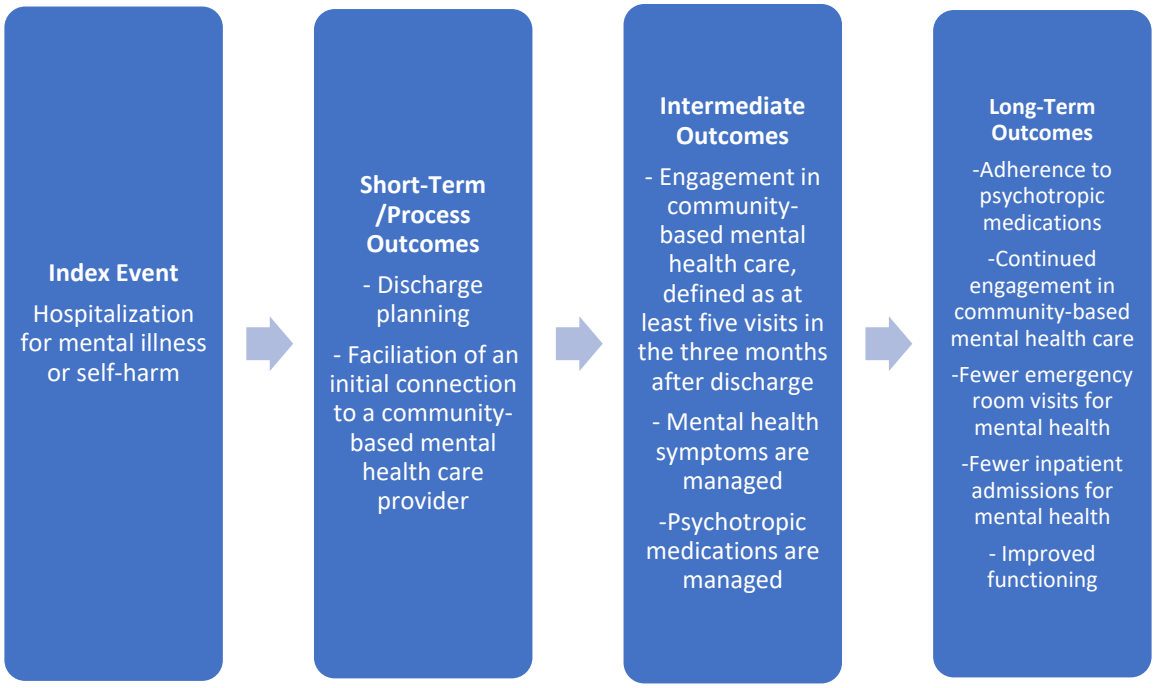
Updated evidence information here.

Previous (Year) Submission:

Evidence from the previous submission here.

1a.01) Provide a logic model.

Briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.



After a person is hospitalized for mental illness or self-harm, managed care plans and hospitals can engage in discharge planning and facilitate an initial condition to a community-based mental health care provider. The community-based mental health

care providers can, with the support of managed care plans, engage people in care and manage mental health symptoms and medications. This should lead to fewer inpatient admissions for mental health, fewer emergency department visits for mental health, continued engagement in community-based mental health care, psychotropic medication adherence, and improved functioning. The outcome being measured by this measure is the intermediate outcome "Engagement in community-based mental health care, defined as at least five visits in the three months after discharge."

1a.02) Select the type of source for the systematic review of the body of evidence that supports the performance measure.

A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data.

☒ Clinical Practice Guideline recommendation (with evidence review)
☐ US Preventive Services Task Force Recommendation
☐ Other systematic review and grading of the body of evidence (e.g., Cochrane Collaboration, AHRQ Evidence Practice Center)
☐ Other (please specify here:)

If the evidence is not based on a systematic review, skip to the end of the section and do not complete the repeatable question group below. If you wish to include more than one systematic review, you may add additional tables to the relevant sections. Please follow the 508 Checklist for tables.

Evidence - Systematic Reviews Table (Repeatable)

1. American Psychiatric Association (APA) Guideline- Schizophrenia

1a.03) Provide the title, author, date, citation (including page number) and URL for the systematic review.
Title: The American Psychiatric Association Practice Guideline for The Treatment of Patients with Schizophrenia, Third Edition
Author: American Psychiatric Association
Date:2019
Citation: American Psychiatric Association (2019). Practice Guideline for the Treatment of Patients With Schizophrenia Third Edition; 2019 Dec. 184 p.
URL:
https://psychiatryonline.org/doi/full/10.1176/appi.books.9780890424841.Schizophrenia03

1a.04) Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

Assessment and Determination of Treatment Plan

1. APA recommends (1C) that patients with schizophrenia have a documented, comprehensive, and person-centered treatment plan that includes evidence-based nonpharmacological and pharmacological treatments.

Pharmacotherapy

1. APA *recommends* (1A) that patients with schizophrenia be treated with an antipsychotic medication and monitored for effectiveness and side effects.*

*This guideline statement should be implemented in the context of a person-centered treatment plan that includes evidence-based nonpharmacological and pharmacological treatments for schizophrenia.

2. APA recommends (1A) that patients with schizophrenia whose symptoms have improved with an antipsychotic medication continue to be treated with an antipsychotic medication.*

Psychosocial Interventions

1. APA *recommends* (1B) that patients with schizophrenia who are experiencing a first episode of psychosis be treated in a coordinated specialty care program.*
2. APA *recommends* (1B) that patients with schizophrenia be treated with cognitive-behavioral therapy for psychosis (CBTp).*
3. APA *recommends* (1B) that patients with schizophrenia receive supported employment services.*
4. APA *recommends* (1B) that patients with schizophrenia receive assertive community treatment if there is a history of poor engagement with services leading to frequent relapse or social disruption (e.g., homelessness; legal difficulties, including imprisonment).*
5. APA *suggests* (2C) that patients with schizophrenia receive interventions aimed at developing self-management skills and enhancing person-oriented recovery.*
6. APA *suggests* (2C) that patients with schizophrenia who have a therapeutic goal of enhanced social functioning receive social skills training.*
7. APA suggests (2C) that patients with schizophrenia be treated with supportive psychotherapy.

1a.05) Provide the grade assigned to the evidence associated with the recommendation and include the definition of the grade.

"Strength of supporting research evidence describes the level of confidence that findings from scientific observation and testing of an effect of an intervention reflect the true effect. Confidence is enhanced by such factors as rigorous study design and minimal potential for study bias. Ratings were determined, in accordance with the AHRQ's "Methods Guide for Effectiveness and Comparative Effectiveness Reviews" (Agency for Healthcare Research and Quality 2014), by the methodologist (L. J. F.) and reviewed by members of the SRG and GWG. Available clinical trials were assessed across four primary domains: risk of bias, consistency of findings across studies, directness of the effect on a specific health outcome, and precision of the estimate of effect.

The ratings are defined as follows:

- High (denoted by the letter A) = high confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
- Moderate (denoted by the letter B)= moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.

- Low (denoted by the letter C) = low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of effect and is likely to change the estimate."

1a.06) Provide all other grades and definitions from the evidence grading system.
All grades were given in the response to the previous question.

1a.07) Provide the grade assigned to the recommendation, with definition of the grade.
Guidelines are either recommend or suggested.
"A recommendation (denoted by the numeral 1 after the guideline statement) indicates confidence that the benefits of the intervention clearly outweigh harms. A suggestion (denoted by the numeral 2 after the guideline statement) indicates greater uncertainty: although the benefits of the statement are still viewed as outweighing the harms, the balance of benefits and harms is more difficult to judge, or the benefits or the harms may be less clear. With a suggestion, patient values and preferences may be more variable, and this can influence the clinical decision that is ultimately made.
Each guideline statement also has an associated rating for the strength of supporting research evidence. Three ratings are used: high, moderate, and low (denoted by the letters A, B, and C, respectively). These ratings reflect the level of confidence that the evidence for a guideline statement reflects a true effect based on consistency of findings across studies, directness of the effect on a specific health outcome, precision of the estimate of effect, and risk of bias in available studies (Agency for Healthcare Research and Quality 2014; Balshem et al. 2011; Guyatt et al. 2006)."

1a.08) Provide all other grades and definitions from the recommendation grading system.
All grades and definitions were given in the response to the previous question.

1a.09) Detail the quantity (how many studies) and quality (the type of studies) of the evidence.
"The Agency for Healthcare Research and Quality's (AHRQ) systematic review *Treatments for Schizophrenia in Adults* (McDonagh et al. 2017) served as the predominant source of information for this guideline. Databases that were searched are Ovid MEDLINE® (PubMed®), the Cochrane Central Register of Controlled Trials, the Cochrane Database of Systematic Reviews, and PsycINFO®. Results were limited to English-language, adult (18 and older), and human-only studies."

"The AHRQ review (McDonagh et al. 2017) adhered to the procedures outlined in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews (Agency for Healthcare Research and Quality 2014). Recent, comprehensive, good- or fair-quality systematic reviews served as a primary source of evidence, supplemented by information from randomized controlled trials (RCTs) published since the systematic reviews or when no systematic reviews were available. For assessment of harms of treatment, systematic reviews of observational trials were also included. Eligibility for inclusion and exclusion of articles adhered to preestablished criteria. Specifically, the AHRQ review included articles that had at least 12 weeks of follow-up and were conducted in outpatient settings in countries that were relevant to the United States' health care system. Articles that addressed benefits of treatment were

included if at least 90% of the sample had a diagnosis of schizophrenia (or schizophreniform disorder), with a schizophrenia spectrum disorder in at least 50% of the sample (minimum sample size > 50) for studies of harms of treatment. For key questions that related to antipsychotic treatment, all of the SGAs were included; for FGAs, studies on fluphenazine, haloperidol, and perphenazine were included. Only head-to-head com-parison studies were included. For studies of psychosocial and other nonpharmacological interventions, studies were included if they compared usual care, standard care, treatment as usual, or a waitlist control group to active treatment with assertive community treatment, cognitive adaptive training, cognitive-behavioral therapy, cognitive remediation, early interventions for first-episode psychosis, family interventions, intensive case management, illness self-management training, interventions for co-occurring schizophrenia and substance use, psychoeducation, social skills training, supported employment, or supportive psychotherapy."

"For key question 1 on antipsychotic treatment, 698 citations were identified, 519 of which were excluded on the basis of title and abstract review, yielding 179 full-text articles that were reviewed, of which 38 were included in the final AHRQ review. For key question 2 on psychosocial and other nonpharmacological interventions, 2,766 citations were identified, 1,871 of which were excluded on the basis of title and abstract review, yielding 895 full-text articles that were reviewed, of which 53 were included in the final AHRQ review."

1a.10) Provide the estimates of benefit, and consistency across studies.

Assessment and Determination of Treatment Plan

Benefits
"In an individual with a possible psychotic disorder, a detailed assessment is important in establishing a diagnosis, recognizing co-occurring conditions (including substance use disorders, other psychiatric disorders, and other physical health disorders), identifying psychosocial issues, and developing a plan of treatment that can reduce associated symptoms, morbidity, and mortality."

"Development and documentation of a comprehensive, person-centered treatment plan assures that the clinician has considered the available nonpharmacological and pharmacological options for treatment and has identified those treatments that are best suited to the needs of the individual patient, with a goal of improving overall outcome. It may also assist in forming a therapeutic relationship, eliciting patient preferences, permitting education about possible treatments, setting expectations for treatment, and establishing a framework for shared decision-making. Documentation of a treatment plan promotes accurate communication among all those caring for the patient and can serve as a re-minder of prior discussions about treatment."

"The potential benefits of this guideline statement were viewed as far outweighing the potential harms."

Pharmacotherapy

Benefits
"Use of an antipsychotic medication in the treatment of schizophrenia can improve positive and negative symptoms of psychosis (high strength of research evidence) and can also lead to

reductions in depression and improvements in quality of life and functioning (moderate strength of research evidence). A meta-analysis of double-blind, randomized, placebo-controlled trials showed a medium effect size for overall efficacy (Leucht et al. 2017), with the greatest effect on positive symptoms. The rates of achieving any response or a good response were also significantly greater in patients who received an antipsychotic medication. In addition, the proportion of individuals who dropped out of treatment for any reason and for lack of efficacy was significantly less in those who were treated with an antipsychotic medication. Research evidence from head-to-head comparison studies and network meta-analysis (McDonagh et al. 2017) showed no consistent evidence that favored a specific antipsychotic medication, with the possible exception of clozapine."

"Use of an antipsychotic medication that has already been associated with symptom response can maintain improvements in symptoms as well as promote enhanced functioning and quality of life (high strength of research evidence). Long-term treatment with an antipsychotic medication has also been associated with a reduction in mortality as compared with no antipsychotic treatment in individuals with schizophrenia. In contrast, discontinuation of antipsychotic treatment can be associated with increases in symptoms and risk of hospitalization and poorer long-term outcomes, including greater mortality in the long term (low strength of research evidence)."

"The potential benefits of this guideline statement were viewed as far outweighing the potential harms."

Psychosocial Interventions

Benefits
Benefits of the psychosocial interventions recommended or suggested in this guideline include lower mortality, lower rates of relapse, better quality of life, better global function, reduction in symptoms, improved satisfaction with mental health services, and greater likelihood of working, being in school, or living independently. The potential benefits of these interventions were viewed as far outweighing the potential harms.

1a.11) Indicate what, if any, harms were identified in the study.

Assessment and Determination of Treatment Plan

Harms
"Some individuals may become anxious, suspicious, or annoyed if asked multiple questions during the evaluation. This could interfere with the therapeutic relationship between the patient and the clinician. Another potential consequence is that time used to focus on a detailed assessment (as outlined in the Practice Guidelines for the Psychiatric Evaluation of Adults, 3rd edition; American Psychiatric Association 2016a) could reduce time available to address other issues of importance to the patient or of relevance to diagnosis and treatment planning."

Pharmacotherapy

Harms

"The harms of using an antipsychotic medication in the treatment of schizophrenia include sedation, side effects mediated through dopamine receptor blockade (e.g., acute dystonia, akathisia, parkinsonism, tardive syndromes, NMS, hyperprolactinemia), disturbances in sexual function, anticholinergic effects, weight gain, glucose abnormalities, hyperlipidemia, orthostatic hypotension, tachycardia, and QTc prolongation. Clozapine has additional harms associated with its use, including sialorrhea, seizures, neutropenia (which can be severe and life-threatening), myocarditis, and cardiomyopathy. Among the antipsychotic medications, there is variability in the rates at which each of these effects occurs, and no specific medication appears to be devoid of possible side effects."

"The harms of continuing use of an antipsychotic medication can vary depending on whether the patient is experiencing any significant side effects from the medication that would have long-term untoward effects. For patients whose medications are well tolerated, long-term risks include tar-dive syndromes from antipsychotic medications. For other patients, long-term risks will vary according to the specific side effect, with metabolic effects of antipsychotic medication serving as a possible contributor to long-term health risks. Some studies have raised concerns about changes in brain region volumes with antipsychotic treatment, but these findings are heterogeneous and in-consistent."

Psychosocial Interventions

Harms
Harms of the psychosocial interventions recommended or suggested in this guideline were not well delineated or systematically studied, but likely to be small.

1a.12) Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.

We are not aware of any studies conducted since the publication of this guideline that would change the recommendations of this guideline.

2. American Psychiatric Association (APA) Guidelines-Bipolar Disorder

1a.03) Provide the title, author, date, citation (including page number) and URL for the systematic review.

Title: Practice Guideline for the Treatment of Patients with Bipolar Disorder, Second Edition
Author: American Psychiatric Association
Date: 2002
Citation: American Psychiatric Association (2002) Practice Guideline for the Treatment of Patients with Bipolar Disorder, Second Edition; 2002 Apr. 82 p.
URL:
https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/bipolar.pdf

1a.04) Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

Psychiatric Management
"Specific goals of psychiatric management include establishing and maintaining a therapeutic alliance, monitoring the patient's psychiatric status, providing education regarding bipolar disorder, enhancing treatment compliance, promoting regular patterns of activity and of sleep, anticipating stressors, identifying new episodes early, and minimizing functional impairments [I] (Recommended with substantial clinical confidence)."
Maintenance Treatment
"Following remission of an acute episode, patients may remain at particularly high risk of relapse for a period of up to 6 months; this phase of treatment, sometimes referred to as continuation treatment, is considered in this guideline to be part of the maintenance phase. Maintenance regimens of medication are recommended following a manic episode [I] (Recommended with substantial clinical confidence)."
"During maintenance treatment, patients with bipolar disorder are likely to benefit from a concomitant psychosocial intervention—including psychotherapy—that addresses illness management (i.e., adherence, lifestyle changes, and early detection of prodromal symptoms) and interpersonal difficulties [II] (Recommended with moderate clinical confidence)."

1a.05) Provide the grade assigned to the evidence associated with the recommendation and include the definition of the grade.

"This practice guideline is based on available evidence and clinical consensus and offers treatment recommendations to help psychiatrists develop plans for the care of adult patients with bipolar disorder."

Each reference cited in the guideline was assigned a letter corresponding to the nature of the supporting evidence. The assigned letters are:
[A] Randomized clinical trial. A study of an intervention in which subjects are prospectively followed over time; there are treatment and control groups; subjects are randomly assigned to the two groups; both the subjects and the investigators are blind to the assignments.
[B] Clinical trial. A prospective study in which an intervention is made and the results of that intervention are tracked longitudinally; study does not meet standards for a randomized clinical trial.
[C] Cohort or longitudinal study. A study in which subjects are prospectively followed over time without any specific intervention.
[D] Control study. A study in which a group of patients and a group of control subjects are identified in the present and information about them is pursued retrospectively or backward in time.
[E] Review with secondary data analysis. A structured analytic review of existing data, e.g., a meta-analysis or a decision analysis.
[F] Review. A qualitative review and discussion of previously published literature without a quantitative synthesis of the data.
[G] Other. Textbooks, expert opinion, case reports, and other reports not included above.

1a.06) Provide all other grades and definitions from the evidence grading system.

All grades and definitions from the evidence grading system were given in the previous question.

1a.07) Provide the grade assigned to the recommendation, with definition of the grade.

"The recommendations are based on the best available data and clinical consensus with regard to a particular clinical decision. The summary of treatment recommendations is keyed according to the level of confidence with which each recommendation is made. In addition, each reference is followed by a letter code in brackets that indicates the nature of the supporting evidence."

[I] Recommended with substantial clinical confidence.
[II] Recommended with moderate clinical confidence.

1a.08) Provide all other grades and definitions from the recommendation grading system.

"Each recommendation is identified as falling into one of three categories of endorsement, indicated by a bracketed Roman numeral following the statement. The three categories represent varying levels of clinical confidence regarding the recommendation:

[I] Recommended with substantial clinical confidence.
[II] Recommended with moderate clinical confidence.
[III] May be recommended on the basis of individual circumstances."

1a.09) Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

"A computerized search of the relevant literature from MEDLINE and PsycINFO was conducted. Sources of funding were not considered when reviewing the literature. The first literature search was conducted by searching MEDLINE and PsycINFO for the period from 1992 to 2000. Key words used were "bipolar disorder," "bipolar depression," "mania," "mixed states," "mixed episodes," "mixed mania," "antimanic," "hypomanic," "hypomania," "manic depression," "prophylactic," "pharmacotherapy," "mood stabilizers," "mood-stabilizing," "rapid cycling," "maintenance," "continuation," "child and adolescent," "antidepressants," "valproate," "lithium," "carbamazepine," "olanzapine," "risperidone," "gabapentin," "topiramate," "lamotrigine," "clonazepam," "divalproex," "psychotherapy," "family therapy," "psychoeducation," "course," "epidemiology," "comorbidity," "anxiety," "anxiety disorders," "attention deficit," "catatonia," "elderly," "family history," "gender," "general medical conditions," "life events," "personality disorders," "pregnancy," "psychosis," "stress," "substance-related disorders," "suicide," "homicide," and "violence."
A total of 3,382 citations were found.

An additional MEDLINE search for the period from 1992 to 2001 used the key words "genetic counseling," "family functioning," "cross-cultural issues," and "pharmacokinetics." A total of 122 citations were found. A search on PubMed was also conducted through 2001 that used the search terms "electroconvulsive," "intravenous drug abuse," "treatment response," "pharmacogenetic," "attention deficit disorder," "violence," "aggression," "aggressive," "suicidal," "cognitive impairment," "sleep," "postpartum," "ethnic," "racial," "metabolism," "hyperparathyroidism," "overdose," "toxicity," "intoxication," "pregnancy," "breast-feeding," and "lactation."

Additional, less formal, literature searches were conducted by APA staff and individual members of the work group on bipolar disorder."

1a.10) Provide the estimates of benefit, and consistency across studies.

"Patients frequently seek treatment during an acute episode, which may be characterized by depression, mania, hypomania, or a mixture of depressive and manic features. Treatment is aimed at stabilization of the episode with the goal of achieving remission, defined as a complete return to baseline level of functioning and a virtual lack of symptoms. (Following remission of a depressive episode, patients may remain at particularly high risk of relapse for a period up to 6 months; this phase of treatment, sometimes referred to as continuation treatment, is considered in this guideline to be part of maintenance treatment.) After successfully completing the acute phase of treatment, patients enter the maintenance phase. At this point, the primary goal of treatment is to optimize protection against recurrence of depressive, mixed, manic, or hypomanic episodes. Concurrently, attention needs to be devoted to maximizing patient functioning and minimizing subthreshold symptoms and adverse effects of treatment."

1a.11) Indicate what, if any, harms were identified in the study.

Harms identified were side effects of recommended medications, such as cognitive problems (e.g., dulling, impaired memory, poor concentration, confusion, mental slowness), tremor, sedation or lethargy, impaired coordination, and gastrointestinal distress.

1a.12) Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.

Many studies relevant to the treatment of bipolar disorder have been conducted since the guideline was published in 2002. The studies do not change the high-level recommendations of this guideline. The guideline has not been updated.

3. American Psychiatric Association (APA) Guidelines-Major Depressive Disorder

1a.03) Provide the title, author, date, citation (including page number) and URL for the systematic review.

Title: Practice Guideline for the Treatment of Patients with Major Depressive Disorder, Third Edition
Author: American Psychiatric Association
Date: 2010
Citation: American Psychiatric Association (2010); Practice Guideline for the Treatment of Patients With Major Depressive Disorder, Third Edition. 2010 Oct. 151 p.
URL: http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf

1a.04) Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

Psychiatric Management
"Psychiatric management consists of a broad array of interventions and activities that psychiatrists should initiate and continue to provide to patients with major depressive disorder through all phases of treatment [I]."

Evaluate functional impairment and quality of life
"In developing a treatment plan, interventions should be aimed at maximizing the patient's level of functioning as well as helping the patient to set specific goals appropriate to his or her functional impairments and symptom severity [I]."

Monitor the patient's psychiatric status
"The patient's response to treatment should be carefully monitored [I]. Continued monitoring of co-occurring psychiatric and/or medical conditions is also essential to developing and refining a treatment plan for an individual patient [I]."

Acute Phase
"Treatment in the acute phase should be aimed at inducing remission of the major depressive episode and achieving a full return to the patient's baseline level of functioning [I]. Acute phase treatment may include pharmacotherapy, depression-focused psychotherapy, the combination of medications and psychotherapy, or other somatic therapies such as electroconvulsive therapy (ECT), transcranial magnetic stimulation (TMS), or light therapy, as described in the sections that follow. Selection of an initial treatment modality should be influenced by clinical features (e.g., severity of symptoms, presence of co-

occurring disorders or psychosocial stressors) as well as other factors (e.g., patient preference, prior treatment experiences) [I]. Any treatment should be integrated with psychiatric management and any other treatments being provided for other diagnoses [I]."
"The patient's response to treatment should be carefully monitored [I]. Continued monitoring of co-occurring psychiatric and/or medical conditions is also essential to developing and refining a treatment plan for an individual patient [I]."

"As with patients who are receiving pharmacotherapy, patients receiving psychotherapy should be carefully and systematically monitored on a regular basis to assess their response to treatment and assess patient safety [I]."

1a.05) Provide the grade assigned to the evidence associated with the recommendation and include the definition of the grade.

"In the listing of cited references, each reference is followed by a letter code in brackets that indicates the nature of the supporting evidence."
"The following coding system is used to indicate the nature of the supporting evidence in the references:
[A] Randomized double-blind clinical trial. A study of an intervention in which subjects are prospectively followed over time, there are treatment and control groups, subjects are randomly assigned to the two groups, both the subjects and the investigators are blind to the assignments.
[A−] Randomized clinical trial. Same as above but not double-blind.
[B] Clinical trial. A prospective study in which an intervention is made and the results of that intervention are tracked longitudinally; study does not meet standards for a randomized clinical trial.
[C] Cohort or longitudinal study. A study in which subjects are prospectively followed over time without any specific intervention.
[D] Case-control study. A study in which a group of patients and a group of control subjects are identified in the present and information about them is pursued retrospectively or backward in time.
[E] Review with secondary data analysis. A structured analytic review of existing data, e.g., a meta-analysis or a decision analysis.
[F] Review. A qualitative review and discussion of previously published literature without a quantitative synthesis of the data.
[G] Other. Textbooks, expert opinion, case reports, and other reports not included above.

1a.06) Provide all other grades and definitions from the evidence grading system.

All grades and definitions from the evidence grading system were given in the previous question.

1a.07) Provide the grade assigned to the recommendation, with definition of the grade.

"In order for the reader to appreciate the evidence base behind the guideline recommendations and the weight that should be given to each recommendation, the summary of treatment recommendations is keyed according to the level of confidence with which each recommendation is made. Each rating of clinical confidence considers the strength of the available evidence. When evidence from randomized controlled trials and meta-analyses is limited, the level of confidence may also incorporate other clinical trials and case reports as well as clinical consensus with regard to a particular clinical decision."
[I] Recommended with substantial clinical confidence.

1a.08) Provide all other grades and definitions from the recommendation grading system.

"Each recommendation is identified as falling into one of three categories of endorsement, indicated by a bracketed Roman numeral following the statement. The three categories represent varying levels of clinical confidence:

[I] Recommended with substantial clinical confidence
[II] Recommended with moderate clinical confidence
[III] May be recommended on the basis of individual circumstances"

1a.09) Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

"Relevant updates to the literature were identified through a MEDLINE literature search for articles published since the second edition of the guideline, published in 2000. For this edition of the guideline, literature was identified through a computerized search of MEDLINE, using PubMed, for the period from January 1999 to December 2006. Using the MeSH headings depression or depressive disorder, as well as the key words major depression, major depressive disorder, neurotic depression, neurotic depressive, dysthymia, dysthymic, dysthymic disorder, endogenous depression, endogenous depressive, melancholia, melancholic, psychotic depression, atypical depression, seasonal depression, postpartum depression, postpartum depressive symptoms, unipolar depression, unipolar depressive, or pseudodementia yielded 39,157 citations. An additional 8,272 citations were identified by using the key words depression or depressive in combination with the MeSH headings affective disorders or psychotic or the key words psychosis, psychotic, catatonic, catatonia, mood disorder, mood disorders, affective disorder, or affective disorders. These citations were limited to English language articles on human treatments using the MeSH headings central nervous system stimulants, hypnotics and sedatives, anticonvulsants, tranquilizing agents, electric stimulation therapy, electroconvulsive therapy, psychotherapy, antidepressive agents, and monoamine oxidase inhibitors or the key words antidepressant, antidepressants, antidepressive, antidepressive agents, antidepressive agents, second generation, antidepressive agents tricyclic, antidepressive agents, tricyclic, fluoxetine, citalopram, escitalopram, paroxetine, sertraline, venlafaxine, duloxetine, mirtazapine, nefazodone, trazodone, imipramine, desipramine, nortriptyline, protriptyline, doxepin, trimipramine, amitriptyline, phenelzine, tranylcypromine, isocarboxazid, moclobemide, antipsychotic agents, testosterone, thyroid, tri iodothyronine, thyroxine, omega 3, s adenosyl methionine, s adenosylmethionine, St. John's wort, hypericum, selegiline, anticonvulsant, anticonvulsants, antipsychotic, antipsychotic agent, antianxiety, anti anxiety, benzodiazepine, benzodiazepines, zolpidem, sedative, sedatives, hypnotic, hypnotics, zaleplon, eszopiclone, valproate, valproic acid, divalproex, carbamazepine, oxcarbazepine, gabapentin, topiramate, lamotrigine, lithium, modafinil, methylphenidate, Adderall, amphetamine, amphetamines, dextroamphetamine, atomoxetine, electroconvulsive, vagal nerve stimulation, vagus nerve stimulation, VNS, rTMS, rapid transcranial magnetic, repetitive transcranial magnetic stimulation, magnetic stimulation, deep brain stimulation, psychotherapy, psychotherapeutic, psychotherapies, behavior therapy, behaviour therapy, cognitive therapy, cognitive behavior therapy, cognitive behavioral analysis system, cognitive behavioral therapy, cognitive behaviour therapy, cognitive behavioural therapy, psychoanalytic, interpersonal therapy, interpersonal psychotherapy, group therapy, family therapy, marital therapy, couples therapy, psychoanalysis, psychodynamic, aversive therapy, desensitization, exposure therapy, relaxation techniques, or progressive muscle relaxation. This yielded 13,506 abstracts, which were screened for relevance with a very modest threshold for inclusion, then reviewed by the Work Group."

"The Psychoanalytic Electronic Publishing database (http://www.p-e-p.org) was also searched using the terms major depression or major depressive. This search yielded 112 references. The Cochrane databases were also searched for the key word depression, and 168 meta-analyses were identified. Additional, less formal, literature searches were conducted by APA staff and individual Work Group members and included references through May 2009. Sources of funding were considered when the Work Group reviewed the literature."

1a.10) Provide the estimates of benefit, and consistency across studies.
Benefits identified were remission or reduction of depression symptoms, increase in functioning, improvement in quality of life, prevention of symptom relapse, and improved interpersonal relationships.

1a.11) Indicate what, if any, harms were identified in the study.
Harms identified were side effects of recommended medications, such as insomnia, weight gain, headaches, sedation, cardiovascular effects, increased risk of falls, and gastrointestinal distress.

Incomplete treatment response was another potential harm. "It is not uncommon for patients to have substantial but incomplete symptom reduction or improvement in functioning during acute phase treatment. A number of studies have provided compelling evidence that even mild residual symptoms at the end of a depressive episode are associated with significant psychosocial disability, compared with asymptomatic remission".

1a.12) Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.
Many studies relevant to the treatment of major depressive disorder have been conducted since the guideline was published in 2010. The studies do not change the high-level recommendations of this guideline. The guideline has not been updated.

4. Treatment for Suicidal Ideation, Self-Harm, and Suicide Attempts Among Youth

1a.03) Provide the title, author, date, citation (including page number) and URL for the systematic review.
Title: Treatment for Suicidal Ideation, Self-Harm, and Suicide Attempts Among Youth
Author: Substance Abuse and Mental Health Services Administration (SAMHSA)
Date: 2020
Citation: Substance Abuse and Mental Health Services Administration (SAMHSA): *Treatment for Suicidal Ideation, Self-harm, and Suicide Attempts Among Youth.* SAMHSA Publication No. PEP20-06-01-002 Rockville, MD: National Mental Health and Substance Use Policy Laboratory. Substance Abuse and Mental Health Services Administration, 2020.
URL: **https://store.samhsa.gov/sites/default/files/pep20-06-01-002.pdf**

1a.04) Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.
"Suicidal ideation, self-harm, and suicide attempts among youth are significant public health concerns. This review of the research literature identified practices and programs used to treat suicidal thoughts and behaviors. The chapter provides an overview of six evidence-based programs, including a discussion of the typical settings, demographic groups, intensity and duration, and outcomes attributed to receipt of the intervention:

- Dialectical Behavior Therapy (DBT) – Strong Evidence
- Attachment-Based Family Therapy (ABFT) – Moderate Evidence
- Multisystemic Therapy-Psychiatric (MST-Psych) – Moderate Evidence
- Safe Alternatives for Teens and Youth (SAFETY) – Moderate Evidence

- Integrated Cognitive Behavioral Therapy (I-CBT) – Moderate Evidence
- Youth-Nominated Support Team-Version II (YST-II) – Moderate Evidence"

1a.05) Provide the grade assigned to the evidence associated with the recommendation and include the definition of the grade.

"Trained reviewers assessed each study to ensure the methodology was rigorous and therefore could demonstrate causation between the programs and the identified outcomes. Reviewers reviewed and documented each study to ensure:

1. Experimental and comparison groups were statistically equivalent, with the only difference being that participants in the experimental group received the intervention and those in the comparison group received treatment as usual or no/minimal intervention.
2. For randomized experiments with high attrition and for quasi-experimental designs, baseline equivalence was established between the treatment and comparison groups.
3. For randomized experiments, randomization was not compromised. For example, ensuring that reassignment of treatment status, usually made to balance the distribution of background variables between treatment and control groups, did not occur.
4. Study did not have any confounding factors (factors that affect the outcome but are not accounted for by the study).
5. Missing data were addressed appropriately:
    - Imputation based on surrounding cases was considered valid.
    - Complete case analysis was considered valid and accounted for as attrition.
    - Using model with dummy for missing as a covariate was considered valid.
    - Assuming all missing data points are either positive or negative was not considered valid.
    - Regression-based imputation was considered valid; mean imputation was not considered valid.
6. Outcome measures were reliable, valid, and collected consistently from all participants.
7. Valid statistical models were used to estimate impacts.
8. Program demonstrated improved outcomes related to suicidal thoughts and behaviors.

Based on the study design and these study characteristics, reviewers gave each study a rating for causal impact. Reviewers used the following scoring metric for each study, based on the eight factors above, to determine if a study rated:

- High support of causal evidence
- Moderate support of causal evidence
- Low support of causal evidence

Only randomized controlled trials, quasi-experimental designs, and epidemiological studies with a strong comparison were eligible to receive a high or moderate study rating.

1a.06) Provide all other grades and definitions from the evidence grading system.

All grades and definitions from the evidence grading system were given in the previous question.

1a.07) Provide the grade assigned to the recommendation, with definition of the grade.

"The reviewers gave each program a rating based on the number of studies with strong, moderate, or emerging support of causal impact. Causal impact is evidence demonstrating that an intervention causes, or is responsible for, the outcome measured in the study's sample population.
The programs were placed into one of the following categories based on the level of causal evidence of its studies:

- Strong Evidence - Causal impact demonstrated by at least two randomized controlled trials, quasi-experimental designs, or epidemiological studies with a high or moderate rating.
- Moderate Evidence - Causal impact demonstrated by at least one randomized controlled trial, quasi-experimental design, or epidemiological study with a high or moderate rating.
- Emerging Evidence - No study received a high or a moderate rating. The program may have been evaluated with less rigorous studies (e.g., pre-post designs) that demonstrate an association between the program and positive outcomes, but additional studies are needed to establish causal impact."

1a.08) Provide all other grades and definitions from the recommendation grading system.

All grades and definitions from the evidence grading system were given in the previous question.

1a.09) Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

"Authors conducted a comprehensive review of published research for each selected intervention to determine its strength as an evidence-based practice. Eligible studies had to:

- Employ a randomized or quasi-experimental design, or
- Be a single sample pre-post design or an epidemiological study with a strong counterfactual (i.e., a study that analyzes what would have happened in the absence of the intervention).

Descriptive studies, implementation studies, and meta-analyses were not included in the review but were documented to provide context and identify implementation strengths and challenges for the programs."

Quantity of studies:

- Dialectical Behavior Therapy (DBT): 8 studies
- Attachment-Based Family Therapy (ABFT): 4 studies
- Multisystemic Therapy-Psychiatric (MST-Psych): 1 study
- Safe Alternatives for Teens and Youth (SAFETY): 2 studies
- Integrated Cognitive Behavioral Therapy (I-CBT): 1 study
- Youth-Nominated Support Team-Version II (YST-II): 2 studies

1a.10) Provide the estimates of benefit, and consistency across studies.

"Studies included in this evidence review demonstrated that use of DBT for youth with suicidal thoughts and behaviors was associated with reductions in one or more of the following outcomes:

- Suicidal ideation
- Self-harm (non-suicidal)
- Self-harm (intent unknown)
- Suicide attempts

The studies included several additional outcomes, including improvement in BPD, reduced psychiatric hospitalizations, reduced depressive symptoms, and improved treatment
completion rate"

"Some studies included in this evidence review demonstrated that use of ABFT for youth with suicidal thoughts and behaviors was associated with reductions in suicidal ideation.
Studies also demonstrated improved outcomes related to treatment retention, reduced depressive symptoms, and improved attachment related anxiety and avoidance."

"The study included in this evidence review demonstrated that use of MST-Psych for youth with suicidal thoughts and behaviors was associated with reductions in suicide attempts."

"Studies included in this evidence review demonstrated that use of SAFETY for youth with suicidal thoughts and behaviors was associated with reductions in:

- Suicidal ideation
- Self-harm (non-suicidal)
- Suicide attempts

One of the studies demonstrated significant reductions in depression and hopelessness for youth and significant reductions in depression for parents involved in the intervention."

"Studies included in this evidence review demonstrated that use of I-CBT for youth with suicidal thoughts and behaviors was associated with reductions in:

- Suicide attempts

The study also demonstrated reductions in the frequency of marijuana use and heavy drinking days, as well as in the number of inpatient hospitalizations, ED visits, and arrests."

"The study included in this evidence review demonstrated that use of YST-II for youth with suicidal thoughts and behaviors was associated with reductions in:

- Suicidal ideation

The study found youth who received YSTII attended more outpatient therapy and medication follow-up sessions and were more likely to participate in outpatient drug or alcohol treatment in the 12 months following their initial hospitalization.
The secondary analysis conducted more than a decade later found YST-II was associated with a reduction in mortality across all causes of death and a reduction in self-injury mortality due to either suicide or drug-related deaths with unknown intent"

1a.11) Indicate what, if any, harms were identified in the study.

Harms of the programs recommended in this guideline were not well delineated, but likely to be small.

1a.12) Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.

We are not aware of any studies conducted since the publication of this guideline that would change the recommendations of this guideline.

Evidence

1a.13) If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, describe the evidence on which you are basing the performance measure.

N/A, evidence is based on clinical practice guidelines.

1a.14) Briefly synthesize the evidence that supports the measure.

N/A, evidence is based on clinical practice guidelines.

1a.15) Detail the process used to identify the evidence.
N/A, evidence is based on clinical practice guidelines.

1a.16) Provide the citation(s) for the evidence.
N/A, evidence is based on clinical practice guidelines.

Importance to Measure and Report: Gap in Care/Disparities (1b.01 - 1b.05)

1b.01) Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care and list the benefits or improvements in quality envisioned by use of this measure.

The HEDIS measure Follow-Up After Hospitalization for Mental Illness captures one outpatient follow-up visit within seven and thirty days after psychiatric discharge. We think that it is important to also measure engagement in community-based mental health care over a longer time period post-discharge. People being discharged from a psychiatric hospitalization are in an acute phase of their mental health condition and need to receive more than one visit for adequate treatment.

This measure will improve the quality of care by encouraging both linkage to community-based mental health care services after an inpatient psychiatric discharge and engagement with community-based mental health services. Follow-up care after an inpatient psychiatric discharge has been associated with reduced inpatient readmissions, increased medication utilization, increased outpatient encounters, and increased functioning.

1b.02) Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

The source for the data presented below is NYS Medicaid administrative data for measurement years 2018 through 2021.

Table 1. Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization, Plan Performance Scores, 2018-2021 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	2018	2019	2020	2021
Number of plans	32	32	32	29
Number of discharges	49,875	52,079	44,929	46,898

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	2018	2019	2020	2021
Mean	46.1	46.4	45.5	46.9
Standard Deviation	12.0	12.0	12.0	13.1
Minimum	11.2	18.8	14.5	12.1
Maximum	66.4	66.7	61.6	63.6
Interquartile Range*	11.8	11.6	14.8	11.8
10th Percentile	25.0	25.5	24.1	21.3
20th Percentile	39.8	38.0	39.5	35.8
30th Percentile	43.6	44.0	40.7	46.2
40th Percentile	45.5	44.9	46.3	48.0
50th Percentile	50.0	49.2	48.8	51.3
60th Percentile	51.0	51.9	49.8	52.1
70th Percentile	52.1	53.3	52.2	53.3
80th Percentile	53.4	55.5	55.7	56.7
90th Percentile	57.5	60.4	58.3	60.2

1b.03) If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.

Performance data was reported in 1b.02.

1b.04) Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev,

min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., “topped out”, disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

The source for the data presented below is NYS Medicaid administrative data for measurement years 2018 through 2021.

Table 2a. Plan Measure Scores by Age Group, 2018 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Adults (Ages 21-64)	Children (Ages 6-20)
Number of plans	32	18
Number of discharges	37,361	12,514
Mean	43.6	51.2
Standard Deviation	11.5	20.8
Minimum	11.3	0.0
Maximum	66.4	69.2
Interquartile Range*	13.0	8.8
10th Percentile	25.5	0.0
20th Percentile	35.6	44.1
30th Percentile	39.8	53.6
40th Percentile	42.9	57.4
50th Percentile	45.9	59.1
60th Percentile	48.5	60.7
70th Percentile	50.6	61.6
80th Percentile	51.5	63.2
90th Percentile	52.8	69.0

Table 2b. Plan Measure Scores by Age Group, 2019 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Adults (Ages 21-64)	Children (Ages 6-20)
Number of plans	32	19
Number of discharges	39,026	13,053
Mean	43.5	64.7
Standard Deviation	12.2	17.7
Minimum	16.1	32.7
Maximum	66.7	100.0
Interquartile Range*	14.6	13.0
10th Percentile	23.1	48.8
20th Percentile	33.8	53.5
30th Percentile	38.5	56.3
40th Percentile	42.4	58.3
50th Percentile	45.6	59.9
60th Percentile	46.4	63.2
70th Percentile	50.1	65.2
80th Percentile	52.4	69.7
90th Percentile	56.0	100.0

Table 2c. Plan Measure Scores by Age Group, 2020 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Adults (Ages 21-64)	Children (Ages 6-20)
Number of plans	32	18
Number of discharges	34,588	10,341

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Adults (Ages 21-64)	Children (Ages 6-20)
Mean	42.9	58.3
Standard Deviation	11.8	9.4
Minimum	14.5	30.0
Maximum	63.8	70.0
Interquartile Range*	13.3	11.2
10th Percentile	22.3	50.0
20th Percentile	35.3	51.3
30th Percentile	39.7	56.7
40th Percentile	40.8	57.7
50th Percentile	44.4	58.6
60th Percentile	46.5	63.3
70th Percentile	48.8	63.7
80th Percentile	54.3	66.4
90th Percentile	56.5	68.1

Table 2d. Plan Measure Scores by Age Group, 2021 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Adults (Ages 21-64)	Children (Ages 6-20)
Number of plans	29	16
Number of discharges	35,573	11,325
Mean	43.5	65.9
Standard Deviation	12.3	12.8

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Adults (Ages 21-64)	Children (Ages 6-20)
Minimum	10.8	38.8
Maximum	60.2	100
Interquartile Range*	14.3	5.3
10th Percentile	21.2	50.0
20th Percentile	35.4	64.2
30th Percentile	42.7	64.7
40th Percentile	43.8	67.0
50th Percentile	46.5	67.6
60th Percentile	48.5	68.7
70th Percentile	51.6	69.7
80th Percentile	52.3	69.8
90th Percentile	56.6	70.9

Table 3a. Plan Measure Scores by Ethnicity, 2018 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
Number of plans	32	32
Number of discharges	10,374	37,054
Mean	50.6	44.9
Standard Deviation	14.8	12.7
Minimum	14.1	10.2
Maximum	83.3	65.5

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
Interquartile Range*	15.0	12.8
10th Percentile	34.6	26.2
20th Percentile	40.9	37.8
30th Percentile	43.9	43.5
40th Percentile	45.5	45.4
50th Percentile	52.3	47.7
60th Percentile	53.7	50.5
70th Percentile	57.7	51.8
80th Percentile	61.1	53.7
90th Percentile	69.2	56.3

Table 3b. Plan Measure Scores by Ethnicity, 2019 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
Number of plans	32	32
Number of discharges	9,446	39,175
Mean	46.6	45.4
Standard Deviation	11.6	12.9
Minimum	18.2	11.9
Maximum	76.0	67.7
Interquartile Range*	13.0	12.8
10th Percentile	32.4	26.4

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
20th Percentile	39.7	36.5
30th Percentile	41.0	41.6
40th Percentile	41.9	44.5
50th Percentile	47.5	49.4
60th Percentile	50.4	51.1
70th Percentile	52.6	51.9
80th Percentile	54.5	56.0
90th Percentile	56.6	58.5

Table 3c. Plan Measure Scores by Ethnicity, 2020 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
Number of plans	32	32
Number of discharges	7,567	34,050
Mean	48.4	44.8
Standard Deviation	16.3	12.0
Minimum	11.1	11.1
Maximum	82.8	62.8
Interquartile Range*	14.3	15.6
10th Percentile	24.8	29.0
20th Percentile	41.3	36.5
30th Percentile	43.2	40.8

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
40th Percentile	46.2	44.8
50th Percentile	47.2	47.4
60th Percentile	50.0	48.4
70th Percentile	55.6	53.1
80th Percentile	57.1	55.1
90th Percentile	71.2	58.7

Table 3d. Plan Measure Scores by Ethnicity, 2021 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
Number of plans	29	29
Number of discharges	7,746	35,497
Mean	47.4	46.5
Standard Deviation	16.1	12.9
Minimum	12.5	10.4
Maximum	73.9	63.6
Interquartile Range*	18.9	12.6
10th Percentile	16.7	25.0
20th Percentile	38.3	34.4
30th Percentile	39.6	43.3
40th Percentile	46.2	48.9
50th Percentile	51.9	50.6

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Hispanic	Not Hispanic
60th Percentile	54.6	52.3
70th Percentile	57.5	53.3
80th Percentile	60.0	57.1
90th Percentile	64.6	60.7

Table 4a. Plan Measure Scores by Race, 2018 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
Number of plans	26	28	32	27	32	32
Number of discharges	492	2,349	17,477	303	23,515	2,355
Mean	51.5	46.5	39.3	34.1	50.9	46.8
Standard Deviation	29.7	27.1	11.3	31.3	11.6	21.8
Minimum	0.0	0.0	11.4	0.0	12.9	0.0
Maximum	100.0	100.0	59.6	100.0	72.6	100.0
Interquartile Range*	32.1	24.1	11.2	60.9	9.6	16.4
10th Percentile	0.0	16.7	21.0	0.0	39.5	21.1
20th Percentile	28.6	29.3	34.1	0.0	43.3	37.5
30th Percentile	37.1	33.3	35.5	0.0	50.1	40.0
40th Percentile	47.4	35.9	37.1	22.0	51.4	42.9
50th Percentile	54.6	42.6	41.5	29.3	53.8	47.8

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
60th Percentile	56.5	50.0	44.7	44.4	55.8	50.0
70th Percentile	58.3	51.7	45.7	50.0	56.1	52.9
80th Percentile	66.7	59.0	46.7	63.9	57.1	59.4
90th Percentile	100.0	100.0	50.8	70.0	61.1	66.7

Table 4b. Plan Measure Scores by Race, 2019 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
Number of plans	26	31	32	26	32	31
Number of discharges	525	2,441	17,832	353	24,622	2,263
Mean	42.8	55.6	38.9	54.7	51.1	45.3
Standard Deviation	25.5	25.1	11.0	30.7	11.2	23.1
Minimum	0.0	0.0	11.8	0.0	19.0	0.0
Maximum	100.0	100.0	55.0	100.0	75.3	100.0
Interquartile Range*	31.1	26.7	15.1	55.0	9.2	18.1
10th Percentile	0.0	31.5	22.1	20.0	32.4	14.3
20th Percentile	22.2	33.3	32.1	25.0	44.9	31.0
30th Percentile	33.6	44.8	32.7	38.0	48.3	39.0

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
40th Percentile	40.0	50.0	37.5	42.9	50.7	43.9
50th Percentile	50.0	52.2	39.8	50.0	53.4	45.9
60th Percentile	51.3	53.8	43.1	55.6	54.8	50.0
70th Percentile	55.6	61.5	46.9	66.7	56.6	50.0
80th Percentile	58.6	70.0	48.2	100.0	57.7	52.9
90th Percentile	66.7	100.0	51.4	100.0	60.9	71.4

Table 4c. Plan Measure Scores by Race, 2020 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
Number of plans	26	30	32	24	32	32
Number of discharges	518	2,152	15,616	280	20,692	1,807
Mean	43.9	46.3	39.8	43.6	50.0	40.4
Standard Deviation	31.4	23.7	11.4	32.4	12.6	25.4
Minimum	0.0	0.0	11.1	0.0	11.8	0.0
Maximum	100.0	100.0	59.6	100.0	69.6	100.0
Interquartile Range*	42.7	16.7	12.6	45.2	11.8	26.6
10th Percentile	0.0	7.6	22.1	0.0	34.8	0.0

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
20th Percentile	16.7	36.2	31.9	17.6	45.3	17.6
30th Percentile	33.3	40.0	36.0	20.0	47.5	26.7
40th Percentile	33.3	44.7	38.2	33.3	50.0	33.3
50th Percentile	37.5	46.6	41.9	34.8	52.0	42.6
60th Percentile	45.0	49.3	44.4	42.9	53.5	48.7
70th Percentile	56.2	54.0	46.3	56.7	57.5	50.7
80th Percentile	70.8	58.4	49.5	66.7	58.5	55.6
90th Percentile	100.0	73.3	51.0	100.0	61.3	66.7

Table 4d. Plan Measure Scores by Race, 2021 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
Number of plans	24	28	29	21	29	29
Number of discharges	486	2,373	15,907	270	21,611	1,882
Mean	42.0	47.9	41.0	55.6	50.5	45.1
Standard Deviation	24.2	23.8	9.7	33.5	16.2	22.4
Minimum	0.0	0.0	18.9	0.0	4.5	0.0
Maximum	100.0	100.0	58.3	100.0	67.9	100.0
Interquartile	29.3	24.2	11.8	49.5	14.5	22.7

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	American Indian/Alaska Native	Asian	Black	Native Hawaiian/ Other Pacific Islander	White	Two or More Races
Range*						
10th Percentile	0.0	5.9	22.1	18.8	19.6	0.0
20th Percentile	20.0	30.0	34.4	25.0	41.3	29.0
30th Percentile	33.3	39.5	35.5	33.3	51.9	39.5
40th Percentile	36.0	48.8	41.2	44.4	53.4	43.8
50th Percentile	46.4	50.7	43.2	50.0	54.4	49.8
60th Percentile	51.7	53.5	43.9	65.1	57.4	50.0
70th Percentile	55.6	58.2	45.5	75.0	60.2	53.2
80th Percentile	57.1	64.7	49.1	100.0	61.9	58.5
90th Percentile	65.5	75.8	52.7	100.0	64.1	71.4

Table 5a. Plan Measure Scores by Sex, 2018 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
Number of plans	32	32
Number of discharges	22,815	27,058
Mean	49.3	43.4
Standard Deviation	11.6	12.4
Minimum	19.4	8.4
Maximum	68.8	67.6
Interquartile Range*	12.7	10.1

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
10th Percentile	29.4	23.2
20th Percentile	40.8	36.1
30th Percentile	44.6	40.0
40th Percentile	50.7	42.6
50th Percentile	53.1	45.2
60th Percentile	54.7	47.5
70th Percentile	56.1	48.8
80th Percentile	58.7	49.6
90th Percentile	60.3	57.1

Table 5b. Plan Measure Scores by Sex, 2019 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
Number of plans	32	32
Number of discharges	24,079	28,000
Mean	50.0	42.8
Standard Deviation	11.9	12.9
Minimum	18.2	16.8
Maximum	67.9	69.0
Interquartile Range*	14.1	14.6
10th Percentile	37.8	22.8
20th Percentile	42.1	32.7

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
30th Percentile	48.7	38.6
40th Percentile	49.8	40.2
50th Percentile	50.9	42.7
60th Percentile	54.1	47.1
70th Percentile	56.3	49.4
80th Percentile	59.8	52.6
90th Percentile	63.3	58.4

Table 5c. Plan Measure Scores by Sex, 2020 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
Number of plans	32	32
Number of discharges	20,802	24,127
Mean	51.3	40.3
Standard Deviation	13.2	11.1
Minimum	18.8	10.8
Maximum	72.4	58.0
Interquartile Range*	16.8	10.2
10th Percentile	31.5	21.9
20th Percentile	41.0	33.1
30th Percentile	47.8	37.8
40th Percentile	50.8	39.3

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
50th Percentile	54.3	41.8
60th Percentile	57.6	44.9
70th Percentile	59.4	46.0
80th Percentile	61.0	50.0
90th Percentile	64.5	51.8

Table 5d. Plan Measure Scores by Sex, 2021 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
Number of plans	29	29
Number of discharges	22,582	24,312
Mean	51.8	42.4
Standard Deviation	13.2	12.4
Minimum	22.6	6.7
Maximum	67.5	58.8
Interquartile Range*	18.3	9.5
10th Percentile	23.8	20.9
20th Percentile	40.6	33.0
30th Percentile	49.3	40.8
40th Percentile	55.1	44.5
50th Percentile	57.1	45.3
60th Percentile	59.2	45.9

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Female	Male
70th Percentile	60.0	48.2
80th Percentile	61.6	49.8
90th Percentile	63.7	56.9

Table 6a. Plan Measure Scores by Primary Diagnosis at Discharge, 2018 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
Number of plans	32	32	32	32
Number of discharges	9,171	16,167	19,022	5,515
Mean	48.4	48.6	43.3	43.5
Standard Deviation	14.2	14.3	12.5	14.5
Minimum	7.8	17.5	9.6	8.7
Maximum	73.1	84.6	68.0	80.0
Interquartile Range*	16.3	21.6	9.6	17.1
10th Percentile	26.3	30.0	22.4	25.8
20th Percentile	39.1	34.7	37.0	32.6
30th Percentile	43.4	40.3	40.8	36.3
40th Percentile	49.3	44.9	43.1	43.8
50th Percentile	52.1	49.2	44.0	44.9
60th Percentile	54.3	54.8	46.8	49.0

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
70th Percentile	56.6	56.9	48.3	50.6
80th Percentile	59.1	60.6	49.2	51.6
90th Percentile	62.2	64.7	57.8	57.1

Table 6b. Plan Measure Scores by Primary Diagnosis at Discharge, 2019 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
Number of plans	32	32	32	32
Number of discharges	9,301	15,710	20,258	6,810
Mean	47.4	49.1	44.1	43.2
Standard Deviation	14.7	14.6	12.6	13.7
Minimum	0.0	15.4	17.7	0.0
Maximum	77.3	90.9	65.8	66.7
Interquartile Range*	15.1	13.1	15.1	16.3
10th Percentile	29.5	32.8	26.0	33.0
20th Percentile	38.9	39.9	34.0	34.7
30th Percentile	42.7	44.8	39.5	35.6
40th Percentile	47.5	45.0	41.7	43.1
50th Percentile	49.4	50.8	43.9	45.9
60th Percentile	51.3	52.9	46.6	48.2

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
70th Percentile	55.1	55.4	50.4	50.9
80th Percentile	56.6	58.7	55.4	53.7
90th Percentile	59.1	62.4	62.7	57.2

Table 6c. Plan Measure Scores by Primary Diagnosis at Discharge, 2020 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
Number of plans	32	32	32	32
Number of discharges	8,118	12,627	18,732	5,452
Mean	49.3	49.4	42.0	43.2
Standard Deviation	16.7	13.4	12.2	14.9
Minimum	12.5	15.8	12.1	0.0
Maximum	92.3	72.0	66.7	78.3
Interquartile Range*	18.3	15.5	12.1	19.9
10th Percentile	25.7	29.3	22.7	26.7
20th Percentile	38.9	37.1	34.8	31.8
30th Percentile	40.0	46.4	38.1	33.3
40th Percentile	47.1	49.4	41.1	40.2
50th Percentile	49.9	51.9	42.5	45.3
60th Percentile	51.9	57.3	44.4	48.7

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
70th Percentile	54.6	58.3	47.8	51.7
80th Percentile	58.4	59.6	51.6	52.6
90th Percentile	70.3	61.0	57.4	57.1

Table 6d. Plan Measure Scores by Primary Diagnosis at Discharge, 2021 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
Number of plans	29	29	29	29
Number of discharges	8,333	13,342	19,612	5,611
Mean	50.5	51.0	41.7	45.1
Standard Deviation	15.7	16.3	11.0	16.4
Minimum	9.1	6.3	16.7	0.0
Maximum	77.8	73.6	58.4	71.4
Interquartile Range*	21.1	19.4	10.9	10.5
10th Percentile	26.2	21.6	21.1	14.3
20th Percentile	38.2	41.5	33.2	37.4
30th Percentile	46.3	44.3	37.7	45.6
40th Percentile	53.1	48.5	42.5	46.2
50th Percentile	54.5	55.6	44.9	47.7
60th Percentile	56.1	59.6	45.5	51.7

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Bipolar Disorder	Depressive Disorders	Schizophrenia Spectrum and Other Psychotic Disorders	Other
70th Percentile	59.4	62.5	47.3	53.8
80th Percentile	63.1	64.4	52.0	57.4
90th Percentile	65.8	67.2	53.0	61.9

1b.05) If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.

Performance data was provided in 1b.04.

Scientific Acceptability: Maintenance (2ma.01 - 2ma.04)

2ma.01) Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

☐ Yes
☐ No

2ma.02) Indicate whether additional empirical validity testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

☐ Yes
☐ No

2ma.03) For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

☐ Yes
☐ No

2ma.04) For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

☐ Yes - Additional risk adjustment analysis is included
☐ No additional risk adjustment analysis included

Scientific Acceptability: Reliability - Testing (2a.01 - 2a.12)

Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact Battelle staff at PQMsupport@battelle.org about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact Battelle staff at PQMsupport@battelle.org with any questions.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the 2021 Measure Evaluation Criteria and Guidance.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet the evaluation criteria for testing.

2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument-based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;

AND

If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

2b3. For outcome measures and other measures when indicated (e.g., resource use):

• an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

• rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.

2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions

Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

Risk factors that influence outcomes should not be specified as exclusions.

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v.$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous (Year) Submission:

Testing from the previous submission here.

2a.01) Select only the data sources for which the measure is tested.

☐ Assessment Data
☒ Claims
☐ Electronic Health Data
☐ Electronic Health Records
☐ Instrument-Based Data
☐ Management Data
☐ Other (please specify here:)
☐ Paper Medical Records
☐ Registry Data

2a.02) If an existing dataset was used, identify the specific dataset.

The dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).
New York State administrative Medicaid data.

2a.03) Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"
1-1-2018 – 12-31-2018

2a.04) Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.

☐ Accountable Care Organization
☐ Clinician: Group/Practice
☐ Clinician: Individual
☐ Facility
☒ Health Plan

☐ Integrated Delivery System
☐ Other (specify)
☐ Population: Community, County or City
☐ Population: Regional and State

2a.05) List the measured entities included in the testing and analysis (by level of analysis and data source).

Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.
Thirty-one Medicaid managed care plans that operate in New York State were included in the analysis. Medicaid fee-for-service discharges were also included as a separate entity.

Question Response:
We believe that our testing done with NYS data may be representative of a national performance measure. New York is the fourth most populous state in the United States with over 20 million people. New York State is diverse in terms of race, ethnicity, severity of mental health conditions, and population density with both urban and rural regions. NYS has a broad mental health service system with representation from multiple types of mental health treatment services.

21% of the United States population is covered by Medicaid/CHIP and 28% of the NYS population is covered by Medicaid/CHIP. A similar proportion of the US (29%) and NYS (27%) populations are less than 200% below the federal poverty level, but NYS is one of 38 states adopting Medicaid expansion. Only four states have higher proportions of their populations covered by Medicaid (1). From the relatively high level of Medicaid coverage in NYS, it would be reasonable to assume that our analysis is not missing segments of the population that are covered by Medicaid in other states.
In the literature, the benefit of outpatient follow-up care after a psychiatric discharge has been seen across different populations. Attending outpatient follow-up care at least once after a psychiatric hospitalization was associated with reduced inpatient readmissions (2-4), increased medication adherence (5), increased retention in outpatient care (5), and higher levels of functioning (6) in studies of Medicaid, commercially insured, United States Veterans Health Administration, and national health system populations. The benefit of follow-up after psychiatric hospitalization is also supported by the nationally adopted HEDIS measure, Follow-Up After Hospitalization for Mental Illness. This measure is used for Medicaid and commercially insured populations throughout the United States. We think the clinical rationale and face validity of this measure would apply more generally to populations outside of NYS Medicaid.

1. Kaiser Family Foundation, Medicaid State Fact Sheets. https://www.kff.org/interactive/medicaid-state-fact-sheets/
2. Kurdyak P, Vigod SN, Newman A, Giannakeas V, Mulsant BH, Stukel T. Impact of Physician Follow-Up Care on Psychiatric Readmission Rates in a

Population-Based Sample of Patients With Schizophrenia. Psychiatr Serv. 2018;69(1):61-68.
3. Marcus SC, Chuang CC, Ng-Mak DS, Olfson M. Outpatient Follow-Up Care and Risk of Hospital Readmission in Schizophrenia and Bipolar Disorder. Psychiatr Serv. 2017;68(12):1239-1246.
4. Okumura Y, Sugiyama N, Noda T. Timely follow-up visits after psychiatric hospitalization and readmission in schizophrenia and bipolar disorder in Japan. Psychiatry Res. 2018;270:490-495.
5. Beadles CA, Ellis AR, Lichstein JC, et al. First outpatient follow-up after psychiatric hospitalization: does one size fit all?. Psychiatr Serv. 2015;66(4):364-372.
6. Greenberg GA, Rosenheck RA. Continuity of care and clinical outcomes in a national health system. Psychiatr Serv. 2005;56(4):427-433.

2a.06) Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

49,873 discharges were included in the analysis. A sample was not used; this number represents all Calendar Year 2018 discharges in the measure eligible population. The source for all data presented is NYS Medicaid administrative data.

Descriptive characteristics of the population included in the analyses are below.

Table 7. Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization, Age at Discharge of the Eligible Population, 2018 NYS Medicaid Discharges

Age Group	Frequency	Percent
<10	971	2.0
10-19	10,285	20.6
20-29	13,398	26.9
30-39	10,488	21.0
40-49	6,862	13.8
50-59	6,429	12.9
60-64	1,442	2.9

Table 8. Sex of the Eligible Population, 2018 NYS Medicaid Discharges

Sex	Frequency	Percent
Female	22,815	45.75
Male	27,058	54.25

Table 9. Ethnicity of the Eligible Population, 2018 NYS Medicaid Discharges

Ethnicity	Frequency	Percent
Hispanic or Latino	10,374	20.80
Non - Hispanic or Latino	37,054	74.29
Unknown	2,447	4.91

Table 10. Race of the Eligible Population, 2018 NYS Medicaid Discharges

Race Group	Frequency	Percent
American Indian/Alaska Native	492	0.99
Asian	2,349	4.71
Black/African American	17,477	35.04
Native Hawaiian/Other Pacific Islander	303	0.61
Unknown	3,384	6.78
White	23,515	47.15
Two or More Races	2,355	4.72

Table 11. Primary Diagnosis at Discharge of the Eligible Population, 2018 NYS Medicaid Discharges

Primary Diagnosis Category at Discharge	Frequency	Percent
Anxiety Disorders	8	0.02
Bipolar and Related Disorders	9,171	18.73

Primary Diagnosis Category at Discharge	Frequency	Percent
Depressive Disorders	16,167	33.02
Disruptive, Impulse-Control, and Conduct Disorders	863	1.76
Dissociative Disorders	12	0.02
Feeding and Eating Disorders	3	0.01
Medication-Induced Movement Disorders	1	0
Neurodevelopmental Disorders	764	1.56
Obsessive-Compulsive and Related Disorders	51	0.1
Personality Disorders	690	1.41
Schizophrenia Spectrum and Other Psychotic Disorders	19,022	38.85
Somatic Symptom and Related Disorders	17	0.03
Substance-Related and Addictive Disorders	8	0.02
Trauma- and Stressor-Related Disorders	2,184	4.46

2a.07) If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.

The data do not differ between different aspects of testing.

2a.08) List the social risk factors that were available and analyzed.

For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter “see validity testing section of data elements”; and enter “N/A” for 2a.11 and 2a.12.
Information on social risk factors is not available in NYS Medicaid administrative data. The NYS Office of Mental Health attempted to collect social risk factor information through an assessment for Health and Recovery Plan (HARP) members (a specialty

Medicaid product line for individuals with serious behavioral health conditions), but the assessment was completed for less than 20% of enrolled individuals annually. We do not feel that this data is representative enough of the measure eligible population to use in measure testing. NYS OMH is exploring other means of collecting social determinants of health information in the future, including an annual assessment for patients of NYS mental health outpatient clinics and/or through EHRs, so that this important data can be incorporated into performance measurement development.

2a.09) Select the level of reliability testing conducted.

Choose one or both levels.
☐ Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)

☒ Accountable Entity Level (e.g., signal-to-noise analysis)

2a.10) For each level of reliability testing checked above, describe the method of reliability testing and what it tests.

Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.

To test reliability, we performed signal to noise analysis using beta binomial models. Specifically, we used the methodology from "The Reliability of Provider Profiling, A Tutorial" by John L. Adams for the National Committee for Quality Assurance, RAND Health, 2009. This method tests how well the performance of one reporting entity can be distinguished from another. Conceptually, it is the ratio of signal to noise and is appropriate for testing the reliability of measures with a binary (yes/no, pass/fail, etc.) outcome. The signal is the proportion of variation attributable to plan performance and the noise is the proportion of variation attributable to measurement error.

A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance.

Our steps to calculate reliability are shown below. Health plans were used as the reporting entity.
The formula for signal-to-noise reliability is:
Signal-to-noise reliability = σ^2plan-to-plan / (σ^2plan-to-plan + σ^2error)
We estimated the variance between plans and variance within plans. The formulas for the variance calculations are:

1. Variance between plans = σ^2plan-to-plan = $(\alpha\ \beta) / (\alpha + \beta + 1)(\alpha + \beta)^2$
   α and β are two shape parameters of the Beta-Binomial distribution, $\alpha > 0$, $\beta > 0$
2. Variance within plans: σ^2error = $\hat{p}(1- \hat{p})/n$
   $\hat{p}$ = observed rate for the plan
   n = plan-specific denominator for the observed rate

We ran a SAS Macro program that was developed by John Adams and based on his tutorial mentioned above to calculate the reliability for each reporting entity. To derive the mean signal-to-noise reliability presented in Table 12, the average of the entity level reliability estimates was taken. The mean signal-to-noise reliability indicates the measure's mean ability to differentiate between reporting entity performance.

Table 13 shows the standard error (SE) and 95% confidence interval (95% CI) of the mean signal-to-noise reliability for all reporting entities. The reporting entities were divided into terciles based on the denominator size (number of eligible discharges per plan). The SE and 95% CI of the mean signal-to-noise reliability provides information about the stability of the reliability estimates. The formula for the 95% CIs is the mean signal-to-noise reliability ± (1.96*SE).

Table 14 shows the minimum, maximum, 10th, 25th, 50th, 75th, and 90th percentiles of the plan-level signal-to-noise reliability estimates. The reporting entities were divided into terciles based on the denominator size (number of eligible discharges per plan). Each plan's reliability estimate was calculated as described above.

2a.11) For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, Measure Evaluation Criteria).

Table 12. Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization, Point Estimate of Mean Signal-to-Noise Reliability, 2018 NYS Medicaid Discharges

Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization	Mean Signal-to-Noise Reliability
All NYS Medicaid Eligible Discharges	0.946

Table 13. Mean Signal-To-Noise Reliability, Standard Error (SE) and 95% Confidence Interval (95% CI) by Terciles of the Denominator Size, 2018 NYS Medicaid Discharges

Tercile	Number of Plans	Number of Eligible Discharges per Plan (min - max)	Mean Signal-To-Noise Reliability	SE	95% CI
All NYS Medicaid Eligible Discharges	**32**	**57-12467**	**0.946**	**0.011**	**(0.924, 0.967)**
Tercile 1	10	57-237	0.870	0.019	(0.833, 0.906)
Tercile 2	11	298-1026	0.969	0.004	(0.960, 0.977)
Tercile 3	11	1178-12467	0.992	0.002	(0.989, 0.995)

Table 14. Distribution of Plan-Level Signal-To-Noise Reliability by Terciles of the Denominator Size, 2018 NYS Medicaid Discharges

		Distribution of Plan Estimates of Signal-to-Noise Reliability						
Tercile	**Number of Plans**	**Min**	**P10**	**P25**	**P50**	**P75**	**P90**	**Max**
All NYS Medicaid Eligible Discharges	**32**	**0.756**	**0.853**	**0.926**	**0.976**	**0.987**	**0.996**	**0.999**
Tercile 1	10	0.756	0.759	0.836	0.881	0.926	0.928	0.928
Tercile 2	11	0.942	0.942	0.955	0.975	0.980	0.982	0.982
Tercile 3	11	0.985	0.985	0.987	0.992	0.996	0.998	0.999

2a.12) Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

A reliability of 0.70 is often considered adequate, 0.80 is generally described as good, and a reliability of greater than 0.90 is thought to be high.

The mean reliability estimate for NYS Medicaid plans was 0.946 with a 95% CI of 0.924 to 0.967, which falls into the high reliability category (Table 13). Across the terciles of plans, mean

reliability estimates ranged from 0.870, 95% CI (0.833-0.906) to 0.992, 95% CI (0.989-0.995), which indicate good to high reliability (Table 13). Plan level reliability increased as denominator size increased. The lowest reliability estimate among all NYS Medicaid plans was 0.756, which is considered adequate. All plans in the top two terciles of denominator size had reliability estimates above 0.90 and fell into the high reliability category.

Scientific Acceptability: Validity - Testing (2b.01 - 2b.04)

2b.01) Select the level of validity testing that was conducted.

☐ Patient or Encounter-Level (data element validity must address ALL critical data elements)

☒ Accountable Entity Level (e.g., hospitals, clinicians)

☒ Empirical validity testing of the measure score

☒ Systematic assessment of face validity of performance measure score as an indicator of quality or resource use (i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance)

2b.02) For each level of testing checked above, describe the method of validity testing and what it tests.

Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.

Construct validity, empirical validity, and face validity were tested for this measure.

To test construct validity, we calculated the Pearson correlation coefficient between Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization and the NCQA HEDIS measure Follow-Up After Hospitalization for Mental Illness. We expected to find a moderate positive correlation between the measures. The eligible population for these measures is similar and there is overlap between the group meeting the criteria for Follow-Up After Hospitalization for Mental Illness (one visit 7 or 30 days after discharge) and meeting the criteria for Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization. In our Calendar Year 2018 discharge cohort, 62% of discharges meeting the criteria for the 30-day Follow-Up After Hospitalization for Mental Illness had five community based mental health follow-up visits in the 90 days after discharge. The two measures are conceptually different, because the HEDIS measure is one of short-term follow-up and the NYS OMH measure is one of longer-term engagement in care. A moderate correlation is desirable in this situation because the two measures are similar but not identical.

The Pearson correlation measures the strength and direction of the linear relationship between two continuous variables. Correlation coefficients range from 1 to -1. A value of 1 represents a perfect positive correlation between two variable and a value of -1 represents a perfect negative correlation. A value of zero means there is no linear relationship between the variables.

To test empirical validity, we calculated Concordance Statistics (or C Statistics) between Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization and three outcomes: mental health inpatient readmissions, psychotropic medication adherence, and continued engagement in care at six months post discharge. A mental health inpatient readmission was defined as an acute admission with a primary mental health or intentional self-harm diagnosis in months four through nine after discharge. Medication adherence was defined as having 80% or more days covered with a medication class appropriate for the primary diagnosis in months four through six following discharge. Continued engagement in care at six months post discharge was defined as having at least one community based mental health care visit in month six post discharge.

The predictive accuracy of a statistical model can be measured by the agreement between observed and predicted outcomes. The C (Concordance) Statistic assesses the ability of a risk factor to predict an outcome. It is commonly used with logistic regression models with a binary outcome. The concept underlying concordance is that a subject who experiences a particular outcome has a higher predicted probability of that outcome than a subject who does not experience the outcome. The C Statistic can be calculated as the proportion of pairs of subjects whose observed and predicted outcomes agree (are concordant) among all possible pairs in which one subject experiences the outcome of interest and the other one does not. The higher the C-statistic, the better the model can discriminate between subjects who do experience the outcome of interest and subjects who do not.

C statistics generally range from 0.5 to 1. A C statistic of 1 represents perfect concordance and 0.5 means the model is no better at predicting the outcome than random chance. Models are typically considered reasonable when the C statistic is higher than 0.7 and strong when it exceeds 0.8.

To ensure face validity, this measure was conceptualized by a workgroup consisting of mental health clinicians and researchers. The concept of follow-up care after a psychiatric discharge being beneficial was supported by a literature review and the use of the HEDIS Follow-Up After Hospitalization for Mental Illness measure by NYS Medicaid to monitor quality of care. The workgroup and NYS OMH program staff agreed that measuring engagement beyond a single follow-up visit was clinically sound. They felt that more than a single visit was needed to adequately treat an individual after a discharge.

During the development process, we presented the measure concepts and testing analyses to the NYS Behavioral Health Clinical Advisory group and gathered their feedback. This group consists of representatives from NYS state health and behavioral health agencies, managed care plans, providers, and behavioral health advocacy organizations.

Question Response:
The names of the people on the NYS Behavioral Health Clinical Advisory group can be

found here - expand the Behavioral Health CAG item: https://www.health.ny.gov/health_care/medicaid/redesign/dsrip/vbp_library/cag/members.htm
The NYS Behavioral Health Clinical Advisory Group members include the Office of Mental Health Chief Medical Officer, medical directors and executive staff from NYS health and behavioral health agencies, executive staff of behavioral health advocacy organizations, executive and quality staff of health plans, and academic experts in behavioral health.

2b.03) Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

Table 15. Pearson Correlation Coefficient for New York State Medicaid Health Plans for Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization and Follow-Up After Hospitalization for Mental Illness, 2018 NYS Medicaid Discharges

	Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization
HEDIS Follow-Up After Hospitalization for Mental Illness	0.56044 N=50,234 discharges over 31 plans and FFS, p value <.0001

Table 16. Concordance Statistics for Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization, NYS Medicaid Calendar Year 2018 Discharges

Outcome	Population	C Statistic	95% Confidence Interval
Inpatient mental health readmissions in months 4-9 after discharge	All eligible discharges	0.5291	(0.5247-0.5335)
Adherence to antipsychotic medications in months 4-6 after discharge	Discharges with a primary diagnosis of schizophrenia	0.6428	(0.6358-0.6498)
Adherence to mood stabilizer medications in months 4-6 after discharge	Discharges with a primary diagnosis of bipolar disorder	0.6461	(0.6360-0.6562)
Adherence to antidepressant medications in months 4-6 after discharge	Discharges with a primary diagnosis of depression	0.6367	(0.6288-0.6446)

Outcome	Population	C Statistic	95% Confidence Interval
At least one community based mental health visit in month six post discharge	All eligible discharges	0.7244	(0.7205-0.7283)

Question Response:

95% confidence intervals for the C Statistics shown in Table 16 were added.

For face validity, a poll was conducted during the Fall 2022 meeting of NYS Behavioral Health Clinical Advisory Group where the following questions were asked about this measure:

1. Does the group think there is a benefit to going beyond the HEDIS FUH measure standard of one visit in thirty days and looking at engagement over a three-month period after discharge?

   [Responses 1. Yes, I think it is beneficial to measure engagement over three months post discharge 2. No, I think the HEDIS FUH measure is enough

2. For people with a recent psychiatric discharge, the data suggests that adequate engagement would be at least five outpatient services in the three months after discharge. Is this a reasonable threshold?

   [Responses 1. Yes, five visits is reasonable 2. No, five visits is not enough 3. No, five visits is too many

3. We have noticed a confounding effect with a group of discharges who have many visits, but whose outcomes don't seem to improve with more treatment. Do you have ideas on how we can control for this confounding?

   [Response: Text box]

In the Spring 2022 meeting of the NYS Behavioral Health Clinical Advisory Group, the following discussion questions were posed to the group (this was not a poll):

1. Is the numerator standard of a minimum of five to seven visits in the three months after discharge reasonable?

2. Are there other outcomes we could examine in reliability and validity analyses? For example, medication adherence, continuing engagement, etc.

2b.04) Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)

For construct validity, the correlation between our measure and the HEDIS FUH

measure is moderate. This result is expected because 62% of the 2018 discharges that met the criteria for the HEDIS measure met the criteria for the NYS OMH measure. The strength of the correlation is reasonable when considering the degree of overlap between the two measures.

In terms of empirical validity, our measure was reasonably predictive of engagement in community based mental health care at six months post discharge. The measure did not meet the 0.7 C Statistic threshold for predictability for acute service use or medication adherence. This lack of predictability may be due to confounding as we did not adjust for any risk factors in our models.

For face validity the measure specifications, descriptive analyses, reliability analyses, and validity analyses were presented to internal workgroups and the NYS Behavioral Health Clinical Advisory group over a several month period. The workgroups agreed that the measure had good face validity.

Question response:
For empirical validity, 95% confidence intervals were added to Table 16. Both the C Statistic and the lower 95% confidence interval for the outcome "At least one community-based mental health visit in month six post discharge" was above the 0.7 threshold for reasonable predictability.

For face validity, in a poll taken during the Fall 2022 meeting of the NYS Behavioral Health Clinical Advisory Group, 92% of respondents thought there was a clinical benefit to going beyond the HEDIS Follow-Up After Hospitalization for Mental Illness measure and measuring engagement over a three-month period. 69% thought a threshold of five visits was reasonable, 15% thought five visits were not enough, and 16% thought five visits were too many.

In the Spring 2022 meeting of the NYS Behavioral Health Clinical Advisory group, the group was asked about the numerator standard. Group members who responded to the discussion question thought the numerator standard of five visits was reasonable and no objections were raised by group members.

Scientific Acceptability: Validity - Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) (2b.05 - 2b.14)

2b.05) Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.

Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.

We calculated an interquartile range to determine if there were practically meaningfully differences in performance measure scores between managed care plans. The IQR is calculated as the rate for the 75th percentile minus the rate for the 25th percentile. To determine if the difference between scores were statistically significant, we conducted an independent sample t-test between two randomly selected plans below the 25th percentile and above the 75th percentile. If the p value of the t-test is less than 0.05, then the performance measure scores of the plans are significantly different.

2b.06) Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.

Table 17. Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization, Variation in Performance Across NYS Medicaid Health Plans, 2018 NYS Medicaid Discharges

	N	Min	P10	P25	Mean	Median	P75	P90	Max	IQR	P value
Engagement in Community-Based Mental Health Care After a Mental Health Hospitalization Rate	**32**	11%	25%	41%	46%	49%	53%	58%	66%	12%	p < 0.001

N = Number of plans reporting

IQR = Interquartile range

p-value = p-value of independent samples t-test comparing randomly selected plans at the 25th percentile to plans at the 75th percentile.

2b.07) Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.

In other words, what do the results mean in terms of statistical and meaningful differences?
There was a difference of 55% between the lowest and highest scoring plans. The interquartile range, or difference between the 75th and 25th percentiles, was 12%. The p value from the t test was $p<0.0001$, which means that the difference between the two randomly selected plan scores was statistically significant.

2b.08) Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

This measure is based on NYS administrative Medicaid data. Claims are used to identify discharges and community-based follow-up visits. It is not expected that a large proportion of claims would be missing because they must be submitted to NYS in order for managed care plans and services providers to receive payment. There is a data lag of about six months in NYS claims. Our analysis used claims data from calendar year 2018 that was extracted after the six-month claims lag period had passed. Performance measures in NYS are always calculated at least six months after the end of the measurement period to minimize the effect of claims lag. When performance measures are calculated, the numerator and denominator counts for the current year are compared to previous years to check for any aberrations that could indicate missing data.

Question response: A six month claims lag would not be part of the measure specifications for all health plans. At NYS state agencies, including NYS OMH, we wait six months before we consider data from health plans complete, but the health plans themselves are expected to have access to complete data in less time. We also expect the claims lag period to vary by health plan and state. We feel that claims lag should be accounted for in the reporting process and in reporting deadlines, but it does not need to be included in the measure specifications.
The general guidelines for HEDIS measures require plans to report their HEDIS measures to NCQA by mid-June of the year following the measurement year. This timeframe is consistent with a claims lag allowance, but the individual HEDIS measure specifications do not mention an allowance for claims lag.

2b.09) Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

It is not possible to know how many claims were not submitted in Calendar Year 2018 because we have no secondary source of the number of services provided. The NYS Medicaid Data Warehouse and NYS Office of Mental Health monitor the number of claims submitted monthly. If there are aberrations in the number of claims compared to prior months, the differences are investigated and changes are made if it is determined that data is missing or was duplicated.

2b.10) Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.

In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.

Note: This item is directed to measures that are risk-adjusted (with or without social risk factors) OR to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eCQMs). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

If aberrations in the monthly number of claims submitted are found, NYS will identify and correct the source of the error. When performance measures are calculated, if quality checks find differences of more than 5-10% in plan level denominator sizes, NYS OMH will identify the source of the difference and correct any data errors found before results are reported.

2b.11) Indicate whether there is more than one set of specifications for this measure.

☐ Yes, there is more than one set of specifications for this measure
☒ No, there is only one set of specifications for this measure

2b.12) Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.

Describe the steps—do not just name a method. Indicate what statistical analysis was used.
N/A, only one set of specifications and data source were used.

2b.13) Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.
N/A, only one set of specifications and data source were used.

2b.14) Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.
N/A, only one set of specifications and data source were used.

Scientific Acceptability: Validity - Other Threats to Validity (Exclusions, Risk Adjustment) (2b.15 - 2b.32)

2b.15) Indicate whether the measure uses exclusions.

☐ N/A or no exclusions
☒ Yes, the measure uses exclusions.

2b.16) Describe the method of testing exclusions and what was tested.

Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?

This measure excludes:

1. Discharges who were dually enrolled in Medicaid and Medicare.
2. Discharges in hospice or using hospice services anytime during the measurement year.

Discharges who were dually enrolled in Medicaid and Medicare were excluded because this is a measure for Medicaid managed care plans and the entirety of services that Medicare/Medicaid duals receive cannot be reliably derived from Medicaid claims data. We would miss discharges or follow-up visits that are reported in Medicare if dual enrollees were included in the measure.

Discharges who received hospice services during the measurement year were excluded because the focus of care and criteria for quality care for this population may be different than the population not receiving hospice services. Engagement in community-based mental health care may not be a priority for individuals receiving hospice.

About 3% of discharges were dually enrolled in Medicaid and Medicare. Individuals with dual Medicaid and Medicare enrollment may have more severe conditions or more co-morbidities than individuals enrolled in Medicaid only. Because we cannot derive all of their service data from Medicaid claims and may miss discharges or follow-up visits, we cannot determine if the dually enrolled population is more or less likely to engage in community-based mental health follow-up care. Excluding this population may slightly increase or decrease the rate of discharges receiving five or more community-based mental health follow-up visits.

It is expected that the hospice exclusion will not have a meaningful impact on the measure because very few discharges met the criteria for this exclusion. Less than 0.1% of discharges were in hospice or used hospice services during the measurement year. Individuals using hospice services may be less likely to receive five or more community-based mental health follow-up visits. Excluding this population is unlikely to change the measure rate detectably because the number of exclusions is so small.

2b.17) Provide the statistical results from testing exclusions.

Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.

These exclusions are common for Medicaid performance measures, including the NCQA HEDIS measure Follow-Up After Hospitalization for Mental Illness, and were not tested.

2b.18) Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.

Excluding Medicare and Medicaid duals is necessary because we cannot use Medicaid claims data to derive all of the services received during the measurement year. It enhances the fairness of the measure to exclude discharges that may have services missing in Medicaid data.

Excluding discharges who received hospice services during the measurement year is necessary because the focus of care and criteria for quality care for this population may be different than the population not receiving hospice services. It enhances the fairness of the measure to exclude discharges where engagement in community based mental health care may not be a priority.

2b.19) Check all methods used to address risk factors.

☐ Statistical risk model with risk factors (specify number of risk factors)
☒ Stratification by risk category (specify number of categories)
☐ Other (please specify here:)
☐ No risk adjustment or stratification

Two age stratifications and a total rate will be reported.

- Stratification 1: Ages 6-20
- Stratification 2: Ages 21-64
- Total rate

2b.20) If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.
N/A, a statistical risk model was not used.

2b.21) If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.
N/A
2b.22) Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

☐ Published literature
☐ Internal data analysis
☐ Other (please specify here:)
N/A
2b.23) Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10 or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

The results of this measure are stratified by age because there are notable differences between the mental health treatment system for children ages 6-20 and for adults ages 21-64. The services offered, service settings, service definitions, clinicians, and training of clinicians are different for children compared to adults.

The HEDIS measure Follow-Up After Hospitalization for Mental Illness also stratifies results into child and adult age groups.

2b.24) Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.
N/A
2b.25) Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.
N/A
2b.26) Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide

the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.
N/A
2b.27) Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.
N/A
2b.28) Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).
N/A
2b.29) Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.
N/A
2b.30) Provide the results of the risk stratification analysis.
N/A
2b.31) Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?
N/A
2b.32) Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.
N/A

Feasibility (3.01 - 3.07)

3.01) Check all methods below that are used to generate the data elements needed to compute the measure score.

☐ Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)
☒ Coded by someone other than person obtaining original information (e.g., DRG, ICD-10 codes on claims)
☐ Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)
☐ Other (Please describe)

3.02) Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields. ALL data elements are in defined fields in electronic health records (EHRs)

☒ ALL data elements are in defined fields in electronic claims
☐ ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)
☐ ALL data elements are in defined fields in a combination of electronic sources
☐ Some data elements are in defined fields in electronic sources
☐ No data elements are in defined fields in electronic sources
☐ Patient/family reported information (may be electronic or paper)

3.03) If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using data elements not from electronic sources.
Not applicable.

3.04) Describe any efforts to develop an eCQM.
There have not been any efforts to develop an eCQM with this measure concept by NYS OMH.

3.05) Complete and attach the eCQM-Feasibility-Scorecard.xls file.
Not applicable.

3.06) Describe difficulties (as a result of testing and/or operational use of the measure) regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

This measure is based on administrative Medicaid data and we did not note any of the difficulties listed. Providers must submit claims to be reimbursed by Medicaid and this measure does not create any additional reporting requirements for providers. Health plans will not be required to collect any additional data.
There is an overall data lag of about six months in NYS Medicaid claims. Our analysis used claims data from calendar years 2018-21 that was extracted after the six-month claims lag period had passed. Performance measures in NYS are always calculated at least six months after the end of the measurement period to minimize the effect of claims lag. When performance measures are calculated, the numerator and denominator counts for the current year are compared to previous years to check for any aberrations that could indicate missing data.
A six month claims lag would not be part of the measure specifications for all health plans. At NYS state agencies, including NYS OMH, we wait six months before we consider data from health plans complete, but the health plans themselves are expected to have access to complete data in less time. We also expect the claims lag period to vary by health plan and state. We feel that claims lag should be accounted for in the reporting process and in reporting deadlines, but it does not need to be included in the measure specifications.
The general guidelines for HEDIS measures require plans to report their HEDIS measures to NCQA by mid-June of the year following the measurement year. This timeframe is consistent with a claims lag allowance, but the individual HEDIS measure specifications do not mention an allowance for claims lag.

3.07) Detail any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm),

Attach the fee schedule here, if applicable.

HEDIS value sets are used to calculate this measure. To use HEDIS value sets, fees must be paid to NCQA. No fees will be collected by NYS OMH.

HEDIS value sets are available through the purchase of the publication HEDIS Technical Specifications for Health Plans, Volume 2. The fees for the electronic version of the specifications vary by the number of users. The fee schedule is below:

Number Of Users
- Single User: $480.00
- 2-4 Users: $1,340.00

- 5-10 Users: $3,120.00
- 11-20 Users: $5,760.00
- 21-30 Users: $8,640.00
- 31-50 Users: $14,400.00

Use (4a.01 – 4a.10)

Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making.

Endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement, in addition to demonstrating performance improvement.

4a.01) Check all current uses. For each current use checked, please provide:

- **Name of program and sponsor**
- **URL**
- **Purpose**
- **Geographic area and number and percentage of accountable entities and patients included**
- **Level of measurement and setting**

☐ Public Reporting
☐ Public Health/Disease Surveillance
☐ Payment Program
☐ Regulatory and Accreditation Programs
☐ Professional Certification or Recognition Program
☐ Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
☐ Quality Improvement (Internal to the specific organization)
☒ Not in use
☐ Use unknown
☐ Other (please specify here:)

4a.02) Check all planned uses.

☒ Public reporting
☐ Public Health/Disease Surveillance
☒ Payment Program
☐ Regulatory and Accreditation Program
☐ Professional Certification or Recognition Program
☐ Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
☒ Quality Improvement (internal to the specific organization)
☐ Measure Currently in Use
☐ Other (please specify here:)

4a.03) If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

This measure is new and has been undergoing testing. It has not yet been publicly reported or integrated into New York State's accountability applications. New Medicaid measures in New York State typically undergo testing for at least 1-2 years before they are publicly reported and/or included in accountability applications.

4a.04) If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.

When the measure is implemented, it will be included in one or more New York State accountability and quality improvement platforms that share performance data by region, county, plan, provider agencies, and/or provider networks. This measure will be embedded within a robust statewide mental health quality improvement infrastructure. These platforms are described below:

NYS Quality Assurance Reporting Requirements
NYS OMH plans to add the measure into New York State's Quality Assurance Reporting Requirements (QARR) for managed care plans. QARR measures are largely adopted from HEDIS with New York State-specific measures added to address public health issues of particular importance in New York. QARR data is collected by health plans and the information is validated by a licensed organization. Only valid information is included in the data.

QARR's purpose is to enable consumers to evaluate the quality of health care services provided by New York State's managed care plans. Using QARR, a consumer can determine how well a health plan performed in the areas of provider network, child and adolescent health, women's health, adult health, behavioral health, and experience with care.

QARR measures are reported publicly and the intended audience is consumers, managed care plans, government agencies, and health care providers. We expect to release the measure specifications for public comment in NYS in fall 2023. Preliminary measure year 2023 rates will be shared with Medicaid managed care plans for their

information and feedback. The measure will be added to QARR for measurement year 2024. NYS Medicaid managed care plans will report on the measure when their 2024 QARR reports are due in mid-2025. NYS QARR data is also used in a quality incentive for Medicaid managed care plans. NYS OMH expects to include this measure in the measurement year 2025 quality incentive.

NYS OMH Vital Signs Dashboard
The NYS OMH Vital signs dashboard allows plans, counties, and providers to examine performance by county and provider and program type. This dashboard highlights disparities in performance by race, ethnicity, gender, age, and region. The measure will be added beginning in measurement year 2024.

Psychiatric Services and Clinical Knowledge Enhancement System (PSYCKES)
The Psychiatric Services and Clinical Knowledge Enhancement System (PSYCKES) assesses performance on quality measures at the state, region, county, provider, and network levels. It allows drill downs to patient caseloads and plan members with quality gaps. New measures are added based on user priorities (provider, plan, county, OMH leadership requests). Measures included in PSYCKES are refreshed monthly, and attributed to managed care plans and to providers to allow them to rapidly identify individuals with quality gaps. The measure will be added beginning in measurement year 2024.

Statewide Quality Improvement Collaboratives (QIC)
NYS OMH Statewide Quality Improvement Collaboratives (QIC) engage providers in quality improvement initiatives and address several QI performance goals. Three to five large scale QICs are run every year. Providers and plans are offered best practice strategies for improving performance on measures. Performance is monitored monthly with feedback to participants. We would like a year or two of baseline data before including the measure in a QIC. Measurement year 2025 or 2026 would be the first year for the measure to be included in a QIC.

4a.05) Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

Detail how many and which types of measured entities and/or others were included. If only a sample of measured entities were included, describe the full population and how the sample was selected.

Blinded plan level measure results for 28 Medicaid Managed Care plans (all Mainstream and Health and Recovery Plans; results for the three HIV Special Needs Plans and fee for service were not included) were presented during the Fall 2022 Behavioral Health Clinical Advisory Group meeting. Six managed care plans have representatives on this group.

Four NYS Medicaid Managed Care plans were offered the opportunity to participate in a field test of this measure. For the field test, we ask participating plans to calculate the measure using our specifications, share their results with us, and answer questions about their experience calculating the measure and their opinion of the measure. A field test is currently being conducted with one New York State Medicaid Managed Care plan. Statewide product line level measure data that detailed denominator counts, numerator counts, and exclusions by age group and NYS region was provided to the plan participating in the field test.

Once the measure is implemented, NYS has an infrastructure in place for providing performance results and assistance with interpretation. NYS QARR results are available on a public website. There is an annual educational webinar, documentation available on the web, and a help desk that can be contacted for assistance. The NYS OMH Vital Signs Dashboard is available on the public domain. The development team is reviewing measures, refining Tableau features to support usability, and field testing with 23 hospital providers. There are statewide provider webinars with multiple offerings annually that introduce providers to the dashboard, explain goals and measures, and demonstrate how to use the data. There are also office hours and a help desk to address questions on interpretation of measure performance and how to use data to drive change. PSYCKES results are available on a restricted website. All provider users (12,000+) receive release notes summarizing new measures added. There are multiple training options offered every quarter to all users on how to use PSYCKES to support clinical decision making and quality improvement. The PSYCKES help desk manages a large volume of questions and technical assistance requests. Tickets are logged with a 24-hour response time. Statewide QICs are supported by a web platform and training and support is offered to providers and plans in monthly webinars, quarterly individual consultation calls, and ad hoc requests for technical assistance.

4a.06) Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

Blinded plan level measure results for 28 Medicaid Managed Care plans (all Mainstream and Health and Recovery Plans; results for the three HIV Special Needs Plans and fee for service were not included) were presented during the Fall 2022 Behavioral Health Clinical Advisory Group meeting. Six managed care plans have representatives on this group.

Measure rates for each plan (with the plan name hidden) were shown to the group. Before presenting the rates, presenters from OMH explained the measure specifications and showed the results of various validation analyses to the group. Group members were given the opportunity to ask questions and comment on the measure specifications and results.

Once the measure is implemented, NYS has an infrastructure in place for providing

results and educational/explanatory assistance to those being measured. NYS QARR holds an annual educational webinar, provides documentation, and has a help desk that can be contacted for assistance. The NYS OMH Vital Signs Dashboard holds statewide provider webinars with multiple offerings annually that introduce providers to the dashboard, explain goals and measures, and demonstrate how to use the data. There are also office hours and a help desk to address questions on interpretation of measure performance and how to use data to drive change. For PSYCKES all provider users (12,000+) receive release notes summarizing new measures added. There are multiple training options offered every quarter to all users on how to use PSYCKES to support clinical decision making and quality improvement. The PSYCKES help desk manages a large volume of questions and technical assistance requests. Tickets are logged with a 24-hour response time. For statewide QICs, training and support are offered to providers and plans in monthly webinars, quarterly individual consultation calls, and ad hoc requests for technical assistance.

4a.07) Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

A poll was conducted during the Fall 2022 meeting of NYS Behavioral Health Clinical Advisory Group where the following questions were asked about this measure:

1. Does the group think there is a benefit to going beyond the HEDIS FUH measure standard of one visit in thirty days and looking at engagement over a three-month period after discharge?
[Responses 1. Yes, I think it is beneficial to measure engagement over three months post discharge 2. No, I think the HEDIS FUH measure is enough
2. For people with a recent psychiatric discharge, the data suggests that adequate engagement would be at least five outpatient services in the three months after discharge. Is this a reasonable threshold?
[Responses 1. Yes, five visits is reasonable 2. No, five visits is not enough 3. No, five visits is too many

3. We have noticed a confounding effect with a group of discharges who have many visits, but whose outcomes don't seem to improve with more treatment. Do you have ideas on how we can control for this confounding?
[Response: Text box]

In the Spring 2022 meeting of the NYS Behavioral Health Clinical Advisory Group, the following discussion questions were posed to the group (this was not a poll):

1. Is the numerator standard of a minimum of five to seven visits in the three months after discharge reasonable?
2. Are there other outcomes we could examine in reliability and validity analyses? For example, medication adherence, continuing engagement, etc.

Feedback received: 92% of poll respondents thought there was a clinical benefit to going beyond the HEDIS Follow-Up After Hospitalization for Mental Illness measure and

measuring engagement over a three-month period. 69% thought a threshold of five visits was reasonable, 15% thought five visits were not enough, and 16% thought five visits were too many.

In the Spring 2022 meeting of the NYS Behavioral Health Clinical Advisory group, the group was asked whether the numerator standard of five visits in the three months after discharge was reasonable. Group members who responded to the discussion question thought the numerator standard was reasonable and no objections were raised by group members.

The field test is ongoing and we do not have final feedback from plans yet.

4a.08) Summarize the feedback obtained from those being measured.

BH CAG poll responses are anonymous and there are meeting participants from entities other than managed care plans. We do not know the responses of plans specifically. The NYS Behavioral Health Clinical Advisory Group members include the Office of Mental Health Chief Medical Officer, medical directors and executive staff from NYS health and behavioral health agencies, executive staff of behavioral health advocacy organizations, executive and quality staff of six health plans, and academic experts in behavioral health. The names of the people on the NYS Behavioral Health Clinical Advisory group can be found here - expand the Behavioral Health CAG item: https://www.health.ny.gov/health_care/medicaid/redesign/dsrip/vbp_library/cag/members.htm

In the poll taken during the Fall 2022 meeting of the NYS Behavioral Health Clinical Advisory Group, 92% of respondents thought there was a clinical benefit to going beyond the HEDIS Follow-Up After Hospitalization for Mental Illness measure and measuring engagement over a three-month period. 69% thought a threshold of five visits was reasonable, 15% thought five visits were not enough, and 16% thought five visits were too many.

In the Spring 2022 meeting of the NYS Behavioral Health Clinical Advisory group, the group was asked about the numerator standard. Group members who responded to the discussion question thought the numerator standard of five visits was reasonable and no objections were raised by group members.

The field test is ongoing and we do not have final feedback from plans yet.

4a.09) Summarize the feedback obtained from other users.
BH CAG poll responses are anonymous and there are participants from entities other than managed care plans. We do not know the responses of other users specifically. The NYS Behavioral Health Clinical Advisory Group members include the Office of Mental Health Chief Medical Officer, medical directors and executive staff from NYS health and behavioral health agencies, executive staff of behavioral health advocacy organizations, executive and quality staff of six health plans, and academic experts in

behavioral health. The names of the people on the NYS Behavioral Health Clinical Advisory group can be found here - expand the Behavioral Health CAG item: https://www.health.ny.gov/health_care/medicaid/redesign/dsrip/vbp_library/cag/members.htm

In the poll taken during the Fall 2022 meeting of the NYS Behavioral Health Clinical Advisory Group, 92% of respondents thought there was a clinical benefit to going beyond the HEDIS Follow-Up After Hospitalization for Mental Illness measure and measuring engagement over a three-month period. 69% thought a threshold of five visits was reasonable, 15% thought five visits were not enough, and 16% thought five visits were too many.

In the Spring 2022 meeting of the NYS Behavioral Health Clinical Advisory group, the group was asked about the numerator standard. Group members who responded to the discussion question thought the numerator standard of five visits was reasonable and no objections were raised by group members.

4a.10) Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

The NYS Behavioral Health Performance Measurement Center reviewed the feedback received during the BH CAG meetings. The measure specifications were not modified because the feedback we received was generally in agreement with the measure concept. We consider the feedback of the BH CAG to be supportive of decisions we made during the measure development process, including the length of the follow-up period after discharge, the service types included in the numerator, the number of visits required by the numerator, and not including a requirement about the regularity of visits. These decisions along with supporting data were shown to the BH CAG in meetings in 2021 and 2022 and they had an opportunity to raise any objections during each meeting.

Once the field test with managed care plans is complete, the NYS Behavioral Health Performance Measurement Center will review the feedback received during the field test. The measure specifications and/or implementation plans will be revised if the field test highlights any problems with either the specifications or implementation of the measure.

Usability (4b.01 - 4b.03)

4b.01) You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

This measure is not currently in use for performance improvement. Performance results can be used by managed care plans to monitor the percentage of psychiatric discharges who are receiving adequate follow-up mental health care. All plans in NYS have room for improvement in this measure. Plans can collaborate with discharging hospitals to ensure that initial linkages to community-based mental health care occur after discharge and can partner with community-based mental health care providers to improve engagement in care in the three months after discharge. Plans can use a data driven approach and determine which community-based providers have higher rates of engagement. These providers can be a source of best practices that can be communicated to and used by providers with lower rates of engagement in order to improve their performance.

The data provided in Importance to Measure and Report: Gap in Care/Disparities identified disparities in age, sex, race, ethnicity, and primary diagnosis at discharge. NYS agencies, plans, and providers will continue to monitor disparities and use a data driven approach to identify best practices to reduce disparities.

4b.02) Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

This measure has not been implemented.

4b.03) Explain any unexpected benefits realized from implementation of this measure.

This measure has not been implemented.

Related and Competing (5.01 - 5.06)

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

5.01) Search and select all endorsed related measures (conceptually, either same measure focus or target population) by going to the PQM website.

(Can search and select measures.)

Follow-Up After Hospitalization for Mental Illness

5.02) Search and select all endorsed competing measures (conceptually, the measures have both the same measure focus or target population) by going to the PQM website.

(Can search and select measures.)

None.

5.03) If there are related or competing measures to this measure, but they are not endorsed, please indicate the measure title and steward.
There are no known related or competing measures to this measure that are not NQF-endorsed.

5.04) If this measure conceptually addresses EITHER the same measure focus OR the same target population as endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

☒ Yes
☐ No

5.05) If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.
This measure and Follow-Up After Hospitalization for Mental Illness use the same value sets, with the exception of the discharge diagnosis value set. This measure has a broader discharge diagnosis value set than the HEDIS measure and includes diagnoses such as anxiety and eating disorders that the HEDIS measure does not include. These additional diagnoses represent about 1% of discharges. NYS OMH felt that there was no clinical rationale for excluding discharges with these diagnoses. With the proportion of diagnoses added being small, it is not expected to impact the interpretability of the measure in a meaningful way. Since all of the data for this measure are derived from claims, these additional diagnoses are not expected to present additional reporting burden.

5.06) Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

No competing measures were identified.

The HEDIS measure Follow-Up After Hospitalization for Mental Illness captures one outpatient follow-up visit within seven and thirty days after psychiatric discharge. We think that it is important to also measure engagement in community-based mental health care over a longer time period post-discharge. People being discharged from a psychiatric hospitalization are in an acute phase of their mental health condition and need to receive more than one visit for adequate treatment.

This measure will improve the quality of care by encouraging both linkage to community-based mental health care services after an inpatient psychiatric discharge and engagement with community-based mental health services. Follow-up care after an inpatient psychiatric discharge has been associated with reduced inpatient readmissions, increased medication utilization, increased outpatient encounters, and increased functioning.

Additional (1 - 9)

1) Provide any supplemental materials, if needed, as an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be collated one file with a table of contents or bookmarks. If material pertains to a specific criterion, that should be indicated.

☐ Available in attached file
☒ No appendix
☐ Available at measure-specific web page URL identified in sp.09

2) List the workgroup/panel members' names and organizations.

Describe the members' role in measure development.
The workgroup who developed this measure included:

Molly Finnerty, MD, New York State Office of Mental Health: provided clinical expertise, oversight of specification development and analytic work.

Xiaoli Fu, MS, New York State Office of Mental Health: conducted analyses and developed data tables

Emily Leckman-Westin, PhD, New York State Office of Mental Health: provided methodological expertise, oversight of analytic work and specification development.

Harold Pincus, MD, Columbia University: provided clinical and methodological expertise

Adrienne Ronsani, MS, New York State Office of Mental Health: conducted analyses, wrote measure specifications, oversight of analytic work.

Wenjun Shao, MS, New York State Office of Mental Health: conducted analyses, developed data tables, wrote measure specifications, oversight of analytic work.

Thomas Smith, MD, New York State Office of Mental Health and Columbia University: provided clinical expertise, oversight of specification development and analytic work.

The names and organizations of the people on the NYS Behavioral Health Clinical Advisory group can be found here - expand the Behavioral Health CAG item: https://www.health.ny.gov/health_care/medicaid/redesign/dsrip/vbp_library/cag/members.htm

3) Indicate the year the measure was first released.
This measure is not currently in use. The specifications as submitted were finalized in 2022.

4) Indicate the month and year of the most recent revision.
The measure specifications as submitted were finalized in November 2022.

5) Indicate the frequency of review, or an update schedule, for this measure.
The measure specifications will be reviewed an updated annually.

6) Indicate the next scheduled update or review of this measure.
This measure will be reviewed and updated if necessary in Fall 2023.

7) Provide a copyright statement, if applicable. Otherwise, indicate "N/A".
N/A

8) State any disclaimers, if applicable. Otherwise, indicate "N/A".
N/A

9) Provide any additional information or comments, if applicable. Otherwise, indicate "N/A".
N/A