NATIONAL QUALITY FORUM

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Purple text represents the responses from measure developers.
Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 3205

Corresponding Measures:

De.2. Measure Title: Medication Continuation Following Inpatient Psychiatric Discharge

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services

De.3. Brief Description of Measure: This measure assesses whether patients discharged from an inpatient psychiatric facility (IPF) with major depressive disorder (MDD), schizophrenia, or bipolar disorder filled a prescription for evidence-based medication within 2 days prior to discharge and 30 days post-discharge. This measure evaluates admissions over a two-year period.

1b.1. Developer Rationale: The aim of the measure is to address gaps in continuity of pharmaceutical treatment during the transition from inpatient to outpatient care. Pharmacotherapy is the primary form of treatment for most patients discharged from an inpatient psychiatric facility (IPF) for bipolar disorder, major depressive disorder (MDD), or schizophrenia. The measure focuses on medication continuation because it is an essential step in medication adherence.

Medication continuation is particularly important in the psychiatric patient population because psychotropic medication discontinuation can have a range of adverse effects, from mild withdrawal to life-threatening autonomic instability and psychiatric decompensation (Ward & Schwartz, 2013). Patients with MDD who do not remain on prescribed medication are more likely to have negative health outcomes, such as relapse and readmission, decreased quality of life, and increased health care costs. If untreated, MDD can contribute to or worsen chronic medical disorders (Geddes et al., 2003; Glue et al., 2010). The literature shows that among patients with schizophrenia, those who were "good compliers" according to the Medication Adherence Rating Scale had better outcomes in terms of rehospitalization rates and medication maintenance (Jaeger et al., 2012). Among patients with bipolar disorder, medication adherence was significantly associated with reduction in manic symptoms (Sylvia et al., 2013), whereas nonadherence was associated with increased suicide risk (OR 10.8, CI 1.57–74.4; Gonzalez-Pinto et al., 2006). Our literature review from January 2016 through August 2020 did not reveal any new evidence regarding performance gaps since the initial endorsement submission.

Current facility-level performance indicates a clear quality gap. Using Medicare claims data from July 1, 2017, through June 20, 2019, the Medication Continuation measure rates ranged from 34.8 to 94.3%, with a median of 76.2%. There was a 21.3 percentage point difference between the 10th and 90th percentiles (63.4–84.7%). Using 2013–2014 Medicare claims data, there was a 21.6 percentage point difference between the 10th and 90th percentiles (66.7–88.3%) and a median score of 79.6%. By calculating the facility-level rates of medication

continuation in Medicare fee-for-service (FFS) claims data, this measure can provide valuable information on areas where care transitions to the outpatient setting can be improved.

Literature about continuation of medication has identified effective interventions that facilities can employ to improve medication adherence among patients discharged from an IPF (Douaihy, Kelly, & Sullivan, 2013; Haddad, Brain, & Scott, 2014; Hung, 2014; Kasckow & Zisook, 2008; Lanouette, Folsom, Sciolla, & Jeste, 2009; Mitchell, 2007; Sylvia et al., 2013). Examples of these interventions include patient education, shared decision making, and text-message reminders. We envision the addition of this measure to the suite of measures for IPFs would help to create a comprehensive picture of the quality of care patients receive at those facilities.

*Douaihy, A. B., Kelly, T. M., Sullivan, C. (2013). Medications for substance use disorders. Social Work in Public Health, 28(3-4), 264-278. doi: 10.1080/19371918.2013.759031

*Geddes, J. R., Carney, S. M., Davies, C., Furukawa, T. A., Kupfer, D. J., Frank, E., & Goodwin, G. M. (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. The Lancet, 361(9358), 653–661. doi:10.1016/s0140-6736(03)12599-8

*Glue, P., Donovan, M. R., Kolluri, S., & Emir, B. (2010). Meta-analysis of relapse prevention antidepressant trials in depressive disorders. Australian and New Zealand Journal of Psychiatry, 44(8), 697-705. doi: 10.3109/00048671003705441

*Gonzalez-Pinto, A., Mosquera, F., Alonso, M., López, P., Ramírez, F., Vieta, E., & Baldessarini, R. J. (2006). Suicidal risk in bipolar I disorder patients and adherence to long-term lithium treatment. Bipolar Disorders, 8(5p2), 618–624. doi:10.1111/j.1399-5618.2006.00368.x

*Haddad, P. M., Brain, C., & Scott, J. (2014). Nonadherence with antipsychotic medication in schizophrenia: Challenges and management strategies. Patient Related Outcome Measures, 5, 43-62. doi: 10.2147/PROM.S42735

*Hung, C. I. (2014). Factors predicting adherence to antidepressant treatment. Current Opinion in Psychiatry, 27(5), 344-349. doi: 10.1097/yco.0000000000000086

*Jaeger, S., Pfiffner, C., Weiser, P., Kilian, R., Becker, T., Langle, G., . . . Steinert, T. (2012). Adherence styles of schizophrenia patients identified by a latent class analysis of the Medication Adherence Rating Scale (MARS): A six-month follow-up study. Psychiatry Research, 200(2-3), 83-88. doi: 10.1016/j.psychres.2012.03.033

*Kasckow, J. W., & Zisook, S. (2008). Co-occurring depressive symptoms in the older patient with schizophrenia. Drugs & Aging, 25(8),631-647.

*Lanouette, N. M., Folsom, D. P., Sciolla, A., Jeste, D. V. (2009). Psychotropic medication nonadherence among United States Latinos: A comprehensive literature review. Psychiatric Services (Washington, DC), 60(2), 157-174. doi: 10.1176/appi.ps.60.2.15724(4).

*Sylvia, L. G., Hay, A., Ostacher, M. J., Miklowitz, D. J., Nierenberg, A. A., Thase, M. E., Perlis, R. H. (2013). Association between therapeutic alliance, care satisfaction, and pharmacological adherence in bipolar disorder. Journal of Clinical Psychopharmacology, 33(3), 343-350. doi: 10.1097/JCP.0b013e3182900c6f

*Ward, M., & Schwartz, A. (2013). Challenges in pharmacologic management of the hospitalized patient with psychiatric comorbidity. Journal of Hospital Medicine, 8(9), 523–529. doi:10.1002/jhm.2059.

S.4. Numerator Statement: The numerator for the measure includes:

- Discharges with a principal diagnosis of MDD in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge
- Discharges with a principal diagnosis of schizophrenia in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge

- Discharges with a principal diagnosis of bipolar disorder in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge

S.6. Denominator Statement: The target population for this measure is Medicare fee-for-service (FFS) beneficiaries with Part D coverage aged 18 years and older discharged from an inpatient psychiatric facility with a principal diagnosis of MDD, schizophrenia, or bipolar disorder.

S.8. Denominator Exclusions: The denominator for this measure excludes discharged patients who:

- Received electroconvulsive (ECT) during the inpatient stay or follow-up period
- Received transcranial stimulation (TMS) during the inpatient stay or follow-up period
- Were pregnant at discharge
- Had a secondary diagnosis of delirium at discharge
- Had a principal diagnosis of schizophrenia with a secondary diagnosis of dementia at discharge

De.1. Measure Type: Process

S.17. Data Source: Claims

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Jun 28, 2017 **Most Recent Endorsement Date:** Jun 28, 2017

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable because this measure is not paired or grouped with another measure.

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence Since the prior evaluation.

1a. Evidence. The evidence requirements for a *structure, process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2016

- This facility-level, claims-based, process measure assesses whether patients discharged from an inpatient psychiatric facility (IPF) with major depressive disorder (MDD), schizophrenia, or bipolar disorder filled a prescription for evidence-based medication within 2 days prior to discharge and 30 days post-discharge. This measure evaluates admissions over a two-year period.
- Developer used the same logic model provided in the 2016 submission for this submission.
- In the 2016 submission, developer provided evidence for medication continuation based on treatment guidelines for major depressive disorder (APA 2010, VA/DoD 2016), schizophrenia (APA 2010), and bipolar disorder (APA 2002, VA/DoD 2010).
- Major depressive disorder (MDD)
    - APA 2010 Guidelines support the use of antidepressant medications for acute and maintenance treatment (except with ECT). **(Grade I Recommendation: substantial clinical confidence)**
    - VA/DoD 2016 Guidelines support the use of antidepressant medications for at least 6 months after remission **(Grade A Recommendation: good evidence, benefits substantially outweigh harm).**
- Schizophrenia
    - APA 2010 Guidelines support use of medications in acute phase **(Grade I Recommendation: substantial clinical confidence)**, for long-acting injectable medications for those with recurrent relapses **(Grade II Recommendation: moderate clinical confidence)**, and for continued medication for at least 6 months if improvement is noted **(Grade I Recommendation: substantial clinical confidence)**
    - The developer noted a new study since the guideline's release comparing longer-term effects and usefulness of a range of antipsychotics, supporting the inclusion of both typical and atypical antipsychotics in this measure.
- Bipolar Disorder
    - APA 2002 Guidelines support use of medications, describing a variety of options/medication choices depending on the situation **(Grade I and Grade II Recommendations: substantial or moderate clinical confidence)**, especially for continuation of medication after remission **(Grade I Recommendation: substantial clinical confidence)**.
    - VA/DoD 2010 Guidelines support the use of various medications **(Grade A, B, and I Recommendations: A-good evidence and benefits substantially outweigh harms; B-fair evidence and benefits outweigh harms; I-evidence on effectiveness is lacking or poor quality or conflicting and balance of benefits and harms cannot be determined)** and the use of medications for continued maintenance after an initial acute manic episode, for at least 6 months **(Grade A Recommendation)**.
        - Note that the Grade I recommendations mostly apply to medications to be used as a secondary choice.
- It was noted that the developer provided guidelines which emphasize the need for continued use of medications, but the evidence described largely focuses on the efficacy or relative advantage of individual medications and not on the timeliness of their use (as is the focus of this measure).
- It was further noted that the VA/DoD guidelines provide some insight to the quality of the studies, but overall the quality of the evidence has not been presented.
- The Committee agreed there is evidence that lack of adherence to medication leads to relapse and negative outcomes. They also noted that claims data related to medication adherence are directly correlated to outcomes

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided updated evidence for medication continuation in patients with schizophrenia based on 2019 APA Schizophrenia Practice Guidelines. The guidelines state that patients with schizophrenia who show improvement following treatment with an antipsychotic medication should continue to be treated with an antipsychotic, leading to improved outcomes and quality of life.
    - "Patients with schizophrenia be treated with an antipsychotic medication and monitored for effectiveness and side effects. APA recommends (**Grade: 1A – high**) that patients with schizophrenia whose symptoms have improved with an antipsychotic medication continue to be treated with an antipsychotic medication. *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."
- Developer did not provide updated evidence for MDD and bipolar disorder; the evidence provided was the same as the 2016 submission.

Questions for the Committee:

- What is the relationship of this measure to patient outcomes?
- Has there been improved evidence that the timelines for the measures are appropriate?

Guidance from the Evidence Algorithm

- (Box 3) Evidence based on a systemic review and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured→ (Box 4) QQC presented→ (Box) 5b MODERATE

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

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

Maintenance measures – increased emphasis on gap and variation

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

- The developer provided performance data from Medicare FFS Part A and Part B claims from July 1, 2017, through June 30, 2019.
- For the Inpatient Psychiatric Facility Quality Reporting (IPFQR) Program sponsored by the Centers for Medicare & Medicaid Services (CMS), the measure is calculated only using testing data from IPFs with at least 75 discharges (1,066 IPFs and 268,673 discharges meet this criteria).
- Medication continuation rate across all IPFs (n=1,680) in the data set:
    - Mean: 75.0%
    - Std dev: 12.8%
    - Min: 0.0%
    - Max: 100.0%
    - Interquartile range: 12.6%
- Medication continuation rate IPFs with at least 75 eligible cases in the denominator (n = 1,066):
    - Mean: 75.1%

- Std dev: 8.3%
- Min: 34.8%
- Max: 94.3%
- Interquartile range: 11.0%

Disparities

- The developer provided disparity data for 182,042 patients from July 1, 2017, through June 30, 2019.
- The data was grouped by sex, SUD diagnosis (diagnosed/not diagnosed), dual status, race, diagnosis (schizophrenia, MDD, bipolar disorder), and age.
- Medication continuation rate across all IPFs (n = 1,680):
  - Sex, male: 72.1%, female: 77.9%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.39
  - Substance use disorder (SUD) diagnosis, diagnosed with SUD: 70.4%, not diagnosed with SUD: 76.9%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.41
  - Dual status, dual: 77.4%, not dual: 69.8%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.51
  - Race, non-White: 71.1%, White: 76.2%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.31
  - Diagnosis, schizophrenia: 75.5%, major depressive disorder (MDD): 74.2%, bipolar disorder: 75.3%
    - Effect size (Eta2) for differences in means between patient groups: 0.001
  - Age, 18–39: 74.0%, 40–59: 74.1%, 60 and older: 75.4%
    - Effect size (Eta2) for differences in means between patient groups: 0.004
- Medication continuation rate across IPFs with at least 75 eligible cases in the denominator (n = 1,066):
  - Sex, male: 72.2%, female: 78.0%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.64
  - SUD diagnosis, diagnosed with SUD: 69.7%, not diagnosed with SUD: 77.4%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.74
  - Dual status, dual: 77.6%, not dual: 69.1%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.85
  - Race, non-White: 71.2%, White: 76.3%
    - Effect size (Cohen's d) for differences in means between patient groups: 0.46
  - Diagnosis, schizophrenia: 76.1%, MDD: 73.2%, bipolar disorder: 75.2%
    - Effect size (Eta2) for differences in means between patient groups: 0.013
  - Age, 18–39: 74.7%, 40–59: 74.8%, 60 and older: 74.9%
    - Effect size (Eta2) for differences in means between patient groups: 0.000
- Note on interpretation of effect size: Cohen's d: 0.2 is considered a small effect size, 0.5 is a medium effect size, and 0.8 is a large effect size, Eta2: 0.01 is small, 0.06 is medium, and 0.14 is large

Questions for the Committee:

- Does the Committee agree with the staff assessment that there is a gap in care that warrants a national performance measure?

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

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures–are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure.

- I agree with the assessment that the guidelines/evidence provided does not directly focus on the timeliness of medication use, but that tangentially adherence to medications is correlated to outcomes. There continues to be limited evidence regarding the timeliness of medications.
- Developer did not provide updated evidence for MDD and bipolar disorder; the evidence provided was the same as the 2016 submission.
- high
- Evidence is sufficient.
- The evidence appears to directly relate to medication adherence as a desired outcome.
- evidence is moderate
- Directionally consistent. But most of the evidence presented looked at absolute effectiveness and not continuation data (but there are lots of data on continuation of meds for MDD).
- No.
- Evidence applies directly to process being measured. The process of (and interventions related to) filling an Rx prior to discharge establishes medication continuation from the inpatient to outpatient setting and is associated with increased medication compliance and prevention of negative outcomes associated with nonadherence. I am not aware of any new information that changes the evidence base for this measure that has not been cited in the submission.
- This developer includes evidence from clinical practice guidelines and research studies conducted on medication continuation post discharge from IPFs. The guidelines contain explicit recommendations for medication adherence for the three diagnoses the developer includes in the measure. The information applies directly. Per the evidence, the process, which is filling a prescription post discharge, is directly related to the outcomes of improved medication adherence, reduced readmissions and improved management of symptomology. I am not aware of additional evidence that would change the evidence base, I was surprised to not see more recent references included for studies.
- Empirical data cited by submission applies directly to the measure (e.g. updated APA guidelines).
- In August 2017, "Committee agreed there is evidence that lack of adherence to medication leads to relapse and negative outcomes. They also noted that claims data related to medication adherence are directly correlated to outcome." Does still seem to be a lack of evidence around "timeliness of use" worth discussing as a committee.

- There is adequate evidence that medication continuation does improve outcomes, although there is not great evidence that the specific window of -2 to 30 days after discharge is the optimal range to use.
- Process

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- I agree that there is a gap in care that supports this measure. There are disparities as presented that this measure would help address.
- The developer reported that literature review from January 2016 through August 2020 did not reveal any new evidence regarding performance gaps since the initial endorsement. But isn't this required for maintenance? Also, all the data is from Medicare claims only.
- high median scores but obvious room for improvement
- Sufficiently demonstrated a performance gap. Data on the measure by population subgroups displayed, demonstrating disparities in care (although would like to see more subdivision amongst the race subgroup).
- The submission includes demonstration of a significant performance gap among surveyed facilities.
- Yes, stratifications also provided by sex, SUD, dual, race, dx, age groups
- Yes, and they actually looked at disparities.
- Gap exists.
- There is high variability in medication continuation rates across the inpatient psychiatric facilities examined, which suggests this measure can provide important feedback to facilities on how they can improve their care transitions. Subgroup data was provided demonstrating lower rates of medication continuation for comorbid SUD, for non-white patients, for male patients, and for younger patients (which is consistent with the literature).
- Yes, performance data was provided following review of Medicare claims data. Data was collected on 1,680 IPFs, with a mean medication continuation at 75% and a standard deviation of 12.8%, demonstrating room for improvement. At the facility-level, there is clearly a gap in performance with regard to medication continuation post discharge. The data was provided for different populations and subgroups. Yes, it does demonstrate disparities between men and women, SUD, diagnosis and race. Less disparities exist among different age groups.
- A gap in care exists. Data on disparities based on gender, race, and diagnosis were provided. Disparities exist between subgroups.
- Does seem to be sizable variability, suggesting major room for improvement. In terms of disparities, are medium to large effect sizes based on groupings by sex, SUD diagnosis (diagnosed/not diagnosed), dual status, race, suggesting it IS important to address disparities as performance focus.
- The performance gap here is smaller than ones we typically endorse. However it is real and important. Disparities data seems real and is important also.
- Yes there is room for improvement. Men are not as likely to refill scripts so additional outreach would be beneficial and more outreach is need for the Medicare population.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if 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 enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

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

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

NQF Staff Scientific Acceptability Preliminary Analysis

Reliability

- Developer calculated signal-to-noise reliability using the beta-binomial model to determine what proportion of variation between IPFs is due to real differences in facility characteristics as opposed to sampling error.
    - The reliability statistic, R, ranges from 0 to 1, with R=0 indicating that all variation is due to sampling error and R=1 indicating that all variation is due to real differences between facilities.
    - The developer defined the threshold for acceptable reliability as 0.7 and calculated the mean and range of the reliability statistic for each individual facility that had at least 75 denominator cases (1,066 facilities).
- Mean reliability was 0.78.
- The 25th percentile across all 1,066 facilities exceeded 0.7. The 75th percentile was 0.81.

Validity

- Developer evaluated validity using the "known group" method.
    - This method determines whether the measure score can be used to discriminate between patient sub-groups that are known to have differences in rates according to the literature.

    - Developer identified predefined patient subgroups known to have lower rates of medication continuation based on evidence from peer-reviewed studies which examined factors related to nonadherence to psychotropic medication in patients with psychiatric disorders
- Differences in mean facility scores were examined for the following patient subgroups:
    - Patients aged 40 or younger, or "younger patients"
    - Male patients
    - Patients with a comorbid Substance Use Disorder (SUD) diagnosis
    - Patients with a diagnosis of schizophrenia
    - Disadvantaged patients with problems accessing medication/limited socioeconomic resources (Medicare-Medicaid status was used as a proxy for socioeconomic status)
- Developer calculated rates for each patient subgroup by facility then calculated the mean rates and standard deviations by subgroup across all facilities.
    - T-tests were used to compare mean group differences for dichotomous variables.
    - Developer also calculated Cohen's d effect size (difference in mean scores divided by the pooled standard deviation) to compare groups.
    - For patient subgroups with more than two categories (age and diagnosis), the developer computed Eta-squared ($\eta 2$) effect size to capture the overall difference in the measure rate between groups.
    - The developer determined that the medication continuation rates for the sub-groups were consistent with the literature.
- Developer observed lower Medication Continuation measure rates for patients with comorbid SUD, for non-white patients, for male patients, and for younger patients.
- The Medication Continuation measure did not follow known group expectations in medication adherence rates between patients 1) enrolled in Medicare only and those with both Medicare and Medicaid coverage and 2) with different principal diagnosis at discharge.

Developer also presented validity results from the previous submission as well:

- Data element validity testing
    - Two psychiatrists reviewed 150 patient records.
        - Clinicians' recorded assessments of principal discharge diagnosis were compared to claims.
        - Positive predictive value (PPV) was calculated using the clinical assessment from the medical record as the gold standard. Note: a high PPV indicates high probability that a claim for a specific condition correctly predicts the diagnosis at discharge in the medical record.
    - Additionally, abstractors at 7 sites indicated whether a prescription was provided at discharge, and if not, to provide a rationale in order to determine if additional exclusions were needed.
        - Data on provision of at least one prescription was compared to claims data.
        - PPV was calculated indicating that most patients who filled a prescription in the follow-up period also received a prescription at discharge.
    - Abstractors at the 7 sites also recorded if the medical record indicated medications were dispensed to the patient free at discharge (since those would not be reflected in claims data).
    - Ten percent of all abstraction cases were reviewed by both clinicians.
    - PPV of claims data was 97%. (MDD – 98%; schizophrenia – 98%; bipolar disorder -96%).
    - For the medical record review, 92% of cases were prescribed medication at discharge ; PPV was 96%.
    - Few discharges included provision medications at discharge.

- Measure score validity testing
    - Measure scores were compared to three related measures using convergent validity analysis:
        - Follow-Up After Hospitalization (7-Day)
        - Follow-Up After Hospitalization (30-Day)
        - IPF All-Cause Unplanned Readmission Measure
    - The developer hypothesized the first two measures would be positively correlated with the medication continuation scores, as they all reflect care coordination. The developer hypothesized that the medication continuation score would be negatively correlated with the all-cause unplanned readmission score, "because readmissions may indicate a lack of care coordination."
    - Developer conducted a Spearman's rank correlation:
        - Follow-Up After Hospitalization 7-day: 0.34
        - Follow-Up After Hospitalization 30-day: 0.43
        - IPF All-Cause Unplanned Readmission Measure (Observed): -0.26
- Face validity
    - Face validity of the measure score was assessed by a technical expert panel. Members were asked if they agreed if the performance rating as specified accurately represents facility-level rates of medication continuation.

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- Do you agree with the staff assessment of the reliability testing for the measure?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- Do you agree with the staff assessment of the validity testing for the measure?

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

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

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- I agree with the staff assessment for reliability and validity. No additional concerns.
- It would be interesting to know more about the relationship between the number of filled prescriptions and actual medication adherence. Given new digital platforms for medication adherence- perhaps it's worth exploring a new "proxy" for medication adherence than filled prescriptions.
- non
- No concerns.
- Data elements appear to be clearly defined. The measure can be consistently implemented.
- clear specifications, good exclusion conditions, data dictionary appreciated
- OK

- Data elements are clear.
- Data elements are clearly defined, with adequate sampling, and good reliability demonstrated through a standard methodology. I do not have concerns about the likelihood that this measure can be consistently implemented.
- Data elements are clearly defined. Codes with descriptors are provided and the steps are clear. I do not have concerns that this measure can be consistently implemented as reliability estimates are high.
- I have no concerns about implementation and the data elements are clearly defined.
- Measure specifications seem adequate.
- It's reasonably reliable.
- No concerns

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- No
- Moderate
- no
- No.
- No
- No. More comprehensive approach. In addition to signal-to-noise, and face validity, examined convergent validity using 3 measures with hypothesized positive and negative correlations, stratifications by pre-defined "known groups" was a plus
- No, it's ok
- No concerns.
- No because signal-to-noise, R, IQR for each IPF were all were within acceptable range.
- I agree with the NQF staff and do not have concerns about reliability.
- No
- Beta-Binomial methodology to measure signal-to-noise ratio seems adequate, though this is new to me. I don't see a need to discuss reliability testing, but also defer to those with more experience here.
- No
- No concerns.

2b1. Validity -Testing: Do you have any concems with the testing results?

- No
- Moderate
- no
- No.
- No
- no
- not really--I do wonder about the list of meds and individual validity by med--but that's a swamp.
- No concerns.
- I do not.
- I do not.
- No

- Face validity and empirical validity tests seem adequate. No concerns.
- Validity seems adequate
- No concerns.

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment) 2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure? 2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- Not applicable
- no concerns
- what happens if no meds are prescribed? or client leaves AMA?
- Acceptable.
- The exclusions are consistent with the evidence. However, it is unclear to me why enrollees in Medicare Advantage plans were excluded from the numerator.
- unable to adjust for social risk factors (e.g., homelessness) given claims data
- The usual concerns about risk adjustment or not
- No concerns.
- No concerns about the patients excluded from the measure. Risk adjustment - n/a.
- All exclusions are in alignment with the evidence and no groups are excluded inappropriately. The developer does not include information on risk adjustment as it is not applicable since it looks at claims for typical care.
- N/A - measure is not risk adjusted
- Exclusions appear appropriate. Risk adjustment or stratification N/A.
- NA
- No risk adjustment in this measure.

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data) 2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- No concerns.
- No concerns
- none
- I am satisfied with the testing of the measure's validity.
- I was not able to discern any threats to validity in the submission.
- no
- Seems reasonable.
- No concerns.

- IQR and FAPH measure demonstrated IPFs' performance is significantly different. Comparability - n/a. Missing data - no identified cases of missing or unreliable data.
- This measure includes data on medication continuation post discharge using Medicare claims. Analyses of this measure indicate that there are meaningful differences in quality at the facility level and among differences in population subgroups. Yes, measure is specified precisely indicating comparable results. The developer reports no discharges with missing data and that missing data should not be present since using claims.
- No
- No concerns. Missing data not expected to be an issue. Confidence interval approach addresses meaningful differences. Comparability N/A as measure has only one set of specifications.
- Nothing significant.
- No threat to validity as data should be available unless they paid cash for a prescription by why do so if they had coverage.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

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.
   - Developer notes that the data are elements are generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)
   - Developer notes that data elements are coded by someone other than person obtaining original information (e.g., DRG, ICD-10 codes on claims)
   - Developer notes that all data elements are in defined fields in electronic claims

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?

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

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?

- The required data are easily generated from EHR/claims data. No concerns.
- Is this measure limited to Medicare patient?
- appears feasible in pilot study
- Feasible.
- This is a limited process measure and should be easy to operationalize.
- feasible using claims data

- Feasible.
- Feasibility is okay.
- No concerns regarding feasibility. Data elements are generated or collected by healthcare personnel during the provision of care and all data elements are in defined fields in electronic claims.
- The data are routinely generated during usual care delivery and then coded afterward. The data are available in electronic form.
- No concerns about feasibility since the measure uses data that are routinely collected
- No concerns about feasibility – data elements are all part of regular claims. As this will be a new measure, may be important to vet with providers in clinical delivery settings.
- It is feasible
- No concerns this is straight forward.

Criterion 4: Usability and Use

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. 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 not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details

- Developer notes that CMS plans to include the measure for use in the IPFQR program, a national pay-for-reporting program with publicly reported results at the facility level, for the first time for FY 2021

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; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- Developer notes that the IPFs have not yet received scores, as the measure will be used in the IPFQR program for the first time for FY 2021. CMS plans to monitor stakeholder feedback going forward.

Additional Feedback: N/A

***Questions for the Committee**:*

- How can the performance results be used to further the goal of high-quality, efficient healthcare?
- How can the measure be vetted in real-world settings by those being measured or others?

Preliminary rating for Use: ☒ Pass ☐ No Pass

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- N/A-The developer notes that the measure will be implemented in FY 2021. Once implemented, facility-level medication continuation scores in Medicare FFS claims data will be calculated and provided to facilities to encourage quality improvement, specifically related to stronger care transitions to outpatient settings.

4b2. Benefits vs. harms. 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

- N/A

Potential harms

- Developer notes that medications used to treat patients with MDD, schizophrenia, and bipolar disorder may cause a range of harmful side effects at varying rates depending on clinical and personal characteristics.
- Developer notes that clinical guidelines indicate that the benefits of using medications associated with the treatment of these conditions outweighs the harm
- Developer asserts that the implementation of this measure will result in quality improvement by identifying patients who do not adhere to medication continuation post-discharge. Improved medication continuation would help reduce negative outcomes.

Additional Feedback: None

***Questions for the 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?

Preliminary rating for Usability: ☐ High ☒ Moderate ☐ Low ☐ Insufficient

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided? 4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or

other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- Yes. I believe the plans to include the measure in the IPFQR as well as feedback processes are well considered.
- Evidence demonstrates usable for Medicare patients only.
- yes
- Yes
- It is unclear from the submission if the measure is being public reported by CMS or if the facilities being measured are given performance results.
- pending use for IPFQR in FY2021
- The fact that this measure is planned for CMS measurement, but not implemented, does concern me.
- Results are reported.
- Measure is being publicly reported and CMS plans to include the measure for use in the IPFQR program for the first time in FY 2021. Those being measured will be provided with the results and have been invited to participate in feedback/measure implementation.
- The measure is not currently publicly reported, but CMS plans to include it in the IPFQR program for FY21. Results at the facility level will be publicly reported, the results will be disclosed and available outside of the organizations and practices whose performance is measured. Data will be provided in 2021 to those being measured. CMS will run data and provide IPFs nationwide with their measure scores, and mean state and national scores. CMS plans to provide user guidelines to IPFs once they receive their data and will offer webinars to explain the data provided. They will monitor stakeholder feedback once this process begins, but have not received feedback yet.
- N/A, as this measure is not in use; however developer states that facilities have not yet received scores, as the measure will be used in the IPFQR program for the first time for FY 2021 and CMS will monitor feedback going forward.
- No immediate concerns, but will be important to monitor IPFQR program reporting and stakeholder feedback, as this is being implemented for the first time in 2021.
- Not a problem
- Yes

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations? 4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- I agree the benefits outweigh any potential harms. If patients are being admitted to hospital for these conditions the use of medications (and adherence to them) outweighs known harmful side effects.
- unclear if benefits outweigh harms due to medication side effects particularly for metabolic diseases.
- is this a system of care measure; or a hospital measure? who is accountable for completion of outpatient dispensing? the data says "facility"--who is that entity?
- This assess prescriptions filled, based on claims, 30-days post discharge. Doesn't the fact that the patient has moved from one setting to another result in an accountability challenge: first prescription in hospital and second (30 -days) in a different setting? How is this info shared and who bears the responsibility for improvement?
- Medication adherence is a major issue in driving quality of care in depression, schizophrenia and bipolar disorder. Benefits far exceed any potential harm.

- agree with unintended consequences stated by developer--this application was thoughtfully written
- Benefits outweigh harms.
- Measure doesn't address if medication is taken, only if it is picked up.
- Once the measure is implemented, Medicare FFS claims data will be calculated and provided to facilities to encourage quality improvement. Benefits of continuous care and use of medication for treatment of the chronic psychiatric conditions in question outweigh the potential for unintended negative consequences and/or harm due to side effects (per clinical guidelines).
- If medication continuation following discharge from IPFs improves health outcomes, then providing IPFs with their scores will allow them to engage in quality improvement to address this aspect of care through intervention strategies such as reminders, patient education, problem solving, etc. Of course, there are risks with these medications as discussed by the developer, but this is addressed through the clinical guidelines provided. Yes, a credible rationale is provided. Since the measure relies on claims data for services already rendered, there are no unintended consequences of reviewing the data. Supporting increased medication continuation does come with some risks, as noted by the developer. There are risks associated with psychiatric medications, but clinical guidelines do recommend their continuation and that the benefits of continuation outweigh the risks.
- No harms identified. Measuring continuity of medication following discharge may help patients stay healthy and reduce readmission.
- No immediate concerns, but will be important to monitor IPFQR program reporting and stakeholder feedback, as this is being implemented for the first time in 2021.
- Usability is fine
- Benefits of the medications for these diagnosis outweigh potential harms.

Criterion 5: Related and Competing Measures

Related or competing measures

- 1879: Adherence to Antipsychotic Medications for Individuals with Schizophrenia
- 1880: Adherence to Mood Stabilizers for Individuals with Bipolar I Disorder
- Antidepressant Medication Management (AMM) from the National Committee for Quality Assurance´s (NCQA) Healthcare Effectiveness Data and Information Set (HEDIS) 2019 (Not NQF endorsed)

Harmonization

- Developer notes the nominator for the Medication Continuation measure has been harmonized with the identified measures to the extent possible because the measure populations of the related measures overlap with the patient population targeted by this measure and the measures share a similar clinical focus on medication use
- Developer compared the medications included in the related measures with medications included in this measure

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?

- Yes as noted. They are reasonably harmonized.
- The measure does not seem fully harmonized. There is overlap.
- reportedly already harmonized.
- The submission notes that the nominator for the Medication Continuation measure has been harmonized with the identified measures to the extent possible because the measure populations of the related measures overlap with the patient population targeted by this measure and the measures share a similar clinical focus on medication use.
- 1879, 1880, and AMM. No anticipated additional steps
- Actually, another example of a plethora of measures, all subtly different, measuring the same concept--I wish NQF could better reduce overlapping measures.
- no.
- 1879, 1880, and Antidepressant Medication Management (AMM) from the National Committee for Quality Assurance´s (NCQA) Healthcare Effectiveness Data and Information Set (HEDIS). The nominator for the Medication Continuation measure has been harmonized with the identified measures to the extent possible.
- Three other measures are listed by the developer and an explanation for how they have been harmonized has been included.
- Potential overlap with NQF1879 and NQF1880 but this measure also includes MDD. The numerator for the Medication Continuation measure has been harmonized with these measures when possible but are not directly competing because the Medication Continuation measure is for those with diagnoses of bipolar disorder, MDD, or schizophrenia.
- While there are related measures, they do not appear to be competing in any way, and measure specifications are adequately harmonized.
- Three individual measures (for MDD, Schizophrenia, and Bipolar Affective Disorder) exist and there has been some attempt to harmonize the specifics. Given that the measurement world is replete with many individual measures and that more keep coming; and that the consequent burden on the delivery system is already difficult to sustain, has there been any consideration in withdrawing the 3 aforementioned measures and replacing them with one measure such as this?
- None

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 01/15/2021

- No NQF Members have submitted support/non-support choices as of this date.
- No Public or NQF Member comments submitted as of this date.

NQF Staff Scientific Acceptability Evaluation

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 3205

Measure Title: Medication Continuation Following Inpatient Psychiatric Discharge

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☒ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☐ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☒ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**
   - None identified by staff.

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2
   - Developer calculated signal-to-noise reliability using the beta-binomial model to determine what proportion of variation between IPFs is due to real differences in facility characteristics as opposed to sampling error
   - The reliability statistic, R, ranges from 0 to 1, with R=0 indicating that all variation is due to sampling error and R=1 indicating that all variation is due to real differences between facilities.
   - Developer calculated signal-to-noise reliability using a three-step approach:
     - The developer first calculated the variance within the individual facilities (noise).

- Using version 2.2. of the BETABIN SAS macro to fit the beta binomial model, the developer calculated an estimate of the variance between facilities (signal).
- Lastly, the developer calculated signal-to-noise reliability as a ratio of the variance between facilities and the sum of the variance between facilities plus the variance within facilities.

7. **Assess the results of reliability testing**
   - Developer defined the threshold for acceptable reliability as 0.7
   - Developer calculated the mean and range of the reliability statistic for each individual facility with at least 75 denominator cases (1,066 facilities).
   - Mean reliability was 0.78.
   - The 25th percentile across all 1,066 facilities exceeded 0.7. The 75th percentile was 0.81.

   Submission document: Testing attachment, section 2a2.3
8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)
   - Signal-to-noise reliability testing is a standard approach for assessing variability due to quality differences among measured entities.
9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)
10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)
11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**
    - (Box 1) → Measure specifications precise, unambiguous, and complete (Box 2) → Empirical testing conducted using statistical tests → (Box 4): Reliability testing conducted with computed performance measure scores → (Box 5): Method described and appropriate for assessing the proportion of variability due to real differences among measured entities → (Box 6a) MODERATE

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.

- None identified by staff.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**
    - None identified by staff.

    Submission document: Testing attachment, section 2b4.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.
    - Not applicable. This measure has a single data source.

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.
    - None identified by staff. The developer evaluated process claims data for this measure and determined that there was no missing data.

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

    16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

    ☒ Yes ☐ No ☐ Not applicable

    16c. **Social risk adjustment:**

    16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

    16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☒ No

    16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☒ No

    16e**. Assess the risk-adjustment approach**
    - Not applicable. Process claims data includes information regarding age, race, gender, and payer. Because this measure is based on processes that are expected to be carried out for all patients, this measure is not risk adjusted.

VALIDITY: TESTING

17. **Validity testing level:** ☒ **Measure score** ☒ **Data element** ☒ **Both**
18. **Method of establishing validity of the measure score:**

    ☒ **Face validity**

    ☒ **Empirical validity testing of the measure score**

    ☐ **N/A (score-level testing not conducted)**

19. **Assess the method(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.2
    - Developer evaluated validity using the "known group" method.
        - This method determines whether the measure score can be used to discriminate between patient sub-groups that are known to have differences in rates according to the literature.
        - Developer identified predefined patient subgroups known to have lower rates of medication continuation based on evidence from peer-reviewed studies which examined factors related to nonadherence to psychotropic medication in patients with psychiatric disorders
    - Differences in mean facility scores were examined for the following patient subgroups:
        - Patients aged 40 or younger, or "younger patients"

    - Male patients
    - Patients with a comorbid Substance Use Disorder (SUD) diagnosis
    - Patients with a diagnosis of schizophrenia
    - Disadvantaged patients with problems accessing medication/limited socioeconomic resources (Medicare-Medicaid status was used as a proxy for socioeconomic status)
- Developer calculated rates for each patient subgroup by facility then calculated the mean rates and standard deviations by subgroup across all facilities.
    - T-tests were used to compare mean group differences for dichotomous variables.
    - Developer also calculated Cohen's d effect size (difference in mean scores divided by the pooled standard deviation) to compare groups.
    - For patient subgroups with more than two categories (age and diagnosis), the developer computed Eta-squared (η2) effect size to capture the overall difference in the measure rate between groups.
    - The developer determined that the medication continuation rates for the sub-groups were consistent with the literature.
- Developer observed lower Medication Continuation measure rates for patients with comorbid SUD, for non-white patients, for male patients, and for younger patients.
- The Medication Continuation measure did not follow known group expectations in medication adherence rates between patients 1) enrolled in Medicare only and those with both Medicare and Medicaid coverage and 2) with different principal diagnosis at discharge.

Developer also presented validity results from the previous submission as well:

- Data element validity testing
    - Two psychiatrists reviewed 150 patient records.
        - Clinicians' recorded assessments of principal discharge diagnosis were compared to claims.
        - Positive predictive value (PPV) was calculated using the clinical assessment from the medical record as the gold standard. Note: a high PPV indicates high probability that a claim for a specific condition correctly predicts the diagnosis at discharge in the medical record.
    - Additionally, abstractors at 7 sites indicated whether a prescription was provided at discharge, and if not, to provide a rationale in order to determine if additional exclusions were needed.
        - Data on provision of at least one prescription was compared to claims data.
        - PPV was calculated , indicating that most patients who filled a prescription in the follow-up period also received a prescription at discharge.
    - Abstractors at the 7 sites also recorded if the medical record indicated medications were dispensed to the patient free at discharge (since those would not be reflected in claims data).
    - Ten percent of all abstraction cases were reviewed by both clinicians.
    - PPV of claims data was 97%. (MDD – 98%; schizophrenia – 98%; bipolar disorder -96%).
    - For the medical record review, 92% of cases were prescribed medication at discharge ; PPV was 96%.
    - Few discharges included provision medications at discharge.
- Measure score validity testing
    - Measure scores were compared to three related measures using convergent validity analysis:
        - Follow-Up After Hospitalization (7-Day)
        - Follow-Up After Hospitalization (30-Day)
        - IPF All-Cause Unplanned Readmission Measure

- The developer hypothesized the first two measures would be positively correlated with the medication continuation scores, as they all reflect care coordination. The developer hypothesized that the medication continuation score would be negatively correlated with the all-cause unplanned readmission score, "because readmissions may indicate a lack of care coordination."
- The developer appears to have done a Pearson correlation, which measures the degree of association between two quantitative variables. For the social sciences, scores of 0.37 or larger are considered to have a "large" correlation effect. (Medium effect is 0.24 – 0.36 and small effect is 0.10 – 0.23.)
  - Follow-Up After Hospitalization 7-day: 0.34
  - Follow-Up After Hospitalization 30-day: 0.43
  - IPF All-Cause Unplanned Readmission Measure (Observed): -0.26

- Face validity
  - Face validity of the measure score was assessed by a technical expert panel. Members were asked if they agreed if the performance rating as specified accurately represents facility-level rates of medication continuation.

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b2.3

- Developer determined that the medication continuation rates for the sub-groups were consistent with the literature.
- Developer observed lower Medication Continuation measure rates for patients with comorbid SUD, for non-white patients, for male patients, and for younger patients.
- The Medication Continuation measure did not follow known group expectations in medication adherence rates between patients 1) enrolled in Medicare only and those with both Medicare and Medicaid coverage and 2) with different principal diagnosis at discharge.
- Results from previous submission provide a complementary picture on the validity of the measure.
- Methods used are common approaches to measure validity testing.

21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

Submission document: Testing attachment, section 2b1.

☒ **Yes**

☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**

NOTE that data element validation from the literature is acceptable.

Submission document: *Testing attachment, section 2b1.*

☒ **Yes**

☐ **No**

☐ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

    (Box 1)-All potential threats to validity assessed → (Box 2) Empirical validity testing conducted using the measure as specified and appropriate statistical testing→ (Box 6) Validity testing conducted with computed performance measure scores of each measured entity → (Box 7) Method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships → (Box 8b) Moderate certainty or confidence that the performance measure scores are a valid indicator of quality-MODERATE

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

    None identified by staff.

Developer Submission

NQF #: 3205

Corresponding Measures:

De.2. Measure Title: Medication Continuation Following Inpatient Psychiatric Discharge

Co.1.1. Measure Steward: Centers for Medicare & Medicaid Services

De.3. Brief Description of Measure: This measure assesses whether patients discharged from an inpatient psychiatric facility (IPF) with major depressive disorder (MDD), schizophrenia, or bipolar disorder filled a prescription for evidence-based medication within 2 days prior to discharge and 30 days post-discharge. This measure evaluates admissions over a two-year period.

1b.1. Developer Rationale: The aim of the measure is to address gaps in continuity of pharmaceutical treatment during the transition from inpatient to outpatient care. Pharmacotherapy is the primary form of treatment for most patients discharged from an inpatient psychiatric facility (IPF) for bipolar disorder, major depressive disorder (MDD), or schizophrenia. The measure focuses on medication continuation because it is an essential step in medication adherence.

Medication continuation is particularly important in the psychiatric patient population because psychotropic medication discontinuation can have a range of adverse effects, from mild withdrawal to life-threatening autonomic instability and psychiatric decompensation (Ward & Schwartz, 2013). Patients with MDD who do not remain on prescribed medication are more likely to have negative health outcomes, such as relapse and readmission, decreased quality of life, and increased health care costs. If untreated, MDD can contribute to or worsen chronic medical disorders (Geddes et al., 2003; Glue et al., 2010). The literature shows that among patients with schizophrenia, those who were "good compliers" according to the Medication Adherence Rating Scale had better outcomes in terms of rehospitalization rates and medication maintenance (Jaeger et al., 2012). Among patients with bipolar disorder, medication adherence was significantly associated with reduction in manic symptoms (Sylvia et al., 2013), whereas nonadherence was associated with increased suicide risk (OR 10.8, CI 1.57–74.4; Gonzalez-Pinto et al., 2006). Our literature review from January 2016 through August 2020 did not reveal any new evidence regarding performance gaps since the initial endorsement submission.

Current facility-level performance indicates a clear quality gap. Using Medicare claims data from July 1, 2017, through June 20, 2019, the Medication Continuation measure rates ranged from 34.8 to 94.3%, with a median of 76.2%. There was a 21.3 percentage point difference between the 10th and 90th percentiles (63.4–84.7%). Using 2013–2014 Medicare claims data, there was a 21.6 percentage point difference between the 10th and 90th percentiles (66.7–88.3%) and a median score of 79.6%. By calculating the facility-level rates of medication continuation in Medicare fee-for-service (FFS) claims data, this measure can provide valuable information on areas where care transitions to the outpatient setting can be improved.

Literature about continuation of medication has identified effective interventions that facilities can employ to improve medication adherence among patients discharged from an IPF (Douaihy, Kelly, & Sullivan, 2013; Haddad, Brain, & Scott, 2014; Hung, 2014; Kasckow & Zisook, 2008; Lanouette, Folsom, Sciolla, & Jeste, 2009; Mitchell, 2007; Sylvia et al., 2013). Examples of these interventions include patient education, shared decision making, and text-message reminders. We envision the addition of this measure to the suite of measures for IPFs would help to create a comprehensive picture of the quality of care patients receive at those facilities.

*Douaihy, A. B., Kelly, T. M., Sullivan, C. (2013). Medications for substance use disorders. Social Work in Public Health, 28(3-4), 264-278. doi: 10.1080/19371918.2013.759031

*Geddes, J. R., Carney, S. M., Davies, C., Furukawa, T. A., Kupfer, D. J., Frank, E., & Goodwin, G. M. (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. The Lancet, 361(9358), 653–661. doi:10.1016/s0140-6736(03)12599-8

*Glue, P., Donovan, M. R., Kolluri, S., & Emir, B. (2010). Meta-analysis of relapse prevention antidepressant trials in depressive disorders. Australian and New Zealand Journal of Psychiatry, 44(8), 697-705. doi: 10.3109/00048671003705441

*Gonzalez-Pinto, A., Mosquera, F., Alonso, M., López, P., Ramírez, F., Vieta, E., & Baldessarini, R. J. (2006). Suicidal risk in bipolar I disorder patients and adherence to long-term lithium treatment. Bipolar Disorders, 8(5p2), 618–624. doi:10.1111/j.1399-5618.2006.00368.x

*Haddad, P. M., Brain, C., & Scott, J. (2014). Nonadherence with antipsychotic medication in schizophrenia: Challenges and management strategies. Patient Related Outcome Measures, 5, 43-62. doi: 10.2147/PROM.S42735

*Hung, C. I. (2014). Factors predicting adherence to antidepressant treatment. Current Opinion in Psychiatry, 27(5), 344-349. doi: 10.1097/yco.0000000000000086

*Jaeger, S., Pfiffner, C., Weiser, P., Kilian, R., Becker, T., Langle, G., . . . Steinert, T. (2012). Adherence styles of schizophrenia patients identified by a latent class analysis of the Medication Adherence Rating Scale (MARS): A six-month follow-up study. Psychiatry Research, 200(2-3), 83-88. doi: 10.1016/j.psychres.2012.03.033

*Kasckow, J. W., & Zisook, S. (2008). Co-occurring depressive symptoms in the older patient with schizophrenia. Drugs & Aging, 25(8),631-647.

*Lanouette, N. M., Folsom, D. P., Sciolla, A., Jeste, D. V. (2009). Psychotropic medication nonadherence among United States Latinos: A comprehensive literature review. Psychiatric Services (Washington, DC), 60(2), 157-174. doi: 10.1176/appi.ps.60.2.157

*Mitchell, A. J. (2007). Understanding medication discontinuation in depression. BMedSci Psychiatric Times, 24(4).

*Sylvia, L. G., Hay, A., Ostacher, M. J., Miklowitz, D. J., Nierenberg, A. A., Thase, M. E., Perlis, R. H. (2013). Association between therapeutic alliance, care satisfaction, and pharmacological adherence in bipolar disorder. Journal of Clinical Psychopharmacology, 33(3), 343-350. doi: 10.1097/JCP.0b013e3182900c6f

*Ward, M., & Schwartz, A. (2013). Challenges in pharmacologic management of the hospitalized patient with psychiatric comorbidity. Journal of Hospital Medicine, 8(9), 523–529. doi:10.1002/jhm.2059.

S.4. Numerator Statement: The numerator for the measure includes:

- Discharges with a principal diagnosis of MDD in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge
- Discharges with a principal diagnosis of schizophrenia in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge
- Discharges with a principal diagnosis of bipolar disorder in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge

S.6. Denominator Statement: The target population for this measure is Medicare fee-for-service (FFS) beneficiaries with Part D coverage aged 18 years and older discharged from an inpatient psychiatric facility with a principal diagnosis of MDD, schizophrenia, or bipolar disorder.

S.8. Denominator Exclusions: The denominator for this measure excludes discharged patients who:

- Received electroconvulsive (ECT) during the inpatient stay or follow-up period
- Received transcranial stimulation (TMS) during the inpatient stay or follow-up period
- Were pregnant at discharge
- Had a secondary diagnosis of delirium at discharge

- Had a principal diagnosis of schizophrenia with a secondary diagnosis of dementia at discharge

De.1. Measure Type: Process

S.17. Data Source: Claims

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Jun 28, 2017 **Most Recent Endorsement Date:** Jun 28, 2017

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable because this measure is not paired or grouped with another measure.

1. Evidence and Performance Gap – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

2020_Medication_Continuation_evidence_attachment.docx,Updated_2020_Medication_Continuation_evidence_attachment.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

No

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 3205

Measure Title: Medication Continuation Following Inpatient Psychiatric Discharge

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here:

Date of Submission: 11/2/2020

Please note: 2016 submission text in blue | 2020 submission text in red

1a.1. This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Health outcome:

☐ Patient-reported outcome (PRO):

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*):

☒ Process: Patient fills prescription, establishing medication continuation from the inpatient to the outpatient setting. (No change for 2020.)

☐ Appropriate use measure:

☐ Structure:

☐ Composite:

1a.12 LOGIC MODEL Diagram or 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.

2020 submission: No change for 2020.

2016 submission: Effective interventions have been identified that can improve medication adherence during the transition from inpatient to outpatient care. Interventions that have been shown to increase medication compliance and prevent negative outcomes associated with nonadherence include patient education, enhanced therapeutic relationships, shared decision-making, and text-message reminders, with emphasis on multidimensional approaches (Douaihy, Kelly, & Sullivan, 2013; Haddad, Brain, & Scott, 2014; Hung, 2014; Kasckow & Zisook, 2008; Lanouette, Folsom, Sciolla, & Jeste, 2009; Mitchell, 2007; Sylvia et al., 2013). Interventions, including those described by the literature, can be implemented during steps 2 and 3 in the logic model to influence medication continuation in step 4. Because the denominator only includes patients who would require continued evidence-based pharmacotherapy and who have few barriers to access, this measure provides an indirect quality indicator of the treatment provided in steps 2 and 3.

1) Patient is admitted for inpatient psychiatric care →
2) Patient receives treatment and is stabilized →
3) Patient is discharged with prescriptions for evidence-based medications and discharge treatment plan →
4) **Patient fills initial prescription, establishing medication continuation from the inpatient to the outpatient setting →**
5) Patient's symptoms are managed by pharmacotherapy →
6) Psychiatric decompensation and adverse outcomes such as emergency department visits, rehospitalization, and suicide are prevented.

*Douaihy, A. B., Kelly, T. M., & Sullivan, C. (2013). Medications for substance use disorders. *Social Work in Public Health*, *28*(3-4), 264-278. doi: 10.1080/19371918.2013.759031
*Haddad, P. M., Brain, C., & Scott, J. (2014). Nonadherence with antipsychotic medication in schizophrenia: Challenges and management strategies. *Patient Related Outcome Measures, 5*, 43-62. doi: 10.2147/PROM.S42735
*Hung, C. I. (2014). Factors predicting adherence to antidepressant treatment. *Current Opinion in Psychiatry, 27*(5), 344-349. doi: 10.1097/yco.0000000000000086
*Kasckow, J. W., & Zisook, S. (2008). Co-occurring depressive symptoms in the older patient with schizophrenia. *Drugs and Aging*, *25*(8), 631-647. doi: 10.2165/00002512-200825080-00002

*Lanouette, N. M., Folsom, D. P., Sciolla, A., & Jeste, D. V. (2009). Psychotropic medication nonadherence among United States Latinos: A comprehensive literature review. *Psychiatric Services, 60*(2), 157-174. doi: 10.1176/appi.ps.60.2.157
*Mitchell, A. J. (2007). Understanding medication discontinuation in depression. *Psychiatric Times, 24*(4).
*Sylvia, L. G., Hay, A., Ostacher, M. J., Miklowitz, D. J., Nierenberg, A. A., Thase, M. E., . . . Perlis, R. H. (2013). Association between therapeutic alliance, care satisfaction, and pharmacological adherence in bipolar disorder. *Journal of Clinical Psychopharmacology, 33*(3), 343-350. doi: 10.1097/JCP.0b013e3182900c6f

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES- State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process (e.g., intervention, or service).

2020 submission: Not applicable

2016 submission: Not applicable

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

What is the source of 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. (IOM)

☒ 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

2020 submission:

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	• Practice guideline for the treatment of patients with schizophrenia: 3rd edition • American Psychiatric Association (APA) • 2019 American Psychiatric Association. (2019). Practice guideline for the treatment of patients with schizophrenia: 3rd ed. Retrieved from https://www.psychiatry.org/psychiatrists/practice/clinical-practice-guidelines Please note, the American Psychiatric Association (APA) and the U.S. Departments of Veterans Affairs and Defense (VA/DOD) have not updated their guidelines for bipolar disorder or major depressive disorder since the initial endorsement submission in 2016. The content from these guidelines remains relevant for this measure.
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	"APA recommends (1A) that patients with schizophrenia be treated with an antipsychotic medication and monitored for effectiveness and side effects. APA recommends (grade: 1A) that patients with schizophrenia whose symptoms have improved with an antipsychotic medication continue to be treated with an antipsychotic medication. *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." p.5
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	2019 APA practice guideline for the treatment of schizophrenia evidence grade: A (high).
Provide all other grades and definitions from the evidence grading system	All other evidence grades from the 2019 APA schizophrenia guideline: B (moderate), C (low).
Grade assigned to the recommendation with definition of the grade	**Schizophrenia** The guideline from the APA (2019) to treat patients with schizophrenia with an antipsychotic medication and to continue to treat such patients whose symptoms have improved with an antipsychotic were graded I: confidence that the benefits of the intervention clearly outweigh harms.
Provide all other grades and definitions from the	All other recommendation grades from the 2019 APA schizophrenia guideline: II, suggestion (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

Systematic Review	Evidence
recommendation grading system	harms might be less clear).
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The 2019 APA schizophrenia guideline was developed in accordance with the Institute of Medicine (now known as the National Academy of Medicine) report "Clinical Practice Guidelines We Can Trust" (Institute of Medicine, 2011) and the "Principles for the Development of Specialty Society Clinical Guidelines" of the Council of Medical Specialty Societies (Institute of Medicine, 2012). The APA solicited input from subject matter experts and patient and family advocates. The guideline includes about 1,000 references. The APA identified evidence through literature searches and systematic review, and research and clinical experts provided input on topics for which there was not high quality evidence.
Estimates of benefit and consistency across studies	The APA found consistent benefits of evidence-based medication administration for those diagnosed with schizophrenia in terms of improved health outcomes and improved quality of life and functioning. It also found the benefits far outweighed the potential harms, which the APA notes can be mitigated. "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." p. 80 "The potential benefits of this guideline statement were viewed as far outweighing the potential harms. Although harms of antipsychotic medications can be significant, the impact of schizophrenia on patients' lives is also substantial, and consistent benefits of antipsychotic treatment were found. Harms of treatment can be mitigated by selecting medications on the basis of individual characteristics and preferences of patients as well as by choosing a medication on the basis of its side-effect profile, pharmacological characteristics, and other factors. For clozapine, the additional benefits of treatment were viewed as outweighing the additional rare but serious harms and the need for ANC [absolute neutrophil count] monitoring to reduce the likelihood of severe neutropenia." p. 81 Leucht, S., Leucht, C., Huhn, M., et al. (2017). Sixty years of placebo-controlled antipsychotic drug trials in acute schizophrenia: systematic review, Bayesian meta-analysis, and meta-regression of efficacy predictors. *Am J Psychiatry 174*(10):927–942.

Systematic Review	Evidence
	McDonagh, M. S., Dana, T., Selph, S., et al. (2017). Treatments for adults with schizophrenia: a systematic review [Comparative Effectiveness Review No 198, AHRQ Publ No 17(18)-EHC031-EF]. Rockville, MD: Agency for Healthcare Research and Quality. Available at https://effectivehealthcare.ahrq.gov/topics/schizophrenia-adult/research-2017. Accessed September 18, 2020.
What harms were identified?	From the 2019 APA schizophrenia guideline: "The harms of using an antipsychotic medication in the treatment of schizophrenia include sedation, side effects mediated through dopamine receptor blockade …, 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." p. 81
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	None.

2016 submission:

Systematic Review	Evidence
Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	American Psychiatric Association. (2002). Practice guideline for the treatment of patients with bipolar disorder, second edition. Retrieved from http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/bipolar.pdf American Psychiatric Association. (2010a). Practice guideline for the treatment of patients with major depressive disorder, 3rd ed. Retrieved from http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf American Psychiatric Association. (2010b). Practice guideline for the treatment of patients with schizophrenia: 2nd ed. Retrieved from http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/schizophrenia.pdf US Department of Veterans Affairs, & US Department of Defense. (2016). Management of major depressive disorder (MDD). Retrieved from http://www.healthquality.va.gov/guidelines/MH/mdd/VADoDMDDCPGFINAL82916.pdf US Department of Veterans Affairs & US Department of Defense. (2010) VA/DOD clinical practice guideline for management of bipolar disorder in adults. Retrieved from http://www.healthquality.va.gov/guidelines/MH/bd/bd_305_full.pdf

Systematic Review	Evidence
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	**Bipolar Disorder** APA 2002 Guidelines *Acute Phase* "The first-line pharmacological treatment for more severe manic or mixed episodes is the initiation of either lithium plus an antipsychotic or valproate plus an antipsychotic [I]. For less ill patients, monotherapy with lithium, valproate, or an antipsychotic such as olanzapine may be sufficient [I]. Short-term adjunctive treatment with a benzodiazepine may also be helpful [II]. For mixed episodes, valproate may be preferred over lithium [II]. Atypical antipsychotics are preferred over typical antipsychotics because of their more benign side effect profile [I], with most of the evidence supporting the use of olanzapine or risperidone [II]. Alternatives include carbamazepine or oxcarbazepine in lieu of lithium or valproate [II]. Antidepressants should be tapered and discontinued if possible [I]. If psychosocial therapy approaches are used, they should be combined with pharmacotherapy [I]." p.9 "Manic or mixed episodes with psychotic features usually require treatment with an antipsychotic medication [II]." p.10 *Maintenance Treatment* "Maintenance regimens of medication are recommended following a manic episode [I]. Although few studies involving patients with bipolar II disorder have been conducted, consideration of maintenance treatment for this form of the illness is also strongly warranted [II]. The medications with the best empirical evidence to support their use in maintenance treatment include lithium [I] and valproate [I]; possible alternatives include lamotrigine [II] or carbamazepine or oxcarbazepine [II]. If one of these medications was used to achieve remission from the most recent depressive or manic episode, it generally should be continued [I]." p.11 VA/DOD 2010 Guidelines "Patients with severe mania should be treated with a combination of antipsychotics and lithium or valproate. These antipsychotics include olanzapine, quetiapine, aripiprazole, or risperidone [B] and may include and ziprasidone [I]." p.8 "Patients with severe mixed episode should be treated with a combination of antipsychotics and lithium or valproate. These antipsychotics include aripiprazole, olanzapine, risperidone, or haloperidol [B] and may include quetiapine or ziprasidone [I]." p.9 "Clozapine, with its more serious side effect profile, may be added to existing medications for severe mania or mixed episode if it has been successful in the past or if other antipsychotics have failed [I]." p.9 "Quetiapine, [A], lamotrigine [B], or lithium [B] monotherapy should be considered as first-line treatment for adult patients with BD depression." p.26 *Maintenance Phase* "Patients who have had an acute manic episode should be treated for at least 6 months after the initial episode is controlled and encouraged to continue on life-long prophylactic treatment with medication. [A]" p.35

Systematic Review	Evidence
	Major Depressive Disorder APA 2010a Guidelines *Acute Phase* "An antidepressant medication is recommended as an initial treatment choice for patients with mild to moderate major depressive disorder [I] and definitely should be provided for those with severe major depressive disorder unless ECT is planned [I]." p.17 "For most patients, a selective serotonin reuptake inhibitor (SSRI), serotonin norepinephrine reuptake inhibitor (SNRI), mirtazapine, or bupropion is optimal [I]." p.17 *Maintenance Treatment* "To reduce the risk of relapse, patients who have been treated successfully with antidepressant medications in the acute phase should continue treatment with these agents for 4–9 months [I]." p.19 VA/DOD 2016 Guidelines "In patients with MDD who achieve remission with antidepressant medication, treatment should be continued at the same dose for at least 6 months to decrease the risk of relapse. [A]" p.106 **Schizophrenia** APA 2010b Guidelines *Acute Phase Treatment* "It is recommended that pharmacological treatment be initiated promptly, provided it will not interfere with diagnostic assessment, because acute psychotic exacerbations are associated with emotional distress, disruption to the patient's life, and a substantial risk of dangerous behaviors to self, others, or property [I]…The selection of an antipsychotic medication is frequently guided by the patient's previous experience with antipsychotics, including the degree of symptom response, past experience of side effects, and preferred route of medication administration. In choosing among these medications, the psychiatrist may consider the patient's past responses to treatment, the medication's side effect profile (including subjective responses, such as a dysphoric response to a medication), the patient's preferences for a particular medication based on past experience, the intended route of administration, the presence of co-morbid medical conditions, and potential interactions with other prescribed medications [I]. Finally, while most patients prefer oral medication, patients with recurrent relapses related to nonadherence are candidates for a long-acting injectable antipsychotic medication, as are patients who prefer this mode of administration [II]." p.11 *Stabilization Phase* "If the patient has improved with a particular medication regimen, continuation of that regimen and monitoring are recommended for at least 6 months [I]. Premature lowering of dose or discontinuation of medication during this phase may lead to a recurrence of symptoms and possible relapse." p.12

Systematic Review	Evidence
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The guideline authors did not grade the evidence or separate the grade for the evidence from the grade from the recommendation.
Provide all other grades and definitions from the evidence grading system	Not applicable
Grade assigned to the recommendation with definition of the grade	**Bipolar Disorder** Guidelines from the APA (2002) on the various treatment approaches related to initiating and continuing the medications in the numerator of this measure following an acute episode of bipolar disorder were graded as either I (recommended with substantial clinical confidence) or II (recommended with moderate clinical confidence). The recommendations for pharmacotherapy in the acute phase and maintenance regimens of medication after a manic episode were both graded as I. Guidelines from the VA/DoD (2010) on the various treatment approaches following an acute episode of bipolar disorder were graded as: B: At least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harm. I: Evidence that the intervention is effective is lacking, or poor quality, or conflicting, and the balance of benefits and harms cannot be determined. Guidelines from the VA/DoD (2010) on the continuation of medications in the numerator of this measure were graded as: A: Good evidence was found that the intervention improves important health outcomes and concludes that benefits substantially outweigh harm. **Major Depressive Disorder** Guidelines from the American Psychiatric Association (APA; 2010a) to initiate and continue the medications in the numerator of this measure following an acute episode of MDD were graded as: I: Recommended with substantial clinical confidence Guidelines from the Department of Veterans Affairs/Department of Defense (VA/DoD; 2016) to continue the medications in the numerator of this measure for at least six months following an acute episode of MDD were graded as: A: Good evidence was found that the intervention improves important health outcomes and concludes that benefits substantially outweigh harm. **Schizophrenia** Guidelines from the APA (2010b) to initiate and continue the medications in the

Systematic Review	Evidence
	numerator of this measure following an acute episode of schizophrenia were graded as: I: Recommended with substantial clinical confidence The guideline from the APA (2010b) to use long-acting injectables for patients hospitalized for schizophrenia was graded as follows: II: Recommended with moderate clinical confidence
Provide all other grades and definitions from the recommendation grading system	APA grade I: Recommended with substantial clinical confidence APA grade II: Recommended with moderate clinical confidence APA grade III: May be recommended on the basis of individual circumstances VA/DoD grade A: Good evidence was found that the intervention improves important health outcomes and concludes that benefits substantially outweigh harm. VA/DoD grade B: At least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harm. VA/DoD grade I: Evidence that the intervention is effective is lacking, or poor quality, or conflicting, and the balance of benefits and harms cannot be determined.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The guidelines are evidence-based rather than expert opinion. Information regarding the quantity, quality, and consistency of the information on the treatment of MDD, bipolar disorder, and schizophrenia is based on extensive literature searches reviewed by expert workgroups and panels, which included practicing clinicians and research experts. The current APA clinical guidelines for the treatment of bipolar disorder were built upon a literature search of articles from 1992 to 2001. A total of 472 citations are included in the current guideline (APA, 2002). The VA/DoD clinical guidelines relied heavily on the APA guidelines and include 276 citations (VA/DoD, 2010). For the treatment of MDD, the current APA guidelines were built upon literature reviews from pervious guidelines with the objective of emphasizing newer treatments. The literature search was conducted on studies published from January 1999 to December 2006. A total of 1,170 citations are reported in the current guideline (APA, 2010a). In a similar manner, the VA/DoD searched literature published from July 2000 to the end of 2006. A total of 253 citations are included in the current guideline (VA/DoD, 2010). The APA clinical guidelines for the treatment of schizophrenia were developed from a literature search conducted for the years 1994 to 2002. A total of 1,391 citations were included in the current guideline (APA, 2010b).
Estimates of benefit and consistency across studies	**Bipolar Disorder** Overall, the literature cited by the guidelines consistently found that pharmacotherapy is effective for the treatment of bipolar disorder. Many studies have demonstrated the efficacy of mood stabilizers (including lithium, anticonvulsants, and typical and atypical antipsychotics) as a treatment for reducing the depressive symptoms and manic episodes associated with bipolar disorder. Five studies found lithium to be a superior treatment for bipolar disorder compared to placebo (Bowden, et al., 1994; Goodwin, Murphy, & Bunney, 1969; Schou, Juel-Nielson, Stroomgreen, & Voldky, 1954; Maggs, 1963; Strokes, Shamoian, Stoll, & Patton, 1971). It should be noted the interpretation of these results is limited due to the use of a cross-over design in four of the trials (Goodwin et al., 1969; Schou et al., 1954; Maggs, 1963; Strokes et al., 1971), non-random

Systematic Review	Evidence
	assignment (Goodwin et al., 1969; Strokes et al., 1971), and variability in diagnostic criteria. In trials comparing lithium to other active pharmacological agents, lithium displayed similar efficacy to carbamazepine (Lerer, Moore, Meyendorff, Cho, & Gershon, 1987; Small et al., 1991), risperidone (Segal, Berk, & Brook, 1998), olanzapine (Berk, Ichim, & Brook, 1999), chlorpromazine, and other typical antipsychotics (Johnson, Gershon, Burdock, Floyd, & Hekimian, 1971; Platman, 1970; Prien, Caffey, & Klett, 1972; Shopsin, Gershon, Thompson, & Collins, 1975; Spring, Schweid, Gray, Steinberg, & Horwitz, 1970; Takahashi, Sakuma, Itoh, K., Itoh, H., & Kurihara, 1975). Open studies (Himmelhoch & Garfinkel, 1986; Kramlinger & Post, 1989; Prien, Himmelhoch, & Kuper, 1988) and randomized active comparator-controlled studies (Bowden, 1995; Freeman, Clothier, Pazzaglia, Lesem, & Swann, 1992; Swann et al., 1997) demonstrate that lithium is an effective treatment for manic states but is less effective in the treatment of mixed states. The efficacy of anticonvulsants (e.g., divalproex, valproate, valproic acid) compared to placebo has been demonstrated in four randomized controlled trials (Bowden, et al., 1994; Brennan, Sandyk, & Borsook, 1984; Emrich, Zerssen, Kissling, Miller, & Windorder, 1981; Pope, McElroy, Keck, & Hudson, 1991) with response rates ranging from 48% to 58%. One randomized, placebo-controlled study has evaluated antipsychotics for the treatment of bipolar disorder. The results indicated that chlorpromazine was superior to placebo in the overall improvement of manic symptoms (Klein, 1967). Typical antipsychotics are comparable to lithium in effectiveness (Platman, 1970; Prien, et al., 1972; Shopsin, et al., 1975; Spring et al., 1970; Takahashi, 1975). Atypical antipsychotics (i.e., risperidone and ziprasidone) have been shown to be superior to placebo and similar to haloperidol in effectiveness (Sachs, 2001). All of the pharmacotherapies evaluated in these studies are included in the numerator definition of this measure to allow for flexibility in prescribing an evidence-based treatment for bipolar disorder. **Major Depressive Disorder** Overall, the literature cited by the guidelines consistently found that pharmacotherapy is effective for the treatment of MDD. Several pharmacotherapies were reviewed through multiple meta-analyses (Anderson, 2000; Cipriani et al., 2005; Cipriani et al., 2009; Edwards & Anderson, 1999; Gartlehner, 2008), systematic reviews (Murdoch & Keam, 2005; Panzer, 2005), and numerous randomized trials that evaluated the efficacy and tolerability of pharmacological treatments for depression. Overall, the results of these studies indicate that SSRIs and SNRIs have relatively similar efficacies and tolerability. There is some evidence that tricyclic antidepressants (TCAs) may be more efficient for inpatient populations. SNRIs have been shown to be superior to placebo in multiple placebo-controlled studies (DeMartinis, Yeung, Entsuah, & Manley, 2007; Nemeroff, Entsuah, Benattia, Demitrack, Sloan, & Thase, 2008; Papakostas, Thase, Fava, Nelson, & Shelton, 2007; Papakostas, Homberger, & Fava, 2008; Septien-Velez, Pitrosky, Padmanabhan, Germain, & Tourian, 2007; Thase, Prtichette, Ossanna, Swindle, Xu, & Detke, 2007; Papakostas, Thase, Fava, Nelson, & Shelt, 2007). Several meta-analyses of controlled trials have documented small (4% – 10%) differences in treatment response for SNRIs compared to SSRIs (Cipriani, Barbui, Brambilla, Furukawa, Hotoph, & Geddes,

Systematic Review	Evidence
	2006; Nemeroff et al., 2008; Papakostas et al., 2008; Smith, 2002; Thase, 2001; Thase et al., 2007). Alternative depression medications have been efficacious in reducing depressive symptoms compared to placebo, including bupropion (Fava, Rush, Thase, Clayton, Stahl, Pradko, & Johnston, 2005) and mirtazapine (Claghorn & Lesem, 1995; Holm & Markham, 1999). Monoamine oxidase inhibitors (MAOIs) have similar efficacy to TCAs (Clayton, McGarvey, Abouesh, & Pinkerton, 2001; Himmelhock, Thase, Mallinger, & Houck, 1991; Masand, Ashton, Gupta, & Frank, 2001; McGrath, Stewart, Harrison, Wager, & Quitkin, 1986; White, Razani, Cadow, Gelfand, Palmer, Simpson, & Sloan, 1984), particularly for patients who have not responded to other antidepressant medication (Himmelhoch, Fuchs, & Symons, 1982; Himmelhoch et al., 1991; White et al., 1984). All of the classes of pharmacotherapies evaluated in these studies are included in the numerator definition of this measure to allow for flexibility in prescribing an evidence-based treatment for MDD. **Schizophrenia** Overall, the literature cited by the guidelines consistently found that pharmacotherapy is effective for the treatment of schizophrenia. According to the APA guidelines for the treatment of schizophrenia (APA, 2010b), evidence supporting the use of typical (i.e., first-generation) antipsychotics was first established in the 1960s (Laskey, Klett, Caffey, Bennett, Rosenblum, & Hollister, 1962) and repeatedly confirmed by subsequent clinical trials (Davis, Barter, & Kane, 1989). These studies compared the efficacy of one or more antipsychotic medications to that of a sedative or a placebo, and nearly all confirmed the antipsychotic medication to be a superior treatment (APA, 2010b). Research on typical antipsychotics has decreased substantially since the development of atypical (i.e., second-generation) antipsychotics. There are a number of atypical antipsychotics that are effective in the treatment of schizophrenia. At the time of the development of the clinical guidelines, clozapine was considered a superior treatment compared to typical antipsychotics in six of eight published double-blind randomized trials (Buchanan, Brier, Kirkpatrick, Ball, & Carpenter, 1998; Essock, Hargreaves, Covell, & Goethe, 1996; Hong, Chen, Chiu, & Sim, 1997; Kane, Honigfeld, Singer, & Meltzer, 1988; Kane et al., 2001; Kumra et al., 1996; Rosenheck et al., 1997; Volavka et al., 2002). A subsequent meta-analysis of five of these studies confirmed that clozapine-treated patients were 2.5 times more likely to improve compared to those treated with a typical antipsychotic. Clinical trials that informed the clinical guidelines demonstrated other atypical antipsychotics to be superior to placebo and to typical antipsychotics, including risperdone (Borison, Pathiraja, Diamond & Meibach, 1992; Chouinard et al., 1993; Marder & Meibach, 1994) and olanzapine (Beasley, Sanger, Satterless, Tollefson, Tran, & Hamilton, 1996; Beasley et al., 1997; Hamilton, Revicki, Genduso, & Beasley, 1998; Lieberman et al., 2003; Tollefson et al., 1997). Quetiapine and aripiprazole were demonstrated to be superior to placebo and typical antipsychotics (Borison, Arvanitis, & Milier, 1996; Fabre, Arvanitis, Pultz, Jones, Malick & Slotnick, 1995; Marder et al., 2003; Small, Kirsch, Arvanitis, Miller, & Link, 1997), although their effectiveness at reducing negative symptoms of schizophrenia is variable (Borison et al., 1996; Fabre et al., 1995; Small et al., 1997; Marder et al., 2003). Meta-analyses of these studies suggest that the efficacy of quetiapine is similar to that of typical antipsychotics (Geddes, Freemantle, Harrison, & Bebbington, 2000; Leucht,

Systematic Review	Evidence
	Pitschel-Walz, Abraham, & Kissling, 1999; Leucht, Wahlbeck, Hamann, & Kissling, 2003). Studies of ziprasidone found that it is superior compared to placebo and typical antipsychotics (Daniel, Zimbroff, Potkin, Reeves, Harrigan, & Lakshminarayanan, 1999; Keck, Buffenstein, Ferguson, Feighner, Jaffe, Harrigan, & Morrissey, 1998), including significantly reducing the risk of relapse (Goff et al., 1998). All of the pharmacotherapies evaluated in these studies are included in the numerator definition of this measure to allow for flexibility in prescribing an evidence-based treatment for schizophrenia. *American Psychiatric Association. (2002). Practice guideline for the treatment of patients with bipolar disorder, second edition. Retrieved from http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/bipolar.pdf *American Psychiatric Association. (2010a). Practice guideline for the treatment of patients with major depressive disorder, 3rd ed. Retrieved from http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf *American Psychiatric Association. (2010b). Practice guideline for the treatment of patients with schizophrenia: 2nd ed. Retrieved from http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/schizophrenia.pdf *Anderson, I. M. (2000). Selective serotonin reuptake inhibitors versus tricyclic antidepressants: A meta-analysis of efficacy and tolerability. *Journal of Affective Disorders*, *58*(1), 19–36. doi:10.1016/s0165-0327(99)00092-0 *Beasley, C. M., Sanger, T., Satterlee, W., Tollefson, G., Tran, P., & Hamilton, S. (1996). Olanzapine versus placebo: Results of a double-blind, fixed-dose olanzapine trial. *Psychopharmacology, 124*(1-2), 159-167. doi:10.1007/bf02245617 *Beasley, C. M., Hamilton, S. H., Crawford, A. M., Dellva, M. A., Tollefson, G. D., Tran, P. V., … Beuzen, J.-N. (1997). Olanzapine versus haloperidol: Acute phase results of the international double-blind olanzapine trial. *European Neuropsychopharmacology, 7*(2), 125-137. doi:10.1016/s0924-977x(96)00392-6 *Berk, M., Ichim, L., & Brook, S. (1999). Olanzapine compared to lithium in mania: A double-blind randomized controlled trial. *International Clinical Psychopharmacology, 14*(6), 339-343. doi:10.1097/00004850-199911000-00003 *Bowden, C. L., Brugger, A. M., Swann, A. C., Calabrese, J. R., Janicak, P. G., Petty, F.…Small, J. G. (1994). Efficacy of divalproex vs lithium and placebo in the treatment of mania: The Depakote Mania Study Group. *JAMA: The Journal of the American Medical Association, 271*(12), 918-924. doi:10.1001/jama.271.12.918 *Bowden, C. L. (1995). Predictors of response to divalproex and lithium. *Journal of Clinical Psychiatry, 56*(3), 25-30.

Systematic Review	Evidence
	*Borison, R. L., Pathiraja, A. P., Diamond, B. I., & Meibach, R. C. (1992). Risperidone: Clinical safety and efficacy in schizophrenia. *Psychopharmacological Bulletin, 28*, 213-218. *Borison, R. L., Arvanitis, L. A., & Milier, B. G. (1996). ICI 204,636, an atypical antipsychotic. *Journal of Clinical Psychopharmacology, 16*(2), 158-169. doi:10.1097/00004714-199604000-00008 *Bowden, C. L., Brugger, A. M., Swann, A. C., Calabrese, J. R., Janicak, P. G., Petty, F., Dilsaver, S. C.…Small, J. G. (1994). Efficacy of divalproex vs lithium and placebo in the treatment of mania. The Depakote Mania Study Group. *JAMA: The Journal of the American Medical Association, 271*(12), 918-924. doi:10.1001/jama.271.12.918 *Brennan, M. J. W., Sandyk, R., & Borsook, D. (1984). Use of sodium valproate in the management of affective disorders: Basic and clinical aspects. In Emrich, H. M., Okuma, T., & Muller, A. A. (Eds.), *Anticonvulsants in Affective Disorders* (pp. 56-65). Amsterdam: Excerpta Medica. *Buchanan, R. W., Breier, A., Kirkpatrick, B., Ball, P., & Carpenter Jr., W. T. (1998). Positive and negative symptom response to clozapine in schizophrenic patients with and without the deficit syndrome. *American Journal of Psychiatry, 155*, 751-760. *Cipriani, A., Brambilla, P., Furukawa, T. A., Geddes, J., Gregis, M., Hotopf, M., … Barbui, C. (2005). Fluoxetine versus other types of pharmacotherapy for depression. *Reviews.* doi:10.1002/14651858.cd004185.pub2 *Cipriani, A., Barbui, C., Brambilla, P., Furukawa, T. A., Hotopf, M., & Geddes, J. R. (2006). Are all antidepressants really the same? The case of Fluoxetine. *The Journal of Clinical Psychiatry, 67*(06), 850–864. doi:10.4088/jcp.v67n0601 *Cipriani, A., Furukawa, T. A., Salanti, G., Geddes, J. R., Higgins, J. P., Churchill, R., … Barbui, C. (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: A multiple-treatments meta-analysis. *The Lancet, 373*(9665), 746–758. doi:10.1016/s0140-6736(09)60046-5 *Chouinard, G., Jones, B., Remington, G., Bloom, D., Addington, D., MacEwan, G. W., … Arnott, W. (1993). A Canadian multicenter placebo-controlled study of fixed doses of risperidone and haloperidol in the treatment of chronic schizophrenic patients. *Journal of Clinical Psychopharmacology, 13*(1), 25???40. doi:10.1097/00004714-199302000-00004 *Claghorn, J. L., & Lesem, M. D. (1995). A double-blind placebo-controlled study of Org 3770 in depressed outpatients. *Journal of Affective Disorders, 34*(3), 165-171. doi:10.1016/0165-0327(95)00014-e *Clayton, A. H., McGarvey, E. L., Abouesh, A. I., & Pinkerton, R. C. (2001). Substitution of an SSRI with bupropion sustained release following SSRI-induced sexual dysfunction. *The Journal of Clinical Psychiatry, 62*(3), 185-190. doi:10.4088/jcp.v62n0309 *Daniel, D. G., Zimbroff, D. L., Potkin, S. G., Reeves, K. R., Harrigan, E. P., Lakshminarayanan, M. (1999). Ziprasidone 80 mg/day and 160 mg/day in the acute exacerbation of schizophrenia and schizoaffective disorder: A 6-week placebo-

Systematic Review	Evidence
	controlled trial. *Neuropsychopharmacology, 20*(5), 491-505. doi:10.1016/s0893-133x(98)00090-6 *Davis, J. M., Barter, J. T., Kane, J. M. (1989). Antipsychotic drugs. In H. I. Kaplan & B. J. Sadock (Eds.), *Comprehensive Textbook of Psychiatry, 5th ed* (pp. 1591-1626). Baltimore, MD. Williams & Wilkins. *DeMartinis, N. A., Yeung, P. P., Entsuah, R., & Manley, A. L. (2007). A double-blind, placebo-controlled study of the efficacy and safety of desvenlafaxine succinate in the treatment of major depressive disorder. *The Journal of Clinical Psychiatry, 68*(05), 677-688. doi:10.4088/jcp.v68n0504 *Edwards, J. G., & Anderson, I. (1999). Systematic review and guide to selection of selective serotonin reuptake inhibitors. *Drugs, 57*(4), 507-533. doi:10.2165/00003495-199957040-00005 *Emrich, H. M., Zerssen, D., Kissling, W., Miller, H.-J., & Windorfer, A. (1980). Effect of sodium valproate on mania. *Archive for Psychiatrie and Nervenkrankheiten, 229*(1), 1-16. doi:10.1007/bf00343800 *Emrich, H. M., Zihl, J., Raptis, C., & Wendl, A. (1990). Reduced dark-adaptation: An indication of lithium's neuronal action in humans. *American Journal of Psychiatry, 147*(5), 629-631. doi: 10.1176/ajp.147.5.629 *Essock, S. M., Hargreaves, W. A. Covell, N. H., & Goethe, J. (1996). Clozapine's effectiveness for patients in state hospitals: Results from a randomized trial. *Psychopharmacological Bulletin, 32*, 683-697. *Fabre, L. F., Arvanitis, L., Pultz, J., Jones, V. M., Malick, J. B., & Slotnick, V. B. (1995). ICI 204,636, a novel, atypical antipsychotic: Early indication of safety and efficacy in patients with chronic and subchronic schizophrenia. *Clinical Therapeutics, 17*(3), 366-378. doi:10.1016/0149-2918(95)80102-2 *Fava, M., Rush, A. J., Thase, M. E., Clayton, A., Stahl, S. M., Pradko, J. F., & Johnston, J. A. (2005). 15 years of clinical experience with bupropion HCl. *The Primary Care Companion to The Journal of Clinical Psychiatry, 07*(03), 106-113. doi:10.4088/pcc.v07n0305 *Freeman, T. W., Clothier, J. L., Pazzaglia, P., Lesem, M. D., & Swann, A. C. (1992). A double-blind comparison of valproate and lithium in the treatment of acute mania. *American Journal of Psychiatry, 149*(1), 108–111. doi:10.1176/ajp.149.1.108 *Gartlehner, G. (2008). Comparative benefits and harms of second-generation antidepressants: Background paper for the American College of Physicians. *Ann Intern Med, 149*(10), 734. doi:10.7326/0003-4819-149-10-200811180-00008 *Geddes, J., Freemantle, N., Harrison, P., & Bebbington, P. (2000). Atypical antipsychotics in the treatment of schizophrenia: Systematic overview and meta-regression analysis. *British Medical Journal, 321*(7273), 1371-1376. doi:10.1136/bmj.321.7273.1371 *Goff, D. C., Posever, T., Herz, L., Simmons, J., Kletti, N., Lapierre, K., ... Ko, G. N. (1998). An exploratory haloperidol-controlled dose-finding study of ziprasidone in hospitalized

Systematic Review	Evidence
	patients with schizophrenia or schizoaffective disorder. *Journal of Clinical Psychopharmacology, 18*(4), 296-304. doi:10.1097/00004714-199808000-00009 *Goodwin, F. K. (1969). Lithium-carbonate treatment in depression and mania. *Archives of General Psychiatry, 21*(4), 486. doi:10.1001/archpsyc.1969.01740220102012 *Hamilton, S. H., Revicki, D. A., Genduso, L. A., Beasley, C. M. (1998). Olanzapine versus placebo and haloperidol: Quality of life and efficacy results of the North American Double-blind Trial. *Neuropsychopharmacology, 18*(1), 41-49. doi:10.1016/s0893-133x(97)00111-5 *Himmelhoch, J. M., Fuchs, C. Z., & Symons, B. J. (1982). A double-blind study of tranylcypromine treatment of major anergic depression. *The Journal of Nervous and Mental Disease, 170*(10), 628-634. doi:10.1097/00005053-198210000-00007 *Himmelhoch, J. M. & Garfinkel, M. E. (1986). Sources of lithium resistance in mixed mania. *Psychopharmacological Bulletin, 22*, 613-620. *Himmelhoch, J. M., Thase, M. E., Mallinger, A. G., & Houck, P. (1991). Tranylcypromine versus imipramine in anergic bipolar depression. *American Journal of Psychiatry, 148*(7), 910-916. doi:10.1176/ajp.148.7.910 *Holm, K. J., & Markham, A. (1999). Mirtazapine. *Drugs, 57*(4), 607–631. doi:10.2165/00003495-199957040-00010 *Hong, C. J., Chen, J. Y., Chiu, H. J., & Sim, C. B. (1997). A double-blind comparative study of clozapine versus chlorpromazine on Chinese patients with treatment-refractory schizophrenia. *International Clinical Psychopharmacology, 12*(3), 123-130. doi:10.1097/00004850-199705000-00001 *Johnson, G., Gershon, S., Burdock, E. I., Floyd, A., & Hekimian, L. (1971). Comparative effects of lithium and chlorpromazine in the treatment of acute manic states. *The British Journal of Psychiatry, 119*(550), 267–276. doi:10.1192/bjp.119.550.267 *Kane, J., Honigfeld, G., Singer, J., & Meltzer, H. (1988). Clozapine for the treatment-resistant schizophrenic. *Archives of General Psychiatry, 45*(9), 789. doi:10.1001/archpsyc.1988.01800330013001 *Kane, J. M., Marder, S. R., Schooler, N. R., Wirshing, W. C., Umbricht, D., Baker, R. W., … Borenstein, M. (2001). Clozapine and haloperidol in moderately refractory schizophrenia. *Archives of General Psychiatry, 58*(10), 965. doi:10.1001/archpsyc.58.10.965 *Keck Jr, P., Buffenstein, A., Ferguson, J., Feighner, J., Jaffe, W., Harrigan, E. P., & Morrissey, M. R. (1998). Ziprasidone 40 and 120 mg/day in the acute exacerbation of schizophrenia and schizoaffective disorder: A 4-week placebo-controlled trial. *Psychopharmacology, 140*(2), 173–184. doi:10.1007/s002130050755 *Klein, D. F. (1967). Importance of psychiatric diagnosis in prediction of clinical drug effects. *Archives of General Psychiatry, 16*(1), 118. doi:10.1001/archpsyc.1967.01730190120016

Systematic Review	Evidence
	*Kramlinger, K. G., & Post, R. M. (1989). Adding lithium carbonate to carbamazepine: Antimanic efficacy in treatment-resistant mania. *Acta Psychiatrica Scandinavica, 79*(4), 378–385. doi:10.1111/j.1600-0447.1989.tb10273.x *Kumra, S., Frazier, J. A., Jacobsen, L. K., McKenna, K., Gordon, C. T., Lenane, M. C.,…Rapoport, J. L. (1996). Childhood-onset schizophrenia. *Archives of General Psychiatry, 53*(12), 1090. doi:10.1001/archpsyc.1996.01830120020005 *Laskey, J. J., Klett, C. J., Caffey, E. M. Jr., Bennett, J. L., Rosenblum, M. P., and Hollister, L. E. (1962). Drug treatment of schizophrenic patients: A comprehensive evaluation of chlorpromazine, chlorprothixene, fluphenazine, reserpine, thioridazine, and trifluprom azine. *Diseases of the Nervous System,* 23. 698-706. *Lerer, B., Moore, N., Meyendorff, E., Cho, S. R., & Gershon, S. (1987). Carbamazepine versus lithium in mania: A double-blind study. *Journal of Clinical Psychiatry, 48,* 89-93. *Leucht, S., Pitschel-Walz, G., Abraham, D., & Kissling, W. (1999). Efficacy and extrapyramidal side-effects of the new antipsychotics olanzapine, quetiapine, risperidone, and sertindole compared to conventional antipsychotics and placebo. A meta-analysis of randomized controlled trials. *Schizophrenia Research, 35*(1), 51-68. doi:10.1016/s0920-9964(98)00105-4 *Leucht, S., Wahlbeck, K., Hamann, J., & Kissling, W. (2003). New generation antipsychotics versus low-potency conventional antipsychotics: A systematic review and meta-analysis. *The Lancet, 361*(9369), 1581–1589. doi:10.1016/s0140-6736(03)13306-5 *Maggs, R. (1963). Treatment of manic illness with lithium carbonate. *The British Journal of Psychiatry, 109*(458), 56–65. doi:10.1192/bjp.109.458.56 *Marder, S. R. & Meibach, R. C. (1994). Risperidone in the treatment of schizophrenia. *American Journal of Psychiatry, 151*, 825-835. *Marder, S. R., McQuade, R. D., Stock, E., Kaplita, S., Marcus, R., Safferman, A. Z., … Iwamoto, T. (2003). Aripiprazole in the treatment of schizophrenia: Safety and tolerability in short-term, placebo-controlled trials. *Schizophrenia Research, 61*(2-3), 123–136. doi:10.1016/s0920-9964(03)00050-1 *Masand, P. S., Ashton, A. K., Gupta, S., & Frank, B. (2001). Sustained-release bupropion for selective serotonin reuptake inhibitor-induced sexual dysfunction: A randomized, double-blind, placebo-controlled, parallel-group study. *American Journal of Psychiatry, 158*(5), 805–807. doi:10.1176/appi.ajp.158.5.805 *McGrath, P. J., Stewart, J. W., Harrison, W., Wager, S., & Quitkin, F. M. (1986). Phenelzine treatment of melancholia. *Journal of Clinical Psychiatry, 47*, 420-422. *Muller, A. A. & Stoll, K. D. (1984). Anticonvulsants in affective disorders. In Emrich, H. M., Okuma, T., & Muller, A. A. (Eds.) *Carbamazepine and Oxcarbazepine in the treatment of manic syndromes: Studies in Germany* (pp. 134-147). Amsterdam, The Netherlands: Exerpta Medica.

Systematic Review	Evidence
	*Murdoch, D. & Keam, S. J. (2005). Escitalopram: A review of its use in the management of major depressive disorder. *Drugs, 65*,2379-2404. doi: 10.2165/00003495-200565160-00013 *Nemeroff, C. B., Entsuah, R., Benattia, I., Demitrack, M., Sloan, D. M., & Thase, M. E. (2008). Comprehensive analysis of remission (COMPARE) with Venlafaxine versus SSRIs. *Biological Psychiatry, 63*(4), 424-434. doi:10.1016/j.biopsych.2007.06.027 *Panzer, M. J. (2005). Are SSRIs really more effective for anxious depression? *Annals of Clinical Psychiatry, 17*(1), 23-29. doi:10.1080/10401230590905317 *Papakostas, G. I., Thase, M. E., Fava, M., Nelson, J. C., & Shelton, R. C. (2007). Are antidepressant drugs that combine serotonergic and noradrenergic mechanisms of action more effective than the selective serotonin reuptake inhibitors in treating major depressive disorder? A meta-analysis of studies of newer agents. *Biological Psychiatry, 62*(11), 1217-1227. doi:10.1016/j.biopsych.2007.03.027 *Papakostas, G., Homberger, C., & Fava, M. (2008). A meta-analysis of clinical trials comparing mirtazapine with selective serotonin reuptake inhibitors for the treatment of major depressive disorder. *Journal of Psychopharmacology, 22*(8), 843–848. doi:10.1177/0269881107083808 *Platman, S. R. (1970). A comparison of lithium carbonate and chlorpromazine in mania. *American Journal of Psychiatry, 127*(3), 351–353. doi:10.1176/ajp.127.3.351 *Pope, Jr., H. G., McElroy, S. L., Keck, Jr., P. E., & Hudson, J. I. (1991). Valproate in the treatment of acute mania: A placebo-controlled study. *Archives of General Psychiatry, 48*, 62-68. *Prien, R. F., Caffey Jr., E. M., Klett, C. J. (1972). Comparison of lithium carbonate and chlorpromazine in the treatment of mania: Report of the Veterans Administration and National Institute of Mental Health Collaborative Study Group. *Archives of General Psychiatry, 26*(2), 146. doi:10.1001/archpsyc.1972.01750200050011 *Prien, R. F., Himmelhoch, J. M., & Kupfer, D. J. (1988). Treatment of mixed mania. *Journal of Affective Disorders, 15*(1), 9–15. doi:10.1016/0165-0327(88)90003-1 *Rosenheck, R., Cramer, J., Xu, W., Thomas, J., Henderson, W., Frisman, L., … Charney, D. (1997). A comparison of clozapine and haloperidol in hospitalized patients with refractory schizophrenia. *New England Journal of Medicine, 337*(12), 809–815. doi:10.1056/nejm199709183371202 *Sachs, G. S. (2001). Emerging data: Atypical antipsychotics in bipolar disorder. In *Program and Abstracts of the 52nd Institute on Psychiatric Services*. Washington, D. C.: American Psychiatric Association. *Schou, M., Juel-Nielsen, N., Stromgren, E., & Voldby, H. (1954). The treatment of manic psychoses by the administration of lithium salts. *Journal of Neurology, Neurosurgery & Psychiatry, 17*(4), 250–260. doi:10.1136/jnnp.17.4.250

Systematic Review	Evidence
	*Shopsin, B., Gershon S., Thompson, H., & Collins, P. (1975). Psychoactive drugs in mania. *Archives of General Psychiatry, 32*(1), 34. doi:10.1001/archpsyc.1975.01760190036004 *Segal, J., Berk, M., & Brook, S. (1998). Risperidone compared with both lithium and haloperidol in mania: A double-blind randomized controlled trial. *Clinical Neuropharmacology, 21*, 176-180. *Septien-Velez, L., Pitrosky, B., Padmanabhan, S. K., Germain, J.-M., & Tourian, K. A. (2007). A randomized, double-blind, placebo-controlled trial of desvenlafaxine succinate in the treatment of major depressive disorder. *International Clinical Psychopharmacology, 22*(6), 338–347. doi:10.1097/yic.0b013e3281e2c84b *Small, J. G. (1991). Carbamazepine compared with lithium in the treatment of mania. *Archives of General Psychiatry, 48*(10), 915. doi:10.1001/archpsyc.1991.01810340047006 *Small, J. G., Hirsch, S. R., Arvanitis, L. A., Miller, B. G., & Link, C. G. (1997). Quetiapine in patients with schizophrenia. *Archives of General Psychiatry, 54*(6), 549. doi:10.1001/archpsyc.1997.01830180067009 *Smith, D. (2002). Efficacy and tolerability of venlafaxine compared with selective serotonin reuptake inhibitors and other antidepressants: A meta-analysis. *The British Journal of Psychiatry, 180*(5), 396–404. doi:10.1192/bjp.180.5.396 *Spring, G., Schweid, D., Gray, C., Steinberg, J., & Horwitz, M. (1970). A double-blind comparison of lithium and chlorpromazine in the treatment of manic states. *American Journal of Psychiatry, 126*(9), 1306–1310. doi:10.1176/ajp.126.9.1306 *Stokes, P., Shamoian, C., Stoll, P., & Patton, M. (1971). Efficacy of lithium as acute treatment of manic-depressive illness. *The Lancet, 297*(7713), 1319–1325. doi:10.1016/s0140-6736(71)91886-1 *Swann, A. C., Bowden, C. L., Morris, D., Calabrese, J. R., Petty, F., …Davis, J. M. (1997). Depression during mania: Treatment response to lithium or divalproex. *Archives of General Psychiatry, 54*(1), 37. doi:10.1001/archpsyc.1997.01830130041008 *Takahashi, R. (1975). Comparison of efficacy of lithium carbonate and chlorpromazine in mania. *Archives of General Psychiatry, 32*(10), 1310. doi:10.1001/archpsyc.1975.01760280108010 *Thase, M. E., Pritchett, Y. L., Ossanna, M. J., Swindle, R. W., Xu, J., & Detke, M. J. (2007). Efficacy of Duloxetine and selective serotonin reuptake inhibitors. *Journal of Clinical Psychopharmacology, 27*(6), 672–676. doi:10.1097/jcp.0b013e31815a4412 *Thase, M. E. (2001). Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *The British Journal of Psychiatry, 178*(3), 234–241. doi:10.1192/bjp.178.3.234 *Tollefson, G. D., Beasley Jr., C. M., Tran, P. V., Street, J. S., Krueger, J. A., Tamura, R. N., … Thieme, M. E. (1997). Olanzapine versus haloperidol in the treatment of schizophrenia and schizoaffective and schizophreniform disorders: Results of an

Systematic Review	Evidence
	international collaborative trial. *American Journal of Psychiatry*, *154*(4), 457–465. doi:10.1176/ajp.154.4.457 *Volavka, J., Czobor, P., Sheitman, B., Lindenmayer, J.-P., Citrome, L., McEvoy, J. P., … Lieberman, J. A. (2002). Clozapine, olanzapine, risperidone, and haloperidol in the treatment of patients with chronic schizophrenia and schizoaffective disorder. *American Journal of Psychiatry, 159*(2), 255–262. doi:10.1176/appi.ajp.159.2.255 *White, K., Razani, J., Cadow, B., Gelfand, R., Palmer, R., Simpson, G., & Sloane, R. B. (1984). Tranylcypromine vs nortriptyline vs placebo in depressed outpatients: A controlled trial. *Psychopharmacology, 82*(3), 258–262. doi:10.1007/bf00427786 *US Department of Veterans Affairs, & US Department of Defense. (2009). management of major depressive disorder (MDD). Retrieved from http://www.healthquality.va.gov/mdd/mdd_full09_c.pdf
What harms were identified?	Medications associated with the treatment of MDD, schizophrenia, and bipolar disorder have been shown to reduce negative symptoms, and the clinical guidelines indicate that the benefits outweigh harms for patients with severe mental illness. However, many of the medications require careful monitoring to avoid harmful side effects. Clinicians prescribing medications for the treatment of these disorders must consider the specific medication and the side effects that might occur. These considerations may vary given a patient's clinical and personal characteristics, as well as the expected improvement in the patient's outcomes. The implementation of this measure will provide the important benefit of quality improvement by helping to identify patients who do not continue their pharmacotherapy post-discharge. Improved medication continuation would help reduce the risk of symptom relapse, prevent future depressive/manic/psychotic episodes, decrease re-hospitalization and suicide rates, and improve the quality of care for individuals with major depressive disorder, schizophrenia, and bipolar disorder.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Since the development of the clinical guidelines for schizophrenia, the Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) Project compared the longer-term effects and usefulness of typical (perphenazine, fluphenazine decanoate) and atypical (olanzapine, quetiapine, risperidone, ziprasidone, clozapine) antipsychotics. A study based on data from that project found that perphenazine, a typical antipsychotic, was equally as effective as the atypical antipsychotics quetiapine, risperidone, and ziprasidone (Lieberman, et al., 2010). This finding further supports the inclusion of both types of antipsychotics in the numerator definition for schizophrenia in this measure. *Lieberman, J. A., Tollefson, G., Tohen, M., Green, A. I., Gur, R. E., Kahn, R., … Hamer, R. M. (2003). Comparative efficacy and safety of atypical and conventional antipsychotic drugs in first-episode psychosis: A randomized, double-blind trial of olanzapine versus haloperidol. *American Journal of Psychiatry*, *160*(8), 1396–1404. doi:10.1176/appi.ajp.160.8.1396

1a.4 OTHER SOURCE OF EVIDENCE

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

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

2020 submission: Not applicable

2016 submission: Not applicable

1a.4.2 What process was used to identify the evidence?

2020 submission: Not applicable

2016 submission: Not applicable

1a.4.3. Provide the citation(s) for the evidence.

2020 submission: Not applicable

2016 submission: Not applicable

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

The aim of the measure is to address gaps in continuity of pharmaceutical treatment during the transition from inpatient to outpatient care. Pharmacotherapy is the primary form of treatment for most patients discharged from an inpatient psychiatric facility (IPF) for bipolar disorder, major depressive disorder (MDD), or schizophrenia. The measure focuses on medication continuation because it is an essential step in medication adherence.

Medication continuation is particularly important in the psychiatric patient population because psychotropic medication discontinuation can have a range of adverse effects, from mild withdrawal to life-threatening autonomic instability and psychiatric decompensation (Ward & Schwartz, 2013). Patients with MDD who do not remain on prescribed medication are more likely to have negative health outcomes, such as relapse and readmission, decreased quality of life, and increased health care costs. If untreated, MDD can contribute to or worsen chronic medical disorders (Geddes et al., 2003; Glue et al., 2010). The literature shows that among patients with schizophrenia, those who were "good compliers" according to the Medication Adherence Rating Scale had better outcomes in terms of rehospitalization rates and medication maintenance (Jaeger et al., 2012). Among patients with bipolar disorder, medication adherence was significantly associated with reduction in manic symptoms (Sylvia et al., 2013), whereas nonadherence was associated with increased suicide risk (OR 10.8, CI 1.57–74.4; Gonzalez-Pinto et al., 2006). Our literature review from January 2016 through August 2020 did not reveal any new evidence regarding performance gaps since the initial endorsement submission.

Current facility-level performance indicates a clear quality gap. Using Medicare claims data from July 1, 2017, through June 20, 2019, the Medication Continuation measure rates ranged from 34.8 to 94.3%, with a median

of 76.2%. There was a 21.3 percentage point difference between the 10th and 90th percentiles (63.4–84.7%). Using 2013–2014 Medicare claims data, there was a 21.6 percentage point difference between the 10th and 90th percentiles (66.7–88.3%) and a median score of 79.6%. By calculating the facility-level rates of medication continuation in Medicare fee-for-service (FFS) claims data, this measure can provide valuable information on areas where care transitions to the outpatient setting can be improved.

Literature about continuation of medication has identified effective interventions that facilities can employ to improve medication adherence among patients discharged from an IPF (Douaihy, Kelly, & Sullivan, 2013; Haddad, Brain, & Scott, 2014; Hung, 2014; Kasckow & Zisook, 2008; Lanouette, Folsom, Sciolla, & Jeste, 2009; Mitchell, 2007; Sylvia et al., 2013). Examples of these interventions include patient education, shared decision making, and text-message reminders. We envision the addition of this measure to the suite of measures for IPFs would help to create a comprehensive picture of the quality of care patients receive at those facilities.

*Douaihy, A. B., Kelly, T. M., Sullivan, C. (2013). Medications for substance use disorders. Social Work in Public Health, 28(3-4), 264-278. doi: 10.1080/19371918.2013.759031

*Geddes, J. R., Carney, S. M., Davies, C., Furukawa, T. A., Kupfer, D. J., Frank, E., & Goodwin, G. M. (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. The Lancet, 361(9358), 653–661. doi:10.1016/s0140-6736(03)12599-8

*Glue, P., Donovan, M. R., Kolluri, S., & Emir, B. (2010). Meta-analysis of relapse prevention antidepressant trials in depressive disorders. Australian and New Zealand Journal of Psychiatry, 44(8), 697-705. doi: 10.3109/00048671003705441

*Gonzalez-Pinto, A., Mosquera, F., Alonso, M., López, P., Ramírez, F., Vieta, E., & Baldessarini, R. J. (2006). Suicidal risk in bipolar I disorder patients and adherence to long-term lithium treatment. Bipolar Disorders, 8(5p2), 618–624. doi:10.1111/j.1399-5618.2006.00368.x

*Haddad, P. M., Brain, C., & Scott, J. (2014). Nonadherence with antipsychotic medication in schizophrenia: Challenges and management strategies. Patient Related Outcome Measures, 5, 43-62. doi: 10.2147/PROM.S42735

*Hung, C. I. (2014). Factors predicting adherence to antidepressant treatment. Current Opinion in Psychiatry, 27(5), 344-349. doi: 10.1097/yco.0000000000000086

*Jaeger, S., Pfiffner, C., Weiser, P., Kilian, R., Becker, T., Langle, G., . . . Steinert, T. (2012). Adherence styles of schizophrenia patients identified by a latent class analysis of the Medication Adherence Rating Scale (MARS): A six-month follow-up study. Psychiatry Research, 200(2-3), 83-88. doi: 10.1016/j.psychres.2012.03.033

*Kasckow, J. W., & Zisook, S. (2008). Co-occurring depressive symptoms in the older patient with schizophrenia. Drugs & Aging, 25(8),631-647.

*Lanouette, N. M., Folsom, D. P., Sciolla, A., Jeste, D. V. (2009). Psychotropic medication nonadherence among United States Latinos: A comprehensive literature review. Psychiatric Services (Washington, DC), 60(2), 157-174. doi: 10.1176/appi.ps.60.2.157

*Mitchell, A. J. (2007). Understanding medication discontinuation in depression. BMedSci Psychiatric Times, 24(4).

*Sylvia, L. G., Hay, A., Ostacher, M. J., Miklowitz, D. J., Nierenberg, A. A., Thase, M. E., Perlis, R. H. (2013). Association between therapeutic alliance, care satisfaction, and pharmacological adherence in bipolar disorder. Journal of Clinical Psychopharmacology, 33(3), 343-350. doi: 10.1097/JCP.0b013e3182900c6f

*Ward, M., & Schwartz, A. (2013). Challenges in pharmacologic management of the hospitalized patient with psychiatric comorbidity. Journal of Hospital Medicine, 8(9), 523–529. doi:10.1002/jhm.2059.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, 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 (4b1) under Usability and Use.

We calculated the measure scores at the facility level using Medicare FFS Part A and Part B claims data from July 1, 2017, through June 30, 2019. The testing data set included 308,556 discharges from 182,042 patients across 1,680 IPFs. For the Inpatient Psychiatric Facility Quality Reporting (IPFQR) Program sponsored by the Centers for Medicare & Medicaid Services (CMS), the measure is calculated only for IPFs with at least 75 discharges eligible for the denominator. The testing data set included 1,066 IPFs and 268,673 discharges that fit this restriction.

The performance score statistics across all facilities in the data set follow, as well as for only those facilities with at least 75 discharges eligible for the denominator.

Medication continuation rate across all IPFs (n=1,680) in the data set:

Mean: 75.0%

Std dev: 12.8%

Min: 0.0%

Max: 100.0%

Interquartile range: 12.6%

Scores by decile:

10%: 61.7%

20%: 68.0%

30%: 71.4%

40%: 74.1%

50%: 76.8%

60%: 79.0%

70%: 81.4%

80%: 83.8%

90%: 87.5%

Medication continuation rate IPFs with at least 75 eligible cases in the denominator (n = 1,066):

Mean: 75.1%

Std dev: 8.3%

Min: 34.8%

Max: 94.3%

Interquartile range: 11.0%

Scores by decile:

10%: 63.4%

20%: 68.4%

30%: 70.2%

40%: 74.5%

50%: 76.2%

60%: 78.1%

70%: 80.0%

80%: 82.2%

90%: 84.7%

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, 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.

Not applicable

1b.4. 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. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) 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 (4b1) under Usability and Use.*

Number of patients in the data: 182,042

Dates of data: July 1, 2017, through June 30, 2019

With large sample sizes, small differences that are statistically significant might not always be practically or clinically meaningful. Therefore, we also computed Cohen's d effect size (the difference in mean scores divided by the pooled standard deviation). A d of 1 indicates the two groups differ by 1 standard deviation, a d of 2 indicates they differ by 2 standard deviations, and so on. Following Cohen's (1988) definitions, we defined effect size values for dichotomous variables as small (0.2), medium (0.5), or large (0.8). For patient subgroups with more than two categories (age and diagnosis), we computed Eta-squared (?2) effect size to capture the overall difference in the measure rate between groups. We categorized corresponding effect size values as small (0.01), medium (0.06), or large (0.14).

Medication continuation rate across all IPFs (n = 1,680):

Sex, male: 72.1%

Sex, female: 77.9%

Effect size (Cohen's d) for differences in means between patient groups: 0.39

Substance use disorder (SUD) diagnosis, diagnosed with SUD: 70.4%

SUD diagnosis, not diagnosed with SUD: 76.9%

Effect size (Cohen's d) for differences in means between patient groups: 0.41

Dual status, dual: 77.4%

Dual status, not dual: 69.8%

Effect size (Cohen's d) for differences in means between patient groups: 0.51

Race, non-White: 71.1%

Race, White: 76.2%

Effect size (Cohen's d) for differences in means between patient groups: 0.31

Diagnosis, schizophrenia: 75.5%

Diagnosis, major depressive disorder (MDD): 74.2%

Diagnosis, bipolar disorder: 75.3%

Effect size (Eta2) for differences in means between patient groups: 0.001

Age, 18–39: 74.0%

Age, 40–59: 74.1%

Age, 60 and older: 75.4%

Effect size (Eta2) for differences in means between patient groups: 0.004

Medication continuation rate across IPFs with at least 75 eligible cases in the denominator (n = 1,066):

Sex, male: 72.2%

Sex, female: 78.0%

Effect size (Cohen's d) for differences in means between patient groups: 0.64

SUD diagnosis, diagnosed with SUD: 69.7%

SUD diagnosis, not diagnosed with SUD: 77.4%

Effect size (Cohen's d) for differences in means between patient groups: 0.74

Dual status, dual: 77.6%

Dual status, not dual: 69.1%

Effect size (Cohen's d) for differences in means between patient groups: 0.85

Race, non-White: 71.2%

Race, White: 76.3%

Effect size (Cohen's d) for differences in means between patient groups: 0.46

Diagnosis, schizophrenia: 76.1%

Diagnosis, MDD: 73.2%

Diagnosis, bipolar disorder: 75.2%

Effect size (Eta2) for differences in means between patient groups: 0.013

Age, 18–39: 74.7%

Age, 40–59: 74.8%

Age, 60 and older: 74.9%

Effect size (Eta2) for differences in means between patient groups: 0.000

Note on interpretation of effect size:

Cohen's d: 0.2 is considered a small effect size, 0.5 is a medium effect size, and 0.8 is a large effect size

Eta2: 0.01 is small, 0.06 is medium, and 0.14 is large

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, 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 1b.4

Not applicable

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

De.6. Non-Condition Specific*(check all the areas that apply):*

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

S.1. Measure-specific Web Page *(Provide a 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.)*

Measure-specific webpage not available at the time of the annual update submission.

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **:** Med_Cont_Data_Dictionary_FY2021.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

We removed the following from the measure´s list of medications for treatment of bipolar disorder as they are not FDA-approved for treatment of bipolar disorder:

- Fluphenazine
- Molindone
- Perphenazine
- Pimozide
- Prochlorperazine
- Thioridazine
- Thiothixene
- Trifluoperazine
- Brexpiprazole
- Iloperidone
- Paliperidone
- Fluphenazine decanoate
- Paliperidone palmitate (1-month extended-release)
- Paliperidone palmitate (3-month extended-release)

This revision is harmonized with Adherence to Mood Stabilizers for Individuals with Bipolar I Disorder (NQF #1880), which does not include these medications as treatments for bipolar disorder.

We also removed paliperidone palmitate (3-month extended-release) from the measure´s list of medications for treatment of schizophrenia due to its questionable clinical appropriateness to be used as the sole therapy for patients recently discharged from IPFs. This medication requires that the patient to be adequately treated with the 1-month extended-release injection for at least 4 months. Therefore, most patients who are admitted to IPFs for acute management of symptoms are unlikely to be candidates for this medication.

We added the International Classification of Diseases-10 (ICD 10) codes F53.0 (postpartum depression) and F53.1 (puerperal psychosis) to the list of codes that define the denominator exclusions. This modification was made because the 2019 code set revised the description for F53 (from "puerperal psychosis" to "mental and behavioral disorders associated with the puerperium, not elsewhere classified"), changed it to the parent code, and added the new codes F53.0 and F53.1. F53 is still in the measure along with F53.0 and F53.1 because all three codes are relevant for the performance period for FY2021 of the Inpatient Psychiatric Facility Quality Reporting (IPFQR) program.

S.4. Numerator Statement *(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.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The numerator for the measure includes:

- Discharges with a principal diagnosis of MDD in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge
- Discharges with a principal diagnosis of schizophrenia in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge
- Discharges with a principal diagnosis of bipolar disorder in the denominator population for which patients were dispensed evidence-based outpatient medication within 2 days prior to discharge through 30 days post-discharge

S.5. Numerator Details *(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 S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The following are lists of evidence-based medications for the treatment of MDD, schizophrenia, and bipolar disorder:

Medications for MDD

- Monoamine Oxidase Inhibitors: isocarboxazid, phenelzine, selegiline (transdermal patch), tranylcypromine
- Selective Serotonin Reuptake Inhibitors (SSRI): citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline
- Serotonin Modulators: nefazodone, trazodone, vilazodone, vortioxetine

- Serotonin Norepinephrine Reuptake Inhibitors (SNRI): desvenlafaxine, duloxetine, levomilnacipran, venlafaxine
- Tricyclic and Tetracyclic Antidepressants: amitriptyline, amoxapine, clomipramine, desipramine, doxepin, imipramine, maprotiline, nortriptyline, protriptyline, trimipramine
- Other Antidepressants: bupropion, mirtazapine
- Psychotherapeutic Combinations: amitriptyline-chlordiazepoxide, amitriptyline-perphenazine, fluoxetine-olanzapine

Medications for Schizophrenia

- First-generation Antipsychotics: chlorpromazine, fluphenazine, haloperidol, haloperidol lactate, loxapine succinate, molindone, perphenazine, pimozide, prochlorperazine, thioridazine, thiothixene, trifluoperazine
- Second-generation (Atypical) Antipsychotics: aripiprazole, asenapine, brexpiprazole, cariprazine, clozapine, iloperidone, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, ziprasidone
- Psychotherapeutic Combinations: amitriptyline-perphenazine, fluoxetine-olanzapine
- Long-Acting (Depot) Injectable Antipsychotics: fluphenazine decanoate, haloperidol decanoate, aripiprazole, aripiprazole lauroxil, olanzapine pamoate, paliperidone palmitate (1-month extended-release injection, risperidone microspheres

Medications for Bipolar Disorder

- Anticonvulsants: carbamazepine, divalproex sodium, lamotrigine, valproic acid
- First-generation Antipsychotics: chlorpromazine, haloperidol, haloperidol lactate, loxapine succinate
- Second-generation (Atypical) Antipsychotics: aripiprazole, asenapine, cariprazine, clozapine, lurasidone, olanzapine, quetiapine, risperidone, ziprasidone
- Lithium Salts: lithium, lithium carbonate, lithium citrate
- Psychotherapeutic Combinations: fluoxetine-olanzapine
- Long-acting (depot) Injectable Antipsychotics: haloperidol decanoate, aripiprazole, aripiprazole lauroxil, olanzapine pamoate, risperidone microspheres

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

The target population for this measure is Medicare fee-for-service (FFS) beneficiaries with Part D coverage aged 18 years and older discharged from an inpatient psychiatric facility with a principal diagnosis of MDD, schizophrenia, or bipolar disorder.

S.7. Denominator Details *(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 S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

The denominator for this measure includes patients discharged from an IPF:

- With a principal diagnosis of MDD, schizophrenia, or bipolar disorder.
- 18 years of age or older at admission.
- Enrolled in Medicare fee-for-service Part A and Part B during the index admission and Parts A, B, and D at least 30-days post-discharge.

- Alive at discharge and alive during the follow-up period.
- With a discharge status code indicating that they were discharged to home or home health care.

The following are ICD-10-CM (clinical modification) diagnosis codes used to identify MDD, schizophrenia, or bipolar disorder:

MDD: F32.0, F32.1, F32.2, F32.3, F32.4, F32.9, F33.0, F33.1, F33.2, F33.3, F33.40, F33.41, F33.8, F33.9

Schizophrenia: F20.0, F20.1, F20.2, F20.3, F20.5, F20.81, F20.89, F20.9, F25.0, F25.1, F25.8, F25.9

Bipolar disorder: F30.10, F30.11, F30.12, F30.13, F30.2, F30.3, F30.4, F30.8, F30.9, F31.0, F31.10, F31.11, F31.12, F31.13, F31.2, F31.30, F31.31, F31.32, F31.4, F31.5, F31.60, F31.61, F31.62, F31.63, F31.64, F31.70, F31.71, F31.72, F31.73, F31.74, F31.75, F31.76, F31.77, F31.78, F31.81, F31.89, F31.9, F32.81, F32.89

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

The denominator for this measure excludes discharged patients who:

- Received electroconvulsive (ECT) during the inpatient stay or follow-up period
- Received transcranial stimulation (TMS) during the inpatient stay or follow-up period
- Were pregnant at discharge
- Had a secondary diagnosis of delirium at discharge
- Had a principal diagnosis of schizophrenia with a secondary diagnosis of dementia at discharge

S.9. Denominator Exclusion Details *(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 S.2b.)*

See Exclusions tab of attached codebook for list of codes used to define exclusions.

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including 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 with at S.2b.)*

Not applicable. The measure is not stratified.

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

Denominator:

1. Pull all IPF discharges from the Part A data.

2. Include IPF discharges for patients who were at least 18 years of age at admission.
3. Identify interim claims having the same beneficiary, provider, admission dates or having an admission date within one day of the discharge date of the previous claim and having a discharge status code of "Still patient." Collapse or combine the interim claims into one hospital stay using the admission date from the earliest claim and the discharge date from the latest claim. The data values from the latest claim are used for the newly combined hospital stay.
4. De-duplicate the IPF inpatient discharges dataset by Patient ID, Sex, Provider ID, Admission Date, and Discharge Date.
5. Remove the IPF inpatient discharges for patients who do not have Part A and Part B coverage at admission, during the entire stay, at discharge, and during the 30 days post-discharge.
6. Remove the IPF inpatient discharges who do not have a principal diagnosis of MDD, bipolar disorder, or schizophrenia using value sets containing ICD-10 codes for each of the disease conditions.
7. Remove the IPF inpatient discharges for patients who expired during the hospital stay or within 30 days of discharge.
8. Remove the IPF inpatient discharges for patients who do not have Part D coverage during the 30 days post-discharge.
9. Remove the IPF inpatient discharges for patients who were not discharged to home or home health.
10. Exclude IPF inpatient discharges who have a secondary diagnosis of pregnancy or delirium.
11. Exclude IPF inpatient discharges who have schizophrenia as the principal diagnosis with a secondary diagnosis of dementia.
12. Exclude IPF inpatient discharges who have ECT or TMS during the hospital stay or within 30 days post-discharge.

Numerator:

1. Pull all Part D claims for the evidence-based medications used for the treatment of MDD, schizophrenia, and bipolar disorder.
2. Pull all Part A and Part B claims for antipsychotic long-acting injectables (LAIs) and add them to the Part D medication claims for schizophrenia and bipolar disorder.
3. Compare the medication claims to the denominator file of eligible IPF inpatient discharges and remove any claims that occur more than two days prior to the discharge date.
4. Determine which claims occur within the follow-up period (two days prior to discharge through 30 days post-discharge) for each of the three disease conditions.
5. Total the denominator cases having at least one medication claim corresponding to the disease condition during the follow-up period.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

This measure is not based on a sample.

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

This measure is not based on survey or patient-reported data.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Claims

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

Medicare administrative data from Parts A, B, and D claims.

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Facility

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Inpatient/Hospital

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

Not applicable because this is not a composite performance measure.

2. Validity – See attached Measure Testing Submission Form

2020_Med_Cont_testing_form.docx,Updated_2020_Med_Cont_testing_form.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

Yes - Updated information is included

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:**
Measure Title: Mediation Continuation Following Inpatient Psychiatric Discharge
Date of Submission: 8/2/2020

Type of Measure:

Measure	Measure (continued)
☐ Outcome (*including PRO-PM*)	☐ Composite – *STOP – use composite testing form*
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	*

*cell intentionally left blank

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing, (e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.17*)	Measure Tested with Data From:
☐ abstracted from paper record	☐ abstracted from paper record
☒ claims	☒ claims
☐ registry	☐ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other:	☐ other:

1.2. 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*).

The Medication Continuation measure uses Medicare fee for service (FFS) Parts A, B, and D claims data.

Medicare administrative claims data

1.3. What are the dates of the data used in testing? July 1, 2017 to June 30, 2019

January 1, 2013- January 31, 2015

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.20*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☐ other:

1.5. How many and which measured entities were 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*)

Our testing dataset included 308,556 patient discharges across 1,680 inpatient psychiatric facilities. To align with other Centers for Medicare & Medicaid Services (CMS) claims-based measures, we removed inpatient claims that met the following criterion during processing prior to testing: Bill Type Code = "110": Hospital Inpatient Part A Nonpayment/Zero Claims – facilities determine an inpatient admission is not medically necessary after discharge.

Table 1.5-A. Distribution of Discharges by IPF Type (July 1, 2017 – June 30, 2019)

IPF Type	N	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max	Discharge count
Acute-care unit	1,118	143.6	136.8	1	20	45	104	199	312	953	160,517
Freestanding	562	229.2	245.4	1	13	40	157	316	569	1,504	128,792
Overall	1,680	172.2	184.9	1	18	44	114	224	409	1,504	289,309

Source: Mathematica analysis of the Medicare Fee for Service (FFS) data for the July 1, 2017 through June 30, 2019, performance period.

The measure was developed and tested using Medicare files for all inpatient psychiatric facility (IPF) discharges that occurred between January 1, 2013 and December 31, 2014. The data include 380,861 discharges from 1,694 IPFs across the United States (Table 1.5-A). IPFs ranged in size from 4 to 771 inpatient beds. Approximately 70% of IPFs in this dataset were units within a larger hospital. The average number of discharges per freestanding IPF was approximately 300 and the average per IPF unit was approximately 200.

Table 1.5-A. Distribution of Discharges by IPF Type (January 1, 2013 – December 31, 2014)

IPF Type	IPFs (N=1,694)	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
Freestanding	515	301.8	322.9	1	20	77	184	416	779	1,760
Unit	1,179	191.2	189.9	1	24	56	135	263	419	1,320
Overall	1,694	224.8	243.6	1	23	60	148	293	529	1,760

To inform the preliminary measure specifications, we conducted alpha testing, which consisted of medical record review in two IPFs at a large academic medical center in the southeast U.S.

To evaluate the validity of key elements in the claims data, we conducted similar medical record abstractions in seven additional IPFs. Test sites varied in size, type, and geographic location (Table 1.5-B).

Table 1.5-B. Characteristics of Test Sites

Study ID	State	Bed Size	Type	Teaching Facility	Type of Medical Record
1	WV	Large	Unit	Yes	EPIC
2	MI	Medium	Unit	Yes	McKesson
3	AZ	Medium	Freestanding	No	Paper Records
4	AZ	Large	Freestanding	No	Paper Records
5	MD	Large	Freestanding	Yes	Allscripts®
6	CA	Small	Unit	No	Cerner
7	LA	Large	Unit	Yes	Epic

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Data included 182,042 patients who had 308,556 discharges from 1,680 facilities within the measurement period:

- 23.6% (42,987) patients were ages 18–39 years, 41.0% (74,690) were 40–59 years old, and 35.4% (64,364) were ages 60 years or older
- 52.2% (94,946) were female and 47.8% (87,096) were male
- 74.6% (135,733) were White, 17.2% (31,251) were Black, 3.6% (6,639) were Hispanic, 2.9% (5,291) were classified as other, and 1.7% (3,128) were classified as unknown
- 58.3% (106,057) were dual Medicare and Medicaid enrollees and 41.7% (75,985) were Medicare only.

On average, 36% of discharges had a principal diagnosis of MDD, 41% of discharges had a principal diagnosis of schizophrenia, and 26% of discharges had a principal diagnosis of bipolar disorder (Table 1.6-A).

The measure is specified to require a minimum denominator size of 75 discharges, as this needed to attain an overall reliability score of at least 0.7. The restricted sample included 1,066 facilities and 268,673 discharges. When limiting to facilities with 75 or more cases during the measurement period, 32% of discharges had a principal diagnosis of MDD, 42% of discharges had a principal diagnosis of schizophrenia, and 26% of discharges had a principal diagnosis of bipolar disorder on average (Table 1.6-B).

Table 1.6-A. Distribution of Bipolar Disorder, MDD, and Schizophrenia Across IPFs

Condition	IPFs	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
Bipolar	1,641	25.9	10.0	2.1	14.3	19.5	25.0	31.6	37.5	100
MDD	1,621	35.7	17.8	1.0	13.3	23.0	34.2	46.4	60.0	100
Schizophrenia	1,651	41.1	19.0	2.8	18.5	27.6	39.5	52.1	66.7	100

Table 1.6-B Distribution of Bipolar Disorder, MDD, and Schizophrenia Across IPFs with Denominator ≥ 75

Condition	IPFs	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
Bipolar	1,092	26.1%	8.4%	2.3%	15.9%	20.6%	25.4%	31.4%	36.1%	81.3%
MDD	1,092	32.0%	14.2%	1.0%	13.6%	21.9%	31.9%	41.1%	49.4%	88.6%
Schizophrenia	1,093	42.0%	15.5%	3.9%	22.8%	31.7%	40.8%	51.6%	62.6%	100.0%

This measure was developed for adult admissions to an IPF with a principal diagnosis of major depressive disorder (MDD), schizophrenia, or bipolar disorder. Eligible patients were enrolled in Medicare Parts A, B, and D during the admission and follow-up period. The final cohort includes 380,861 discharges. On average, 35% of discharges had a principal diagnosis of MDD, 40% of discharges had a principal diagnosis of schizophrenia, and 27% of discharges had a principal diagnosis of bipolar disorder (Table 1.6-A). When limiting to facilities with 75 or more cases during the measurement period (rationale provided in Section 2a.2), 30% of discharges had a principal diagnosis of MDD, 43% of discharges had a principal diagnosis of schizophrenia, and 27% of discharges had a principal diagnosis of bipolar disorder on average (Table 1.6-B). The patients in the claims data were 51% male, 84% under age 65, and 70% dually enrolled. The racial and ethnic groups represented were 72% white, 21% black, and 4% Hispanic.

Table 1.6-A. Distribution of Bipolar Disorder, MDD, and Schizophrenia Across IPFs

Condition	IPFs	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
MDD	1,651	34.8	19.0	0.8	11.7	21.4	32.5	45.7	61.4	100
Schizophrenia	1,655	40.2	19.9	0.6	15.2	25.7	38.0	52.7	67.4	100
Bipolar Disorder	1,658	27.3	11.8	1.0	14.3	20.0	26.1	33.3	40.6	100

Table 1.6-B Distribution of Bipolar Disorder, MDD, and Schizophrenia Across IPFs with Denominator ≥ 75

Condition	IPFs	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
MDD	1,182	29.5	14.7	0.8	11.2	19.3	28.7	38.6	48.1	91.3
Schizophrenia	1,184	43.1	17.3	0.6	23.1	30.7	40.8	54.0	67.2	96.1
Bipolar Disorder	1,184	27.4	9.5	1.0	15.7	21.1	26.9	33.3	39.7	76.3

1.7. 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 reported below.

Not applicable.

Most data analysis was conducted in claims data. As noted in Section 1.5, alpha testing data from medical record review at two sites helped to inform the measure specifications. Medical records for 166 discharges were abstracted by two clinicians.

The field testing that informed the validity of key data elements was conducted by two nurses at each facility. Each nurse abstracted medical records for 75 discharges each for a total of 150. Twenty percent of each nurse's discharges were randomly selected and assigned to the other nurse abstractor to assess the reliability of the nurse abstractions. Additionally, two clinicians per facility reviewed a sub-sample (10 percent) of the medical records of the 150 discharges to determine the validity of the principal diagnosis, based on information contained in the record. Fifty percent of each clinician's discharges were randomly selected and assigned to the other clinician abstractor to assess the reliability of the clinician abstractions. Reliability scores between the two clinicians were calculated.

At the start of testing, each test site received a one-hour training by HSAG on the abstraction instructions and process and a one-hour follow-up meeting after review of the first 10 medical records to provide clarifications, if needed.

The abstraction tool that was used by all field testing sites is provided in the measure technical report in the supplemental materials.

1.8 What were 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.

As described in section 1.6, the following variables are collected with claims data: gender, age, race, and payer. This measure is based on a process that should be carried out for all patients (except those excluded), so no adjustment for patient mix is necessary.

Not applicable. The measure is not risk-adjusted or stratified.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level 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*)

Signal-to-noise reliability. The signal-to-noise (SNR) statistic, R (ranging from 0 to 1), summarizes the proportion of the variation between facility scores on a measure that is due to real differences in underlying facility characteristics (such as differences in medical care) as opposed to background-level or random variation (for example, due to measurement or sampling error). If R = 0, all observed variation is due to sampling error. In this case, the measure is not useful to distinguish between entities with respect to healthcare quality. Conversely, if R = 1, all entity scores are free of sampling error, and all variation represents real differences between entities in the measure result.

We estimated SNR reliability for the Medication Continuation measure in three steps (Adams 2009; Adams 2014; NQF 2016). First, we calculated facility-specific Medication Continuation variance ("noise") as a function of the rate at each facility and the facility sample size (number of discharges from that facility), *n*:

$$\sigma^2_{within} = \frac{\hat{p}(1-\hat{p})}{n} (1);$$

Second, we used version 2.2 of the BETABIN SAS macro written by Wakeling to fit the beta-binomial model to the Medication Continuation dataset (Wakeling n/d). The macro produced the estimated average pass rate across all facilities, as well as the Alpha (α) and Beta (β) parameters that describe the shape of the fitted beta-binomial distribution. We calculated the "signal" (between-facility variation on the Medication Continuation measure) using these parameters:

$$\sigma^2_{between} = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} (2);$$

Third, we calculated the SNR reliability as the ratio of the between-level variance and the total variance (that is, the sum of the between-level and within-level variances) of the Medication Continuation measure rate:

$$Reliability = \frac{\sigma^2_{between}}{\sigma^2_{between}+\sigma^2_{within}} (3);$$

To examine the reliability of the measure score, we utilized the approach proposed by Adams (2009) and Scholle et al. (2008) to assess measure precision in the context of the observed variability across IPFs. The following is quoted from the tutorial published by Adams:

> "Reliability is a key metric of the suitability of a measure for [provider] profiling because it describes how well one can confidently distinguish the performance of one physician from another. Conceptually, it is the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in performance. There are three main drivers of reliability: sample size, differences between physicians, and measurement error. At the physician level, sample size can be increased by increasing the number of patients in the physician's data as well as increasing the number of measures per patient."

For this measure, the signal-to-noise ratio was calculated as a function of the variance between IPFs (signal) and the variance within an IPF (noise). Reliability was estimated using a beta-binomial model. This approach has two basic assumptions:

1. Each measured entity has a true pass rate, p, which varies; and,
2. The measured entity's score is a binomial random variable conditional on the measured entity's true value, which comes from the beta distribution.

Reliability scores vary from 0.0 to 1.0. A score of 0.0 implies that all variation is attributed to measurement error (noise); whereas, a reliability of 1.0 implies that all variation is caused by a real difference in performance (across IPFs). In a simulation, Adams showed that differences between physicians started to be seen at reliability of 0.7, and significant differences could be seen at reliability of 0.9. Our rationale was based on Adams' work; thus, a minimum reliability score of 0.7 was used to indicate sufficient signal strength to discriminate performance between IPFs.

Using methodology described by Scholle et al. (2008), reliability estimates were computed separately, based on the mean denominator size for IPFs within each denominator category. As Scholle described in the article, the reliability estimate at the mean denominator for each category should reflect "the typical experience of IPFs in this population."

*Adams, J. L. The reliability of provider profiling: A tutorial. Santa Monica, California: RAND Corporation. TR-653-NCQA, 2009.

*Scholle, S. H., Roski, J., Adams, J. L., Dunn, D. L., Kerr, E. A., Dugan, D. P., et al. (2008). Benchmarking physician performance: Reliability of individual and composite measures. *American Journal of Managed Care, 14*(12), 833-838.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

Table 2a2.3 summarizes the mean and range of the reliability statistic for the Medication Continuation measure, which was calculated separately by facility. The mean reliability across all 1,066 facilities with at least 75 denominator cases exceeded the 0.70 threshold for acceptable reliability. The 25th percentile for the measure reliability was 0.70, and the 75th percentile was 0.81.

Table 2a.2.3. Comparison of IPF Measure Score Distribution by Denominator Minimum

Denominator	# IPFs (%)	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
Denominator >=75	1,066 (63.5%)	75.1	8.3	34.8	63.4	70.1	76.2	81.1	84.7	94.3
Overall	1,680 (100%)	75.0	12.8	0.0	61.8	70.0	76.8	82.6	87.5	100.0

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017–June 30, 2019 performance period.

A minimum denominator size of 75 discharges is needed to attain an overall reliability score of at least 0.7 (Table 2a.2.3-A), which is within acceptable norms and indicates sufficient signal strength to discriminate performance between facilities, using the method of mean denominator and volume categories. With a minimum denominator of 75 discharges, 1,184 IPFs (70%) have enough discharges within a two-year measurement period for public reporting. The removal of smaller facilities does not have an appreciable impact on the distribution of measure scores (Table 2a.2.3-B).

Table 2a.2.3-A. IPF Reliability and Assessment of Adequacy for Tests Conducted

Measures	Minimum Denominator	# of IPFs N=1,694 (%)	Mean Rate (%) of IPFs	Reliability Score
Overall	75	1,184 (69.9)	78.0	0.77

Table 2a.2.3-B. Comparison of IPF Measure Score Distribution by Denominator Minimum

Measures	# IPFs	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
Overall	1,694	78.0	11.1	0.0	66.7	73.6	79.6	84.4	88.3	100.0
Denominator ≥ 75	1,184	78.0	7.9	21.1	68.3	73.9	79.1	83.4	86.5	98.5

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

The mean reliability, as well as the 25th percentile, across all facilities exceeded the 0.70 threshold for acceptable reliability. Reliability above 0.7 indicates that the measure can be judged to be reliable (Glance et al. 2019).

References:

Glance, L.G., K.J. Maddox, K. Johnson, D. Nerenz, D. Cella, B. Borah, J. Kunisch, et al. 2019. "National Quality Forum Guidelines for Evaluating the Scientific Acceptability of Risk-Adjusted Clinical Outcome Measures." A Report From the National Quality Forum Scientific Methods Panel. *Annals of Surgery*: June 2020 – vol. 271, no. 6, June 2020, -pp. 1048–1055. Available at https://doi.org/10.1097/SLA.0000000000003592. Accessed July 1, 2020.

National Quality Forum. "Guidance for Measure Testing and Evaluating Scientific Acceptability of Measure Properties." 2011. Available at https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=70943. Accessed July 9, 2019.

The results indicate the measure score is reliable by adjusting the minimum case size for the denominator to require at least 75 cases during the measurement period. To increase the number of IPFs that have at least 75 cases during the measurement period, we recommend using a two-year measurement period.

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☒ **Empirical validity testing**

☐ **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*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. 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)*

We examined validity of the Medication Continuation measure using the known-group method. A measure is considered to exhibit known-group validity if the measure score can be used to discriminate between subgroups of patients known to have differences in the measure rates based on findings from the literature. We investigated known-group validity by evaluating differences in mean Medication Continuation facility

scores among predefined groups of patients based on the evidence from peer-reviewed studies. These studies examined factors related to nonadherence to psychotropic medication among patients with major psychiatric disorders. Consistent with the literature, IPF-level Medication Continuation measure scores were hypothesized to be lower based on evidence demonstrated (that is. worse medication adherence) among younger patients (<40 years old) (Garcia et al. 2016; Sajatovic et al. 2007); male patients (Chakrabarti 2017; Lacasta-Tintorer 2011) patients with a comorbid Substance Use Disorder (SUD) diagnosis (Garcia et al. 2016; Chakrabarti 2017; Sajatovic et al. 2007; Velligan et al. 2017); non-White patients (Fleck et al. 2005; Sajatovic et al. 2007); patients with a diagnosis of schizophrenia (Chakrabarti 2017; Higashi et al. 2013; Sajatovic et al. 2007); and more disadvantaged patients with problems accessing medication and limited socioeconomic resources (Lanouette et al. 2009; Jawad et al. 2018). We used the beneficiaries' dual Medicare-Medicaid status as a proxy for socioeconomic status.

To test for the differences in the Medication Continuation measure rates by patient subgroups, we first calculated measure rates for each subgroup by facility. Then, we computed mean rate and standard deviations by subgroup across all facilities. For dichotomous variables, we used t-tests to compare mean group differences. With large sample sizes, small differences that are statistically significant may not always be practically or clinically meaningful. Therefore, we additionally computed Cohen's (1988) d effect size (the difference in mean scores divided by the pooled standard deviation). A d of 1 indicates the two groups differ by 1 standard deviation, a d of 2 indicates they differ by 2 standard deviations, and so on. Following Cohen's (1988) definitions, we defined effect size values for dichotomous variables as small (0.2), medium (0.5), or large (0.8). For patient subgroups with more than two categories (age and diagnosis), we computed Eta-squared (η2) effect size to capture the overall difference in the measure rate between groups. We categorized corresponding effect size values as small (0.01), medium (0.06), or large (0.14).

References:

Chakrabarti, S. (2017). "Medication Non-Adherence in Bipolar Disorder: Review of Rates, Demographic and Clinical Predictors." Available at https://doi.org/10.13105/wjma.v5.i4.103. Accessed July 1, 2020.

Fleck, D.E., P.E. Keck, K.B. Corey, and S.M. Strakowski. (2005). "Factors Associated With Medication Adherence in African American and White Patients with Bipolar Disorder. *Journal of Clinical Psychiatry,* vol. 66, no. 5, 2006, pp. 646–652. Available at https://doi.org/10.4088/JCP.v66n0517. Accessed July 1, 2020.

García, S., M. Martínez-Cengotitabengoa, S. López-Zurbano, I. Zorrilla, P. López, E. Vieta, E., and A. González-Pinto, (2016). "Adherence to Antipsychotic Medication in Bipolar Disorder And Schizophrenic Patients." *Journal of Clinical Psychopharmacology*, vol. 36, no. 4, August 2016, pp. 355–371. Available at https://doi.org/10.1097/JCP.0000000000000523. Accessed July 1, 2020.

Higashi, K., G.Medic, K.J. Littlewood, T. Diez, O. Granström, O., and M. de Hert. (2013). Medication Adherence in Schizophrenia: Factors Influencing Adherence and Consequences of Nonadherence: A Systematic Literature Review. *Therapeutic Advances in Psychopharmacology,* vol. 3, no. 4, August 2013, pp. 200–218. Available at https://doi.org/10.1177/2045125312474019. Accessed July 1, 2020.

Jawad, I., S. Watson, P.M. Haddad, P.S. Talbot, and R.H. McAllister-Williams. (2018). Medication Nonadherence in Bipolar Disorder: A Narrative Review. Available at https://doi.org/10.1177/2045125318804364. Accessed July 1, 2020.

Lacasta-tintorer, D. R. García-lecina, D. Flamarich-zampalo, T. Font-canal, and M.I.F. Martín MIF. Do Depressed Patients Comply with Treatments Prescribed? A Cross-Sectional Study of Adherence to the Antidepressant Treatment." *Actas Esp Psiquiatr*, vol. 39, no. 5, 2011, pp. 288–293.

Lanouette, N., D. vFolsom, A. Sciolla, A., and D. Jeste. "Psychotropic Medication Nonadherence Among United States Latinos: A Comprehensive Literature Review." *Psychiatric Services*, vol. 60, no. 2, 2009, pp. 157–174. Available at https://doi.org/10.1176/appi.ps.60.2.157. Accessed July 1, 2020.

McCann, T.V., G. Boardman, E. Clark, and S. Lu. "Risk Profiles for Nonadherence to Antipsychotic Medications." *Journal of Psychiatric and Mental Health Nursing*, vol. 15, no. 8, 2008, pp. 622–629.

Sajatovic, M., O. Elhaj, E.A. Youngstrom, S.R. Bilali, D.J. Rapport, S.J. Ganocy, and J.R. Calabrese. "Treatment Adherence in Individuals With Rapid Cycling Bipolar Disorder." *Journal of Clinical Psychopharmacology*, vol. 27, no. 4, 2007, pp. 412–414.

Velligan, D.I., M. Sajatovic, A. Hatch, P. Kramata, and J. Docherty, (2017). "Why do Psychiatric Patients Stop Antipsychotic Medication? A Systematic Review of Reasons for Nonadherence to Medication in Patients with Serious Mental Illness." *Patient Preference and Adherence*, Vol. 3. no. 4, March 2017, pp. 200-218. Available at https://doi.org/10.2147/PPA.S124658. Accessed July 1, 2020.

Critical data elements

Two psychiatrists reviewed 150 patients' medical records to ensure that the claims data are accurate in identifying several key data elements for calculating the measure. First, the clinicians recorded their assessment of the patient's principal discharge diagnosis based on information in the medical record. These findings were compared to the principal diagnoses in the claims. We evaluated the positive predictive value using the clinical assessment from the medical record as the "gold standard" because this shows how often a diagnosis in the claims agrees with the diagnosis from the medical record. A high positive predictive value indicates a high probability that a claim for a certain condition (e.g., schizophrenia) correctly predicts the principal discharge diagnosis in the medical record.

Next, at the seven test sites, abstractors were asked to indicate whether a prescription was provided at discharge. When an evidence-based prescription was not provided, they were asked to provide the rationale from the medical record to determine if additional exclusion criteria should be applied to the measure. The information on whether at least one prescription for an evidence-based medication was provided at discharge was compared to the numerator based on claims data. We evaluated the positive predictive value using the prescription at discharge as the "gold standard". The positive predictive value indicates that most patients who filled an evidence-based prescription during the follow-up period also received an evidence-based prescription from the IPF at discharge.

Finally, abstractors from the seven test sites were asked to record whether there was an indication in the medical record that medications had been dispensed to the patient free at discharge, as those medications would not appear in the claims data.

To ensure that the abstraction results were reliable, 10% of the cases were reviewed by both clinicians, and their results were compared to assess agreement.

Performance measure score

Measure scores were compared to three related measures:

1. Follow-Up After Hospitalization (7-Day)
2. Follow-Up After Hospitalization (30-Day)
3. IPF All-Cause Unplanned Readmission Measure

We tested the measure distributions for normality at each unit of analysis, selected the appropriate statistical test for the distribution, and assessed the significance of the correlation coefficient. We would expect the scores for the 7- and 30-day Follow-Up After Hospitalization measure to be positively correlated with the medication continuation scores because these are care coordination measures and higher scores indicate higher quality. We would expect the medication continuation scores to be negatively correlated with the all-cause unplanned readmission measure scores, because readmissions may indicate a lack of care coordination and higher scores on the readmission measure indicate lower quality.

Face validity of the measure score was assessed by the IPF Technical Expert Panel (TEP). Specifically, the TEP members were asked whether they agreed, disagreed, or were unable to rate the following statement:

> *The performance rating from the continuation of medication measure, as specified, represents an accurate reflection of facility-level rates of evidence-based medication continuation for MDD, schizophrenia, or bipolar disorder following discharge from an IPF.*

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

As shown in Table 2b1.3, we found multiple instances of known-group validity for the Medication Continuation measure.

Table 2b1.3. Differences in the Medication Continuation rates by patient group

Grouping variable	Patient subgroups	Medication continuation measure rates (%): All facilities	Medication continuation measure rates (%): Facilities with ≥ 75 discharges	Effect size (Cohen's d) for differences in means between patient groups: All facilities	Effect size (Cohen's d) for differences in means between patient groups: Facilities with ≥ 75 discharges
Sex	Male patients (hypothesized lower rate)	72.1%	72.2%	0.39	0.64
*	Female patients	77.9%	78.0%	*	*
SUD diagnosis	SUD (hypothesized lower rate)	70.4%	69.7%	0.41	0.74
*	No SUD	76.9%	77.4%	*	*
Dual status	Dual (hypothesized lower rate)	77.4%	77.6%	0.51	0.85
*	Non-dual	69.8%	69.1%	*	*
Race	Non-White (hypothesized lower rate)	71.1%	71.2%	0.31	0.46
*	White	76.2%	76.3%	*	*
Grouping variable	**Patient subgroups**	**Medication continuation measure rates (%): All facilities**	**Medication continuation measure rates (%): Facilities with ≥ 75 discharges**	**Effect size (Eta^2) for differences in means between patient groups: All facilities**	**Effect size (Eta^2) for differences in means between patient groups: Facilities with ≥ 75 discharges**
Diagnosis	Schizophrenia (hypothesized lower rate)	75.5%	76.1%	0.001	0.013
*	MDD	74.2%	73.2%	*	*
*	Bipolar disorder	75.3%	75.2%	*	*

Grouping variable	Patient subgroups	Medication continuation measure rates (%): All facilities	Medication continuation measure rates (%): Facilities with ≥ 75 discharges	Effect size (Cohen's d) for differences in means between patient groups: All facilities	Effect size (Cohen's d) for differences in means between patient groups: Facilities with ≥ 75 discharges
Age	18-39 (hypothesized lower rate)	74.0%	74.7%	0.004	0.0001
*	40-59	74.1%	74.8%	*	*
*	≥60	75.4%	74.9%	*	*

*cell intentionally left blank

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period. Facilities with less than 75 discharges were excluded from the analysis. Results based on 1,680 inpatient psychiatric facilities with a total of 308,556 eligible discharges (full sample data), and 1,066 inpatient psychiatric facilities and 268,673 discharges (≥75 discharges).

Notes: The differences in the measure rates by sex, SUD diagnosis, dual Medicare-Medicaid enrollment and race were significant at $p \leq 0.01$ for all hospitals and hospitals with ≥75 discharges. The differences in the measure rates by age groups were statistically significant at $p \leq 0.05$ for all hospitals but were not statistically significant for hospitals with ≥75 discharges. The differences in the measure rates by diagnosis code were statistically significant at $p \leq 0.05$ for all hospitals and $p \leq 0.01$ for hospitals with ≥75 discharges.

Critical data elements

The positive predictive value of the claims data was 97% (921/945) (Table 2b2.3-A). The positive predictive values were similar across all three conditions, with 98% (289/294) for MDD, 98% (328/335) for schizophrenia, and 96% (304/316) for bipolar disorder.

Table 2b2.3-A. Agreement Between Medical Record and Claims for Diagnoses

Measures	Diagnosis In Medical Record	Diagnosis Not in Medical Record	Total
MDD	*	*	*
MDD in claims	289	5	294
No MDD in claims	6	0	6
Total MDD	295	5	300
Schizophrenia	*	*	*
Schizophrenia in claims	328	7	335
No schizophrenia in claims	9	0	9
Total schizophrenia	329	7	344
Bipolar Disorder	*	*	*
Bipolar disorder in claims	304	12	316
No bipolar disorder in claims	3	0	3
Total bipolar disorder	307	12	319
Total Overall	939	24	963

*cell intentionally left blank

During the medical record review at the 7 test sites, 92% (873/945) of cases were prescribed an evidence-based medication at discharge (Table 2b2.3-B). Among the patients who were not prescribed an evidence-based medication, the majority of reasons identified by the medical record abstractors indicated quality deficits. For example, 61% of the cases without an evidence-based medication at discharge had medications prescribed that were not indicated for the principal discharge diagnosis, 11% did not have any medications prescribed, and 5% were clearly the result of medical errors. No reason was identified by the abstractors for 9% of the cases, which could also indicate potential quality deficits. The remaining cases do not represent quality deficits but do indicate opportunities for improvement in cases where prescriptions could have been provided in addition to medications dispensed at discharge or could have been provided to patients who declined pharmacotherapy because the patient may decide differently and want to continue pharmacotherapy after leaving the IPF.

When comparing numerator positive cases from the claims data to the medical record, the positive predictive value was 96% (622/646) as calculated from Table 2b2.3-B.

Table 2b2.3-B. Comparison of Medications Prescribed at Discharge to Fills During the Follow-Up Period in Claims Data

Measures	Evidence-Based Prescription at Discharge	No Evidence-Based Prescription at Discharge	Total
Numerator Positive	622	24	646
Numerator Negative	251	48	299
Total	873	72	945

The medical record review found that there were few discharges where the facility provided medications to patients at discharge. Among those discharges, some of the medications provided were filled for the patient through an outpatient pharmacy and appeared in the claims data.

Performance measure score

Results of the analysis for correlations of medication continuation scores with the three conceptually related Inpatient Psychiatric Facility Quality Reporting (IPFQR) measures are included in Table 2b2.3-C. The medication continuation scores were moderately correlated with the scores for 7- and 30-day follow-up after hospitalization for mental illness scores as expected ($\rho = 0.34$ and 0.43). The medication continuation scores were negatively correlated with readmission scores as expected ($\rho = -0.26$). All correlations are statistically significant at p-value < 0.0001.

After reviewing these results and the proposed measure specifications, all of the 10 TEP members who were present for the face validity vote agreed that the measure score had face validity.

Table 2b2.3-C. Performance Measure Score Correlation

Measures	IPFs	Correlation
Follow-Up After Hospitalization 7-day (7/1/2014 – 6/30/2015)	1,145	0.34312
Follow-Up After Hospitalization 30-day (7/1/2014 – 6/30/2015)	1,145	0.43065
IPF All-Cause Unplanned Readmission Measure (Observed) (1/1/2013 – 12/31/2014)	1,184	-0.26059

2b1.4. What is 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?*)

The known group validity of the Medication Continuation measure was shown by comparing adherence rates to medication between groups of patients with a priori expected differences in adherence to psychotropic medication (i.e., by age, sex, race, presence of comorbid SUD diagnosis, SES (dual status), and principal diagnosis. Consistent with our hypotheses, we observed lower Medication Continuation measure rates (i.e. worse adherence to medication post-discharge) for patients with comorbid SUD, for non-white patients, for male patients, and for younger patients. Other studies reported similar patterns of differences in the adherence rates by these sub-groups of patients, which confirms the validity of the Medication Continuation measure in discriminating between these subgroups of patients (see e.g. Chakrabarti, 2017; Garcia et al., 2016; Higashi et al., 2013; Lacasta-Tintorer, 2011; Sajatovic et al., 2007; Velligan et al., 2017). The Medication Continuation measure was also able to detect differences in medication adherence rates between patients 1) enrolled in Medicare only and those with both Medicare and Medicaid coverage and 2) with different principal diagnosis at discharge, although the pattern of differences in the rates was in the direction opposite from what we expected. Overall, observed ability of the Medication Continuation measure to discriminate between the compared groups in respect to their adherence to prescribed medication supports its validity.

Consistent with the literature, we observed substantially lower Medication Continuation measure rates (that is, worse adherence to medication post-discharge) for patients with comorbid SUD, non-White patients, male patients, and younger patients.

Critical data elements

The medical record review in the two initial test sites confirmed that the principal discharge diagnoses in the administrative claims data are a valid source for identifying the primary cause of admission to the IPF.

The medical record review from the additional 7 test sites confirmed that the construct of medication continuation is valid for assessing IPF quality because most patients who filled a prescription during the follow-up period received a prescription from the IPF at discharge. A quality deficit was identified for most patients who were not provided a prescription for an evidence-based medication at discharge so no additional exclusion criteria were applied to the measure as the result of this analysis.

Finally, the medical record review at the seven test sites confirmed that the claims data are valid for identifying all prescription fills in this patient population because medications provided at discharge were filled using the patient's insurance, which would appear in the claims data. We anticipate that free medications are provided to the patient population for this measure less frequently because all patients included in the measure denominator are enrolled in Medicare Part D. Low-income Medicare patients can receive assistance with co-pays, and patients who are dually enrolled in Medicaid (70% of this cohort) receive additional assistance covering the costs of medications that are not covered by Medicare. Notes from the medical record abstractors indicate that all of the medications provided at discharge were for 30-day supplies or less. Therefore, the patients who received medications at discharge on Day 0 would need to fill a prescription for an evidence-based medication before the end of the 30-day follow-up period to avoid gaps in treatment. Those fills would also appear in the claims data.

Performance measure score

The moderate strength of the correlations, conceptually supported directionality, and unanimous face validity assessment add further support that the measure is valid as specified.

_____

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section 2b4*

The denominator for this measure excludes discharged patients who:

- Received electroconvulsive (ECT) therapy during the inpatient stay or follow-up period.
- Received transcranial magnetic stimulation (TMS) during the inpatient stay or follow-up period.
- Were pregnant during the inpatient stay.
- Had a secondary diagnosis of delirium.
- Had a principal diagnosis of schizophrenia with a secondary diagnosis of dementia.

2b2.1. Describe the method of testing exclusions and what it tests (*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*)

To assess the effect of these exclusions, we examined the number of IPF discharges affected by each exclusion and calculated and compared the measure rates with and without each exclusion.

All exclusion analyses were conducted using Medicare claims data from inpatient psychiatric stays at IPFs where the patients were discharged alive with Parts A, B, and D enrollment during the follow-up period.

1. **Electroconvulsive therapy (ECT)**

   We compared the medication continuation rates of patients with ECT during the admission or follow-up period to those of patients without ECT during the admission or follow-up period. We also conducted a medical record review to evaluate whether evidence-based medications were prescribed at discharge to patients who received ECT or a recommendation for ECT.

2. **Transcranial magnetic stimulation (TMS)**

   We compared the medication continuation rates for patients with TMS during the admission or follow-up period to those of patients without TMS during the admission or follow-up period.

3. **Pregnancy**

   We compared the medication continuation rates for patients who were pregnant during the admission to those of patients who were not pregnant during the admission.

4. **Secondary diagnosis of delirium**

   We compared the medication continuation rates for patients with delirium during the admission to those of patients without delirium during the admission.

5. **Principal diagnosis of schizophrenia with secondary diagnosis of dementia**

   Antipsychotics may be contraindicated for patients with dementia. Antipsychotics are included in the numerator for schizophrenia and bipolar disorder. However, alternative pharmacotherapies are available for bipolar disorder that meet the numerator criteria, so we only compared the medication continuation rates for patients with a principal diagnosis of schizophrenia and a secondary diagnosis of dementia to those of patients with no dementia.

2b2.2. What were 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*)

Tables 2b2.2A–2b2.2E summarize the IPF discharges omitted by exclusion type.

Table 2b2.2A. Frequency of exclusion for ECT and performance rates with and without exclusion

Principal condition	All IPF discharges: Frequency	All IPF discharges: Perf rate	Discharges with ECT: Frequency	Discharges with ECT: % all discharges for condition	Discharges with ECT: Perf rate	Discharges without ECT: Frequency	Discharges without ECT: % all discharges for condition	Discharges without ECT: Perf rate
MDD	95,494	73.8	4,525	4.7	82.1	90,969	95.3	73.4
Schizophrenia	131,409	74.6	1,826	1.4	82.1	129,583	98.6	74.5
Bipolar disorder	81,653	74.5	2,480	3.0	77.3	79,173	97.0	74.4
Overall	308,556	74.3	8,831	2.9	80.7	299,725	97.1	74.1

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

Table 2b2.2B. Frequency of exclusion for TMS and performance rates with and without exclusion

Principal condition	All IPF discharges: Frequency	All IPF discharges: Perf rate	Discharges with TMS: Frequency	Discharges with TMS: % all discharges for condition	Discharges with TMS: Perf rate	Discharges without TMS: Frequency	Discharges without TMS: % all discharges for condition	Discharges without TMS: Perf rate
MDD	95,494	73.8	216	0.2	80.6	95,278	99.8	73.8
Schizophrenia	131,409	74.6	15	0.0	80.0	131,394	100.0	74.6
Bipolar disorder	81,653	74.5	56	0.1	76.8	81,597	99.9	74.5
Overall	308,556	74.3	287	0.1	79.8	308,269	99.9	74.3

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

Table 2b2.2C. Frequency of exclusion for pregnancy and performance rates with and without exclusion

Principal condition	All IPF discharges: Frequency	All IPF discharges: Perf rate	Discharges with pregnancy: frequency	Discharges with pregnancy: % all discharges for condition	Discharges with pregnancy: Perf rate	Discharges without pregnancy: Frequency	Discharges without pregnancy: % all discharges for condition	Discharges without pregnancy: Perf rate
MDD	95,494	73.8	47	0.0	48.9	95,447	100.0	73.8
Schizophrenia	131,409	74.6	93	0.1	69.9	131,316	99.9	74.6
Bipolar disorder	81,653	74.5	101	0.1	62.4	81,552	99.9	74.5
Overall	308,556	74.3	241	0.1	62.7	308,315	99.9	74.3

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

Table 2b2.2D. Frequency of exclusion for delirium and performance rates with and without exclusion

Principal condition	All IPF discharges: Frequency	All IPF discharges: Perf rate	Discharges with delirium: Frequency	Discharges with delirium: % all discharges for condition	Discharges with delirium: Perf rate	Discharges without delirium: Frequency	Discharges without delirium: % all discharges for condition	Discharges without delirium: Perf rate
MDD	95,494	73.8	2,618	2.7	74.6	92,876	97.3	73.8
Schizophrenia	131,409	74.6	2,649	2.0	78.8	128,760	98.0	74.5
Bipolar disorder	81,653	74.5	2,011	2.5	77.4	79,642	97.5	74.4
Overall	308,556	74.3	7,278	2.4	76.9	301,278	97.6	74.3

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

Table 2b2.2E. Frequency of exclusion for primary diagnosis of schizophrenia and secondary diagnosis of dementia and performance rates with and without exclusion

Principal condition	All IPF discharges: Frequency	All IPF discharges: Perf rate	Schizophrenia discharges with secondary diagnosis of dementia: Frequency	Schizophrenia discharges with secondary diagnosis of dementia: % all discharges for condition	Schizophrenia discharges with secondary diagnosis of dementia: Perf rate	Schizophrenia discharges without secondary diagnosis of dementia: Frequency	Schizophrenia discharges without secondary diagnosis of dementia: % all discharges for condition	Schizophrenia discharges without secondary diagnosis of dementia: Perf rate
Schizophrenia	131,409	74.6	3,375	2.6	76.5	128,034	97.4	74.5

Source: Mathematica analysis of the Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

Table 2b2.2F summarizes the mean, 95% confidence interval, and illustrates the difference from the national rate (75.0%).

Table 2b2.2F. Mean performance rate and 95% confidence interval by exclusion type

Exclusion type	Mean	95% CI	Difference from National Rate
All exclusions applied	75.03	74.42 – 75.64	*
No exclusions applied	75.07	74.46 – 75.68	No Difference
All exclusions applied except received ECT	75.07	74.46 – 75.69	No Difference
All exclusions applied except received TMS	75.03	74.42 – 75.64	No Difference
All exclusions applied except pregnant	75.02	74.41 – 75.63	No Difference
All exclusions applied except delirium	75.00	74.39 – 75.62	No Difference
All exclusions applied except schizophrenia with dementia	75.02	74.41 – 75.63	No Difference

Source: Mathematica analysis of Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

*cell intentionally left blank

Table 2b3.2-A. Frequency of ECT During or After the Index Admission

Principal Condition	All IPF Admissions: Frequency	All IPF Admissions: % Rx	ECT During Admission Or Follow-Up Period: Frequency	ECT During Admission Or Follow-Up Period: % Total	ECT During Admission Or Follow-Up Period: % Rx	No ECT During Admission Or Follow-Up Period: Frequency	No ECT During Admission Or Follow-Up Period: % Total	No ECT During Admission Or Follow-Up Period: % Rx
MDD	139,355	71.7	7,414	5.3	76.3	131,941	94.7	71.4
Schizophrenia	217,417	75.6	3,086	1.4	77.3	214,331	98.6	75.5
Bipolar disorder	132,376	75.5	4,474	3.4	74.6	127,902	96.6	75.6
Overall	489,148	74.5	14,974	3.1	76.0	474,174	96.9	74.4

Table 2b3.2-B. Frequency of TMS During the Stay or After the Index Admission for MDD, Schizophrenia, or Bipolar Disorders

Principal Condition	All IPF Admissions: Frequency	All IPF Admissions: % Rx	TMS During Admission Or Follow-Up Period: Frequency	TMS During Admission Or Follow-Up Period: % Total	TMS During Admission Or Follow-Up Period: % Rx	No TMS During Admission Or Follow-Up Period: Frequency	No TMS During Admission Or Follow-Up Period: % Total	No TMS During Admission Or Follow-Up Period: % Rx
Overall	489,148	74.5	76	0.0	76.3	489,072	100.0	74.5

Table 2b3.2-C. Follow-Up Rates for Patients Who Are and Are Not Pregnant

Condition	All IPF Admissions: Frequency	All IPF Admissions: % Rx	Pregnant: Frequency	Pregnant: % Total	Pregnant: % Rx	Not Pregnant: Frequency	Not Pregnant: % Total	Not Pregnant: % Rx
MDD	139,355	71.7	59	0.0	59.3	139,296	99.9	71.7
Schizophrenia	217,417	75.6	138	0.1	59.4	217,279	99.9	75.6
Bipolar disorder	132,376	75.5	134	0.1	61.9	132,242	99.9	75.5
Overall	489,148	74.5	331	0.1	60.4	488,817	99.9	74.5

Table 2b3.2-D. IPF Admissions with Secondary Delirium Diagnosis

Principal Condition	All IPF Admissions: Frequency	All IPF Admissions: % Rx	Delirium: Frequency	Delirium: % Total	Delirium: % Rx	No Delirium: Frequency	No Delirium: % Total	No Delirium: % Rx
MDD	139,355	71.7	3,420	2.5	66.5	135,935	97.5	71.8
Schizophrenia	217,417	75.6	3,837	1.8	71.9	213,580	98.2	75.6
Bipolar disorder	132,376	75.5	2,385	1.8	73.2	129,991	98.2	75.6
Overall	489,148	74.5	9,642	2.0	70.3	479,506	98.0	74.5

Table 2b3.2-E. IPF Admissions with Principal Diagnosis of Schizophrenia and Secondary Diagnosis of Dementia

Principal Condition	All IPF Admissions: Frequency	All IPF Admissions: % Rx	Secondary Dementia: Frequency	Secondary Dementia: % Total	Secondary Dementia: % Rx	No Dementia: Frequency	No Dementia: % Total	No Dementia: % Rx
Schizophrenia	217,417	75.6	6,971	3.2	65.3	210,446	96.8	75.9

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., 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*)

Applying all the exclusions did not significantly change the mean measure rate. This finding is largely in keeping with testing from the initial endorsement submission, which also found similar mean scores and relatively low rates of exclusions, particularly for pregnancy and receipt of TMS. We believe these exclusions should be retained as the clinical reasoning behind them has not changed.

1. **ECT**

   ECT procedures are used as a form of treatment in the IPF patient population (3.1%), and many patients receiving ECT filled evidence-based medications during the follow-up period. However, given that ECT may be used as an alternative when patients fail pharmacotherapy and that the medical record review showed that patients receiving ECT did not always receive an evidence-based prescription, the TEP and workgroup recommended the exclusion from the denominator of patients receiving ECT during the index admission or follow-up period.

2. **TMS**

   TMS is a newer procedure and is still rare. Many patients receiving TMS also filled evidence-based medications during the follow-up period. However, since TMS may be used as an alternative when patients fail pharmacotherapy, the TEP and workgroup recommended the exclusion of patients receiving TMS during the index admission or follow-up period from the denominator.

3. **Pregnancy**

   Pregnancy was rare in this patient population (0.1%). The results showed that pregnant patients had empirically lower rates of filling evidence-based medications within 30 days of discharge than patients who were not pregnant (60.4% compared to 74.5%), which supports the TEP and workgroup recommendations to exclude from the denominator. Therefore, we excluded pregnant patients from the measure.

4. Secondary diagnosis of delirium

Patients with secondary diagnoses of delirium are rare (2.0%). The results showed that patients with delirium had empirically lower rates of filling evidence-based medications within 30 days of discharge than patients without delirium (70.3% compared to 74.5%), which supports the TEP and workgroup recommendations to exclude from the denominator. Therefore, we excluded patients with delirium from the measure.

5. Principal diagnosis of schizophrenia with secondary diagnosis of dementia

Patients with schizophrenia and secondary diagnoses of dementia were rare (3.2%). The results showed that patients with schizophrenia and a secondary diagnosis of dementia had empirically lower rates of filling evidence-based medications within 30 days of discharge than patients without dementia (65.3% compared to 75.9%), which supports the TEP and workgroup recommendations to exclude from the denominator. Therefore, we excluded patients with schizophrenia and a secondary diagnosis of dementia from the measure.

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with risk factors**

☐ **Stratification by risk categories**

☐ **Other,**

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Not applicable.

Not applicable because the measure is not risk adjusted.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable.

Not applicable because this is a process measure.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

Not applicable

Not applicable because this measure is not risk adjusted.

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

Not applicable.

Not applicable because this measure is not risk adjusted.

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

Not applicable.

Not applicable because this measure is not risk adjusted.

2b3.5. 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*)

Not applicable.

Not applicable because this measure is not risk adjusted or stratified.

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Not applicable.

Not applicable because this measure is not risk adjusted.

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Not applicable.

Not applicable because this measure is not risk adjusted.

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Not applicable.

Not applicable because this measure is not risk adjusted.

2b3.9. Results of Risk Stratification Analysis:

Not applicable.

Not applicable because this measure is not stratified.

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

Not applicable.

Not applicable because this measure is not risk adjusted or stratified.

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not applicable.

Not applicable because this measure is not risk adjusted or stratified.

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. 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 related to performance gap in 1b)*

To examine differences in performance, we calculated measure rates across 1,066 facilities with at least 75 discharges within the performance period. We excluded facilities with <75 discharges because estimates for facilities with fewer cases are less reliable. We computed a confidence interval for each facility's rate and if it did not contain the mean Medication Continuation rate across all facilities, the facility was identified as better or worse than average.

To evaluate whether there is currently a performance gap and variation in performance across facilities, we applied all inclusion and exclusion criteria to calculate facility-level measure scores. We observed the distribution of medication continuation rates and the difference between IPFs in the 90th percentile of performance and IPFs in the 10th percentile. To identify statistically significant differences in performance, we

calculated 95% confidence intervals (95% CI) around the measure scores for each IPF and compared the 95% CI to the national medication continuation rate across all IPFs. If the confidence intervals did not overlap with the national medication continuation rate, the difference was considered statistically significant.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *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*)

Based on 1,066 facilities with at least 75 discharges, the Medication Continuation measure rates in our sample ranged from 34.8% to 94.3% (with a median of 76.2%). Fifty percent of facilities fell within the interquartile range of 70.1% and 81.9%. Thus, there is substantial variation in measure scores across facilities.

Table 2b4.2-A. Distribution of the Medication Continuation measure rates

Measure	Number of facilities	Mean rate	Min	10th Pct.	25th Pct.	Median	75th Pct.	90th Pct.	MAX	IQR
Facilities with ≥ 75 discharges	1,066	75.1%	34.8%	63.4%	70.1%	76.2%	81.9%	84.7%	94.3%	0.118
All facilities	1,680	75.0%	0.0%	61.8%	70.0%	76.8%	82.6%	87.5%	100.00%	0.126

Source: Mathematica analysis of Medicare fee-for-service (FFS) data for the July 1, 2017– June 30, 2019 performance period.

Of the 1,066 facilities, 21% (N=228) were statistically significantly worse than average and 27% (N=283) were better than average.

Table 2b4.2-B. Performance distribution of facilities on the FAPH measure relative to the sample average

Performance group	N and % of facilities	Mean performance rate
Worse than the national rate	228 (21%)	63%
No different than the national rate	555 (52%)	75%
Better than the national rate	283 (27%)	84%
All IPFs	1,066	75%

Source: Mathematica analysis of Medicare fee-for-service (FFS) data for the July 1, 2017–June 30, 2019, performance period.

Note: Facilities were determined as having statistically worse or better than average performance if the 95% confidence interval for each facility's measure rate did not include the national mean rate. Percentages are rounded off to the nearest whole integer.

An analysis of 2013-2014 Medicare claims data indicated performance varied between high- and low-performing facilities across more than 1,600 IPFs for each of the three diagnoses (Table 2b5.2-A). For the combined measure score, there is about a 22 percentage point difference between the 10th and 90th percentiles (66.7%–88.3%) and a median score of 79.6%.

Table 2b5.2-A. Distribution of Facility Performance

Diagnosis	# IPFs	Mean	SD	Min	10th Pctl	Lower Quartile	Median	Upper Quartile	90th Pctl	Max
MDD	1,651	75.5	13.9	0.0	60.0	69.6	77.1	83.3	89.7	100.0
Schizophrenia	1,655	79.1	15.3	0.0	63.6	73.1	81.5	87.9	95.5	100.0
Bipolar disorder	1,658	78.3	14.4	0.0	63.9	72.5	80.0	86.4	93.5	100.0
Overall	1,694	78.0	11.1	0.0	66.7	73.6	79.6	84.4	88.3	100.0

About 24% of facilities had medication continuation rates that were statistically better than the national rate, and about 13% of facilities had medication continuation rates that were statistically worse than the national rate (Table 2b5.2-B).

Table 2b5.2-B. Distribution of IPFs Compared to the National Medication Continuation Rate

Performance Categorization	Count IPFs	Percent IPFs
Total IPFs	1,694	100.0
Better than national rate	399	23.6
No different than national rate	572	33.8
Worse than national rate	213	12.6
Fewer than 75 discharges during the performance period	510	30.1

2b4.3. What is 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? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

There was substantial variability in measure rates across facilities. The measure was also able to detect facilities with better and worse than average performance. We computed the average Medication Continuation score for all facilities in a sample as well as a 95% confidence interval for each facility's score on the measure. If confidence intervals did not contain the average Medication Continuation score, the facility was identified as better or worse than average.

The results indicate ample room for improvement and meaningful differences in the quality of care between the highest and lowest performing facilities.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

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 eMeasures). 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.

2b5.1. 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; what statistical analysis was used*)

Not applicable.

Not applicable because there is only one set of specifications.

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

Not applicable.

Not applicable because there is only one set of specifications.

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

Not applicable.

Not applicable because there is only one set of specifications.

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to 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 (*describe the steps—do not just name a method; what statistical analysis was used*)

During measure testing, we did not find any cases of missing or unreliable data. The measure uses processed claims, and we do not expect missing or unreliable data to be an issue.

Not applicable because this measure is based on claims data.

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

We did not find any discharges with missing data.

Not applicable because this measure is based on claims data.

2b6.3. What is 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**?** (i.*e., 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, provide rationale for the selected approach for missing data*)

Missing data are not a problem given that the measure uses processed claims.

Not applicable because this measure is based on claims data.

3. Feasibility

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.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

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-9 codes on claims)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*)
Update this field for **maintenance of endorsement**.

ALL data elements are in defined fields in electronic claims

3b.2. 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 other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. 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.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

During measure testing, we found no cases of missing or unreliable data. The measure uses processed claims, and we do not expect missing or unreliable data to be an issue.

3c.2. Describe 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)*.

None

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. 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 not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-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 performance improvement.

Specific Plan for Use	Current Use (for current use provide URL)
Public Reporting Payment Program	*

*cell intentionally left blank

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

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

CMS, the measure's sponsor, has included the measure for use in the IPFQR program, a national pay-for-reporting program with publicly reported results at the facility level, for the first time for FY 2021.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

Not applicable

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*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.)

Not applicable

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

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.

IPFs nationwide will receive their measure scores, as well as mean state and national scores, via CMS's IPFQR program preview period this fall, and facility-level results will be publicly reported in 2021. CMS plans to monitor stakeholder feedback.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

CMS will supply IPFs with their measure scores in fall 2020 via a Microsoft Excel workbook that will provide detailed information on all discharges included in an IPF's measure score. CMS will release a user guide for the IPF report that explains all data provided to IPFs, and CMS plans to hold an on-demand webinar that will also explain all data provided.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

The measure is new to the IPFQR Program, IPFs have not yet received scores, and CMS plans to monitor stakeholder feedback going forward.

4a2.2.2. Summarize the feedback obtained from those being measured.

Not applicable

4a2.2.3. Summarize the feedback obtained from other users

Not applicable

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Not applicable

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b 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, what are the reasons? 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.

As we note in Section 1b.1, measure rates for the performance period from July 1, 2017, through June 20, 2019, ranged from 34.8 to 94.3%, with a median of 76.2%. There was a 21.3 percentage point difference between the 10th and 90th percentiles (63.4%–84.7%). Using 2013–2014 Medicare claims data, there was a 21.6 percentage point difference between the 10th and 90th percentiles (66.7%–88.3%) and a median score of 79.6%.

This measure is new and is being implemented for the first time for FY 2021. By calculating the facility-level medication continuation scores in Medicare FFS claims data and providing them to facilities, CMS aims to encourage quality improvement, specifically relating to stronger care transitions to outpatient settings.

4b2. Unintended Consequences

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).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

This measure is being implemented for the first time for FY 2021.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

This measure is being implemented for the first time for FY 2021.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

1879 : Adherence to Antipsychotic Medications for Individuals with Schizophrenia

1880 : Adherence to Mood Stabilizers for Individuals with Bipolar I Disorder

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

Antidepressant Medication Management (AMM) from the National Committee for Quality Assurance´s (NCQA) Healthcare Effectiveness Data and Information Set (HEDIS) 2019

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):

Are the measure specifications harmonized to the extent possible?

Yes

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

The numerator for the Medication Continuation measure has been harmonized with these measures when possible because the measure populations of the three related measures overlap with the patient population targeted by this measure and the measures share a similar clinical focus on medication use. We compared the medications included in the related measures with medications included in the Medication Continuation measure.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
The related measures that we identified are not competing measures because the Medication Continuation measure is for those with diagnoses of bipolar disorder, MDD, or schizophrenia.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

No appendix **Attachment:**

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare & Medicaid Services

Co.2 Point of Contact: Helen, Dollar-Maples, Helen.Dollar-Maples@cms.hhs.gov, 410-786-7214-

Co.3 Measure Developer if different from Measure Steward: Mathematica

Co.4 Point of Contact: Jason, Smoot, jsmoot@mathematica-mpr.com, 734-205-3109-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.

During initial measure development, the following groups provided input on design of the measure denominator, exclusions, and list of medications.

Inpatient Psychiatric Facility (IPF) Outcome and Process Measure Development and Maintenance Technical Expert Panel (TEP):

Alisa Busch, MD, MS

Director, Integration of Clinical Measurement & Health Services Research

Chief, Health Services Research Division, Partners Psychiatry and Mental Health

Assistant Professor of Psychiatry and Health Policy, Harvard Medical School

Kathleen Delaney, PhD, PMH-NP, RN

Professor, Rush College of Nursing

Jonathan Delman, PhD, JD, MPH

Assistant Research Professor, Systems and Psychosocial Advances Research Center, University of Massachusetts Medical School

Frank Ghinassi, PhD, ABPP

Vice President, Quality and Performance Measurement, Western Psychiatric Institute and Clinic

Associate Professor in Psychiatry, University of Pittsburg

Eric Goplerud, PhD

Senior Vice President, Director of Public Health Department, NORC at the University of Chicago

Geetha Jayaram, MD

Associate Professor, Schools of Medicine, Health Policy and Management and the Armstrong Institute for Patient Safety, Johns Hopkins University

Charlotte Kauffman, MA, LCPC

Service Systems Coordinator, State of Illinois-Division of Mental Health

Tracy Lenzini, BS

Executive Director, Grand Traverse Health Advocates

Kathleen McCann, RN, PhD

Director of Quality and Regulatory Affairs, National Association of Psychiatric Health Systems

Gayle Olano-Hurt, MPH, CPHQ, PMC

Director Data Management, Outcomes Measurement & Research Administration, Sheppard Pratt Health System

Mark Olfson, MD, MPH

Professor of Psychiatry, Columbia University Medical Center Department of Psychiatry; New York State Psychiatric Institute

Irene Ortiz, MD, MSW

Medical Director, Molina Healthcare of New Mexico

Thomas Penders, MS, MD, DLFAPA

Medical Director, Inpatient Psychiatry, Vident Medical Center

Associate Professor, Brody School of Medicine Department of Psychiatry, East Carolina University

Lucille Schacht, PhD

Senior Director, Performance and Quality Improvement, National Association of State Mental Health Program Directors Research Institute, Inc.

Lisa Shea, MD

Medical Director, Butler Hospital

Thomedi Ventura, MS, MSPH

Program Evaluator, Telligen

Elvira Ryan, MBA, BSN, RN

Associate Project Director, Division of Healthcare Quality Evaluation, The Joint Commission

Measure work group members:

TEP members:

Frank Ghinassi, PhD

Geetha Jayaram, MD

Charlotte Kauffman, MA

Kathleen McCann, PhD, RN

Gayle Olano-Hurt, MPH

Thomedi Ventura, MSPH

UF members:

Regina Bussing, MD

Professor and Chair, Department of Psychiatry, University of Florida College of Medicine

Mathew Nguyen, MD

Assistant Professor and Medical Director, Department of Psychiatry, University of Florida College of Medicine

Gary Reisfield, MD

Associate Professor, Department of Psychiatry, University of Florida College of Medicine

Almut Winterstein, PhD, RPh, FISPE

Professor and Chair, Pharmaceutical Outcomes and Policy, University of Florida College of Medicine

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2020

Ad.3 Month and Year of most recent revision:

Ad.4 What is your frequency for review/update of this measure? CMS plans to review and update this measure annually.

Ad.5 When is the next scheduled review/update for this measure? 2021

Ad.6 Copyright statement: None

Ad.7 Disclaimers: This performance measure is not a clinical guideline, does not establish a standard of medical care, and has not been tested for all potential applications. The measure and specifications are provided without warranty.

The measure specifications also contain limited proprietary coding. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets.

Ad.8 Additional Information/Comments: None.