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

Brief Measure Information

NQF #: 0560

Corresponding Measures:

Measure Title: HBIPS-5 Patients discharged on multiple antipsychotic medications with appropriate justification

Measure Steward: The Joint Commission

Brief Description of Measure: The proportion of patients, age greater than and equal to 1 year, discharged from a hospital-based inpatient psychiatric setting on two or more antipsychotic medications with appropriate justification.

Developer Rationale: Research studies have found that 4-35% of outpatients and 30-50% of inpatients treated with an antipsychotic medication concurrently received 2 or more antipsychotics (Covell, Jackson, Evans, & Essock, 2002; Ganguly, Kotzan, Miller, Kennedy, & Martin, 2004; Gilmer, Dolder, Folsom, Mastin, & Jeste, 2007; Kreyenbuhl, Valenstein, McCarthy, Ganocyz, & Blow, 2006; Stahl & Grady, 2004). One study reported 4.6% of patients concurrently received 3 or more antipsychotics (Jaffe & Levine, 2003). These findings are seen across diverse sectors: state mental health authorities, the Veterans Health System and Medicaid-financed care. Antipsychotic polypharmacy can lead to greater side effects, often without improving clinical outcomes (Ananth, Parameswaran, & Gunatilake, 2004; Stahl & Grady, 2004). As a result, a range of stakeholders have called for efforts to reduce unnecessary use of multiple antipsychotics (Centorrino, Gören, Hennen, Salvatore, Kelleher, & Baldessarini, 2004; Gilmer, Dolder, Folsom, Mastin, & Jeste, 2007; National Association of State Mental Health Program Directors, 2001; University HealthSystem Consortium, 2006). Practice guidelines recommend the use of a second antipsychotic only after multiple trials of a single antipsychotic have proven inadequate (American Psychiatric Association [APA] Practice Guidelines, 2004). Randomized controlled trials (RCTs) provide some evidence to support augmentation with a second antipsychotic in treatment resistant patients. Most of these studies were limited to augmentation of clozapine with another second-generation antipsychotic (Tranulis, Skalli, Lalonde, & Nicole, 2008). Among patients without a documented history of previous treatment failures of antipsychotic monotherapy, multiple RCTs and other controlled trials failed to show a benefit of antipsychotic polypharmacy over monotherapy (Ananth, Parameswaran, & Gunatilake, 2004; Centorrino, Gören, Hennen, Salvatore, Kelleher, & Baldessarini, 2004; Potkin, Thyrum, Alva, Bera, Yeh, & Arvanitis, 2002; Shim et al., 2007; Stahl,& Grady, 2004). Clinical circumstances, such as shorter inpatient stays, may require hospitals to discharge a patient on multiple antipsychotics with an aftercare plan to transition to monotherapy. In such cases, effective communication between the inpatient and aftercare clinician is an essential element of care.

As stated above, recent literature supports three appropriate justifications for prescribing multiple antipsychotic medications: previous failed trials of monotherapy, cross-tapering to monotherapy and

augmentation of clozapine. A review of the justifications for prescribing more than one antipsychotic medication will help hospitals determine if their practice is supported by the evidence-base.

The measure will assist health care organizations (HCOs) to track the number of patients prescribed two or more antipsychotic medications at the time of discharge with appropriate justification.

Numerator Statement: Psychiatric inpatients discharged on two or more routinely scheduled antipsychotic medications with appropriate justification.

Denominator Statement: Psychiatric inpatient discharges

Denominator Exclusions: • Patients who expired

• Patients with an unplanned departure resulting in discharge due to elopement

• Patients with an unplanned departure resulting in discharge due to failing to return from leave

• Patients with a length of stay less than or equal to 3 days

Measure Type: Process

Data Source: Electronic Health Records, Paper Medical Records

Level of Analysis: Facility, Other

Original Endorsement Date: Aug 05, 2009 **Most Recent Endorsement Date:** Feb 28, 2014

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**

Evidence Summary or **Summary of prior review in 2014**

- A single systematic review (SR) from 2004 by the APA is detailed along with several other studies published as recently as 2012. A meta-analysis that is presumed to be part of the SR coalesces information from at least 9 randomized control trials support the conclusion that "the reduction in

antipsychotic polypharmacy without an appropriate justification…" as highly credible. The meta-analysis further concludes that polypharmacy increases patient risk of sudden cardiac death, metabolic illnesses, urogentital illnesses, and neurologic sensory syndromes. The quality consistency and quantitity of the evidence are all described as high. The quoted guidelines focus on intitial trial durations of 4-6 weeks and on the used of clozapine if two other second generation antipsychotics fail to demonstrate efficacy.

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: More recent literature was survey which did not counter their original conclusions.

Exception to evidence

N/A

Questions for the Committee:

- *Is the committee satisfied that the 2004 guidelines along with the more recent literature support the connection between this measure and better heatlh outcomes for persons prescribed antipsychotics?*

Guidance from the Evidence Algorithm

(Not outcome)→ Box 3 (QQC High) → 5 (High/Moderate) see below:

- Systematic review is somewhat dated, and the connection to the meta-analysis and completeness of that meta-analysis is unclear from the discourse presented in the table under Evidence section 1a.3. Also note that all of the evidence is not systematically graded for the reader to discern directly.

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.

- Rates below are, across years, well below 1 with marked standard deviations. Rates are presumed to be percent (e.g., .29737 = 29.737% performance rate on the measure). Percentailes also are provided by developer, but not displayed below.

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
N	296	316	463	477	505	651	1015	1035	729	688
Mean	0.297	0.401	0.400	0.484	0.509	0.537	0.570	0.587	0.609	0.619 (truncated to 3 digits)
StDV.	0.26	0.33	0.32	0.33	0.33	0.33	0.33	0.32	0.31	0.33 (truncated to 2 digits)

Disparities

- No disparities on the measure were evident in the literature, per the developer, and the developer also provided the following tabulations (at the provider level using the same samples as the table above):

Gender	2013	2014	2015	2016	2017
Male	0.541	0.577	0.620	0.617	0.618
Female	0.540	0.591	0.621	0.601	0.623

Hispanic Ethnicity	2013	2014	2015	2016	2017
Hispanic	0.647	0.585	0.658	0.625	0.643
Non-Hispanic	0.531	0.583	0.618	0.610	0.619

Race	2013	2014	2015	2016	2017
White	0.522	0.565	0.608	0.602	0.618
African American	0.541	0.581	0.634	0.636	0.647
American Indian	0.434	0.506	0.534	0.441	0.433
Asian	0.536	0.604	0.683	0.626	0.621
Pacific Islander	0.539	0.390	0.611	0.530	0.644

Age Category	2013	2014	2015	2016	2017
1-12 years	0.575	0.562	0.584	0.628	0.434
13-17 years	0.505	0.522	0.590	0.588	0.509
18-64 years	0.537	0.569	0.631	0.620	0.632
65+ years	0.463	0.512	0.563	0.561	0.594

- The above tabulations do suggest a few disparaties, some of the largest of which are highlighted.

Questions for the Committee:

none

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

Comments:

**The evidence on the process measure presented in the previous review was judeged to be high in quantitiy, quality and consistency. Developers report that their updated literature review did not yield any new guidelines or significant research related to antipsychotic medications that would warrant a change in the measure. They also report that an updated guideline for the treatment of patients with Schizophrenia is under development by the APA. The current guideline dates from 2004, and it would be useful to have coroboration from this more recent source. When is it expected to be released?

**Evidence is high that this measure is empirically based

**I'm not completely satisfied that the 2004 APA guidelines with the more recent literature necessarily supports improved outcomes for all patients prescribed antipsychotics. while it's highly desirable to avoid

polypharmacy, there may be times where it makes clinical sense. for example, low dose aripiprazole to counteract elevated prolactin with risperdal. this specific example is not included in the 'three appropriate justifications' (failed trials of monotherapy, cross tapering to monotherapy, and augmentation of clozapine). other uncommon examples exist.

**Literature and practice guidelines support the measure. The evidence is dated but still applicable. One question I have is what the overall prevalence of multiple antipsychotic prescribing is at the current time. No new studies looking at the percentage of patients being discharged on multiple antipsychotics was presented.

1b. Performance Gap

Comments:

**The performance data presented from 2009 - 2018 demonstrate considerable improvement with an increase on average from .297 to .619 during this period which also demonstrates that there still remains room for improvement. There was substantial variation in the number of reporting entities over this time frame, ranging from 296 in 2009, increasing to a high of 1035 in 2016, and dropping back to 688 in 2018. What is the reason for this variation? Does the number vary depending on how many facilities participate. If so, what is the number of non-reporting facilities each year? Data is presented by gender, ethnicity, race and age for the period 2013 - 2017. There is considerable variation among age groups. Those 1 - 12 years actually have declined from .575 to .434 during this period. Those 13 - 17 years remained about the same. Adults 18 - 64 and 65+ years showed improvement over this time period.

**No disparities; current performance data provided; populations subgroups divided though unclear numbers that fall into different demographics/ages

**No concerns noted.

**There is variability in performance. However, since HBIPS-4 was discontinued, it is difficult to ascertain how problematic this issue remains. It would also be interesting to examine the percentage of patients admitted on a multiple antipsychotic regiment to get an idea of the current community prevalence of multiple antipsychotic prescribing.

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.

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

Evaluator: NQF Staff

Summary of Reliability and Validity testing and results

- Data element level reliability test-retest assessment was conducted. Testing included a review of 191 records and showed apparent perfect agreement for 5 numerator observations and 191 observations for denominator elements.
- Validity involved correlation analysis to other HBIPS measures (score level), and feedback from hospitals using a Likert scale to rate each of the data elements.

Questions for the Committee regarding reliability:

- Were enough numerator records tested?
- Is apparent perfect agreement credible, or was their testing too simplistic?

Questions for the Committee regarding validity:

- Was the slight correlation with a separate HBIPS screening measure enough to validate this as a quality measure?

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

Evaluation A: Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0560

Measure Title: HBIPS-5 Patients discharged on multiple antipsychotic medications with appropriate justification

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

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 <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☐ **Yes** ☐ **No** **n/a**

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

   Submission document: Testing attachment, section 2a2.2

   - Test-retest ('abstract- re-abstract' in this case) of the chart abstraction to get data elements, including the denominator and the information listed below in the results.
   - All sampled cases were re-abstracted by trained Joint Commission staff. Re-abstracted data are compared with originally abstracted data on a data element by data element basis. Agreement rates for individual data elements appear in the table below:

Data Elements	Total Numerator	Total Denominator	Agreement Rate
Numerator Data Element			
Appropriate justification for multiple antipsychotic medications	5	5	100%
Denominator Data Elements			
Admission Date	191	191	100%
Birthdate	191	191	100%
Discharge Date	191	191	100%
Discharge Disposition	191	191	100%
ICD-9-CM Other Diagnosis Codes*	191	191	100%
ICD-9-CM Principal Diagnosis Code*	191	191	100%
Number of Antipsychotic Medications Prescribed at Discharge	191	191	100%
Patient Status at Discharge	191	191	100%
Psychiatric Care Setting	191	191	100%

** The measure was tested with ICD-9-CM codes. A crosswalk from ICD-9-CM diagnosis codes to ICD-10-CM diagnosis codes was done and reviewed by the Technical Advisory Panel. The panel determined that the intent of the measure was not changed as a result of the conversion."*

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - Certainly it is the case that test-retest using EHRs demonstrates reliability. However, was tested for the numerator (justication was provided for multiple AP use) only 5 times. Additionally, the following question remains: is it really so easy to perfectly reproduce this measure? Is it because the measure is too simplistic? Are there no nuances regarding reproducibility which need to be tested, perhaps with test-retest of translating clinical notes into indicators about whether poly-AP use was properly justified?

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)

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.**

    - Committee should discuss, and the developer should defend, that the EHR test-retest (on only 5 observations for the numerator) is sufficient.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    Here is the developer's report of the magnitude of the excluded populations:

    "2017 Discharges N= 520,778

    - Patients who expired= 0.07%
    - Patients with an unplanned departure resulting in discharge due to elopement or failing to return from leave= 2.2%
    - Patients with a length of stay equal to or less than 3 days= 16.0%
    - Number of antipsychotic medications prescribed at discharge ≤ 1= 76.6%"

    One concern is the third bullet above: 16% are short stays which are excluded. The concern here is that these patients may be discharged with multiple antipsychotics absent proper justification. That might well be an event worth including in the quality measure. This concern would be assuaged greatly if the developer could at least show that APs are never or rarely prescribed to those who have such a short stay. The developer notes that the reason for this exclusion is that a short length of stay does not allow time for evaluation of a patient's response to medication changes. Cases in which the length of stay is ≤ 3 days are excluded up front and the number of prescribed antipsychotic medications is not abstracted.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.

    None.

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.

- N/A per the developer, even as they get data from multiple hospitals and thus multiple EHRs, the records are said to be standardized by their interaction with those hosptials.

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

Submission document: Testing attachment, section 2b6.

- Only that 16% removal of those with <3 day inpatient event.

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 (see information under 1b)

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

16d.**Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

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

- Survey feedback on face validity was tendered from 40 hospitals, supplemented by site visits to three (all in 2006). Respondents used a 5-point Likert scale to consider numerator, denominator and exclusions to both.
- Correlations with other HBIPS measures were conducted, but a priori hypotheses were not well articulated in the application. Presumably the following correlations where anticipated:

HBIPS-5 (The current measure of documented poly-antipsychtic use) should be:

…directly correlated to: HBIPS-1 (Admission screening…)

…inversely correlate to each of: HBIPS-2 (hours of restraint); HBIPS-3 (hours of seclusion)

Both these methods seem reasonable.

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

    Submission document: Testing attachment, section 2b2.3

    "Tests for correlations between HBIPS-5 and the remaining HBIPS measures (HBIPS-1, HBIPS-2, HBIPS-3) are 0.13857(p=0.0002), -0.04720 (p=0.2068), and -0.04642 (p=0.2164), respectively. This indicates that there are no statistically significant correlations between HBIPS-5 and HBIPS-2 and HBIPS-3. There is a slight positive correlation between HBIPS-5 and HBIPS-1."

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.**

    - A higher rating could have been achieved with better external standards (besides other HBIPS), and with a better explanation of *a priori* hypotheses and how/how not they were fulfilled.

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.**

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

2a1. Reliability – Specifications
Comments:
**All data elements are clearly defined and have accompanying codes. The calculation algorithms, and other specifications, such as exclusions are clearly defined. No concerns about the liklihood that this measure can be consistently implemented.
**Codes provided; data elements clearly defined; no concerns about this being consistently implemented
**Not my area of expertise; but no concerns noted.
**The specifications are clear; the exclusions seem reasonable. It would be interesting to see what the prevalence of multiple antipsychotic prescribing is for patients discharged in less than 3 days.

2a2. Reliability – Testing
Comments:
**No
**No
**Concerns about only 5 observations for pts discharged on >1 antipsychotic with appropriate justification is sufficient.
**The reliability testing was 100%. However the numerator was 5 out of 191 cases (2.6%). This brings up the issue as to how much of a concern this issue remains.

2b1. Validity –Testing
Comments:
**No
**No
**Not my area of expertise;
**16% of the sample excluded due to discharge < 3 days. It would be helpful to understand the rate of multiple antipsychotic prescribing in this subset. Additionally the correlations with the other HBIPS data set did not seem compelling in demonstrating validity.

2b2-3. Meaningful Differences
Comments:
**2b2. The exclusions are not evidence-based, but represent a reasonable approach. No risk-adjustment was performed.
**Social risk factors weren't clearly articulated. Risk adjustment strategy is not well articulated.
**See above. Also, general concern that this is a process measure based off of 2004 APA guidelines, and not clear that this necessarily correlates with improved outcomes.
**N/A

2b4-7. Threats to Validity
Comments:
**2b6. See previous comment about the possibiliity of missing facilities. There does not appear to be missing data from the facilities that did report.
**This was noted already but does present a concern - 16% removal of those with <3 day inpatient is concerning event
**Concerns about those with length of stay of 3 days or less are excluded (16%). if the goal is to decrease polypharmacy, should this subset not be excluded...
**Differences are demonstrated.

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.

- Information is derived from EHRs, and seems to be standardized.
- Data is said to be typically sourced from paper medical records that are then transferred to e-records.

Questions for the Committee:

None.

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

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility
Comments:

**Electronic Health Records and Paper Medical Records can be used. The measure is not specified as an eMeasure, and the developer stated that it would be difficult abd resource intensive to do so.

**Appears feasible; required data is available electronically; no concerns about data collection strategy

**No concerns; info from paper and electronic records.

**Feasibility has been demonstrated

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**

Accountability program details

Public Reporting - ORYX Performance Measurement Reporting Program, https://www.qualitycheck.org/
Payment Program - Inpatient Psychiatric Facility Quality Reporting (IPFQR) Program, https://www.medicare.gov/hospitalcompare/search.html

Regulatory and Accreditation Programs -Hospital Accreditation Program, http://jointcommission.org
Quality Improvement (external benchmarking to organizations) - America's Hospitals: Improving Quality and Safety – The Joint Commission's Annual Report 2017, https://www.jointcommission.org/annualreport.aspx
Quality Improvement (Internal to the specific organization) - ORYX Performance Measurement Report

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

- The Joint Commission utilizes an automated feedback system to track and consider all feedback.
    - Modifications to this measure have not been required based on feedback received.
- Queries submitted via the automated feedback system have decreased significantly for the HBIPS measure set in the past 3 years. (522 in 2016, 288 in 2017, 187 for 2018 YTD).
- There have been no issues with the data elements for this measure and no updates needed to the data element specifications based upon feedback received.
- All measure specifications are reviewed twice a year and updates are made as needed based on feedback from the measure users, input from the TAP, or changes in the guidelines.

Additional Feedback: n/a

***Questions for the Committee**:*

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

- "Though 2009 to 2nd quarter 2018, a binomial random effects model was used to determine if there was a change in rates over time with time as a fixed effect and healthcare organization as a random effect. The results of the model show statistical significant over time ($P<0.001$) and an odd ratio estimate of time to be 1.302."
- This description is interpreted as a 30% increase in performance, after adjusting for individual healthcare organization effects, over the 10 year period studied-- i.e., ~3% per year improvement.
- The work of Rsinksi et al., (2018) showed improvement (details not given) in 4 separate hospital discharge cohorts.
- The developer states that the comparison of rates over cohorts was also adjusted for hospital covariates. In contrast, the trend analysis of the national data did not look within cohort, nor was it adjusted for hospital covariates. Therefore the unexpected increase in measure rates over time observed with the national data could be due to cohort and covariate effects and changes in the mix of hospitals reporting over time.

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 none

Potential harms none

Additional Feedback: N/A

Questions for the Committee:

Preliminary rating for Usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

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

4a1. Use - Accountability and Transparency

Comments:

**The measure is used in public reporting and is part of the Inpatient Psychiatric Facility Quality Reporting Program sponsored by the Centers for Medicare and Medicaid Services. Measures are reported to the hospitel on a quarterly basis. It was acknowledged that almost 15% of facilities reporting this measure deomonstrate unfavorable performance. Although overall there was considerable improvement over time, there is a concern about the poor scoring on thei subset of facilities. It is not clear if feedback on improvement is being given to underperfoming facilities.

**Yes

**Currently in use and publicly reported.

**The measure is used for public reporting by CMS. The measure developer notes that the measure specifications have remained stable. The developer notes decreased feedback on the Joint Commission website over the past few years which is taken as indicating that the measure has been deemed acceptable. The developer does not specify how many of the comments were related directly to HBIPS-5.

4b1. Usability – Improvement

Comments:

**This measure tracks an important quality dimension. It seems to be operating reasonably well.

**Seems the benefits of the measure outweigh the harms. doesnt appear there are unintended consequences. ensuring that the hospitals do a stronger job tracking those being discharged on more than one anti-psychotic med

**30% improvement noted over 10 year period. Potential unintended consequences: times where 2 antipsychotics without the listed appropriate justifications might be appropriate.

**The measure is usable. No unintended consequences were noted.

Criterion 5: Related and Competing Measures

Related or competing measures

- Developer did not identify any related or competing measures
- NQF staff identified the following related measures:
  - 1879: Adherence to Antipsychotic Medications for Individuals with Schizophrenia
  - 2801: Use of First-Line Psychosocial Care for Children and Adolescents on Antipsychotics
  - 3205: Medication Continuation Following Inpatient Psychiatric Discharge.

Harmonization
None.

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

5. Related and Competing
Comments

**NO

**None

**1879: Adherence to antipsychotics for individuals with schizophrenia, 2801: use of first line psychosocial care for children on antipsychotics; 3205: medication continuation following inpatient psychiatric discharge. (on the surface, these seem like they would be more highly associated with improved outcomes).

**N/A

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 06/17/2019

- There have been no public comments or support/non-support choices as of this date.

Brief Measure Information

NQF #: 0560

Corresponding Measures:

De.2. Measure Title: HBIPS-5 Patients discharged on multiple antipsychotic medications with appropriate justification

Co.1.1. Measure Steward: The Joint Commission

De.3. Brief Description of Measure: The proportion of patients, age greater than and equal to 1 year, discharged from a hospital-based inpatient psychiatric setting on two or more antipsychotic medications with appropriate justification.

1b.1. Developer Rationale: Research studies have found that 4-35% of outpatients and 30-50% of inpatients treated with an antipsychotic medication concurrently received 2 or more antipsychotics (Covell, Jackson, Evans, & Essock, 2002; Ganguly, Kotzan, Miller, Kennedy, & Martin, 2004; Gilmer, Dolder, Folsom, Mastin, & Jeste, 2007; Kreyenbuhl, Valenstein, McCarthy, Ganocyz, & Blow, 2006; Stahl & Grady, 2004). One study reported 4.6% of patients concurrently received 3 or more antipsychotics (Jaffe & Levine, 2003). These findings are seen across diverse sectors: state mental health authorities, the Veterans Health System and Medicaid-financed care. Antipsychotic polypharmacy can lead to greater side effects, often without improving clinical outcomes (Ananth, Parameswaran, & Gunatilake, 2004; Stahl & Grady, 2004). As a result, a range of stakeholders have called for efforts to reduce unnecessary use of multiple antipsychotics (Centorrino, Gören, Hennen, Salvatore, Kelleher, & Baldessarini, 2004; Gilmer, Dolder, Folsom, Mastin, & Jeste, 2007; National Association of State Mental Health Program Directors, 2001; University HealthSystem Consortium, 2006). Practice guidelines recommend the use of a second antipsychotic only after multiple trials of a single antipsychotic have proven inadequate (American Psychiatric Association [APA] Practice Guidelines, 2004). Randomized controlled trials (RCTs) provide some evidence to support augmentation with a second antipsychotic in treatment resistant patients. Most of these studies were limited to augmentation of clozapine with another second-generation antipsychotic (Tranulis, Skalli, Lalonde, & Nicole, 2008). Among patients without a documented history of previous treatment failures of antipsychotic monotherapy, multiple RCTs and other controlled trials failed to show a benefit of antipsychotic polypharmacy over monotherapy (Ananth, Parameswaran, & Gunatilake, 2004; Centorrino, Gören, Hennen, Salvatore, Kelleher, & Baldessarini, 2004; Potkin, Thyrum, Alva, Bera, Yeh, & Arvanitis, 2002; Shim et al., 2007; Stahl,& Grady, 2004). Clinical circumstances, such as shorter inpatient stays, may require hospitals to discharge a patient on multiple antipsychotics with an aftercare plan to transition to monotherapy. In such cases, effective communication between the inpatient and aftercare clinician is an essential element of care.

As stated above, recent literature supports three appropriate justifications for prescribing multiple antipsychotic medications: previous failed trials of monotherapy, cross-tapering to monotherapy and augmentation of clozapine. A review of the justifications for prescribing more than one antipsychotic medication will help hospitals determine if their practice is supported by the evidence-base.

The measure will assist health care organizations (HCOs) to track the number of patients prescribed two or more antipsychotic medications at the time of discharge with appropriate justification.

S.4. Numerator Statement: Psychiatric inpatients discharged on two or more routinely scheduled antipsychotic medications with appropriate justification.

S.6. Denominator Statement: Psychiatric inpatient discharges

S.8. Denominator Exclusions: • Patients who expired

• Patients with an unplanned departure resulting in discharge due to elopement

• Patients with an unplanned departure resulting in discharge due to failing to return from leave

• Patients with a length of stay less than or equal to 3 days

De.1. Measure Type: Process

S.17. Data Source: Electronic Health Records, Paper Medical Records

S.20. Level of Analysis: Facility, Other

IF Endorsement Maintenance – Original Endorsement Date: Aug 05, 2009 **Most Recent Endorsement Date:** Feb 28, 2014

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

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

0560_evidence_attachment_7.1_HBIPS5-636840321120368547.docx

1a.1 <u>For Maintenance of Endorsement:</u> 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*)**:** 0560

Measure Title: Patients discharged on multiple antipsychotic medications with appropriate justification

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 12/20/2018

Instructions

- ***Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.***
- ***Complete EITHER 1a.2, 1a.3 or 1a.4 as applicable for the type of measure and evidence.***
- ***For composite performance measures:***
  - ***A separate evidence form is required for each component measure unless several components were studied together.***
  - ***If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.***

- **All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.**
- **If you are unable to check a box, please highlight or shade the box for your response.**
- **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.**
- **Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.**
- **Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.**
- **Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.**
- **Efficiency: [6] evidence not required for the resource use component.**
- **For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.**
- **Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.**

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

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

Outcome

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the 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*): Click here to name the intermediate outcome

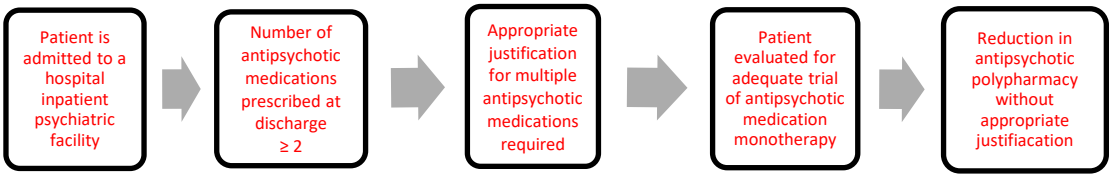
☒ Process: Click here to name what is being measured

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 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.



The focus of the measure is to evaluate all psychiatric inpatients who are discharged on two or more antipsychotic medications to determine if appropriate justification exists for this practice. A reduction in antipsychotic polypharmacy without an appropriate justification will reduce the likelihood of developing serious side effects, thus reducing the overall cost of ongoing health care.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

Not applicable

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

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

Not applicable

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) 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

Updated literature search did not yield any new guidelines or significant research related to antipsychotic medications that would warrant a change in the measure. An updated guideline for the Treatment of Patients with Schizophrenia is currently under development by the American Psychiatric Association.

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**Title: Practice Guideline for the Treatment of Patients with Schizophrenia Second Edition** **Author: American Psychiatric Association Work Group on Schizophrenia** **Date: February 2004** **Citation, including page number: American Psychiatric Association (APA). Practice guideline for the treatment of patients with schizophrenia. 2nd ed. Washington (DC): American Psychiatric Association (APA); 2004 Feb. 114 p. [1391 references]** **URL: https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/schizophrenia.pdf** **Rationale for Using this Guideline Over Others: APA began developing practice guidelines in 1991. The development process is detailed in a document available from the APA Department of Quality Improvement and Psychiatric Services, the "APA Guideline Development Process." Key features of this process include the following:** • **A comprehensive literature review** • **Development of evidence tables** • **Initial drafting of the guideline by a work group that included psychiatrists with clinical and research expertise in psychiatric evaluation** • **Production of multiple revised drafts with widespread review** • **Approval by the APA Assembly and Board of Trustees** • **Planned revisions at regular intervals**

	This guideline represents a synthesis of current scientific knowledge and rational clinical practice on the psychiatric evaluation of adults. It strives to be as free as possible of bias toward any theoretical approach.
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.	In assessing treatment resistance or partial response, it is important to carefully evaluate whether the patient has had an adequate trial of an antipsychotic medication, including whether the dose is adequate and whether the patient has been taking the medication as prescribed. An initial trial of 4–6 weeks generally is needed to determine if the patient will have any symptomatic response, and symptoms can continue to improve over 6 months or even longer periods of antipsychotic treatment [II]. Given clozapine's superior efficacy, a clozapine trial should be considered for a patient who has had no response or partial and suboptimal response to two trials of antipsychotic medication (including at least one second-generation agent) or for a patient with persistent suicidal ideation or behavior that has not responded to other treatments [I].
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	**Grading of Strength/Quality of the Body of Evidence.** Has the body of evidence been graded? No **System Used for Grading the Body of Evidence:** Other **If other, identify and describe the grading scale with definitions:** Although grading of the evidence was not determined during the systematic review, it was determined that the guideline developers accounted for a balanced representation of information, looked beyond one specialty group or discipline, and provided information that was accessible and met the requirements set out in the NQF criteria.
Provide all other grades and definitions from the evidence grading system	Not applicable
Grade assigned to the **recommendation** with definition of the grade	**If guideline recommendation graded, identify the entity that graded the evidence including balance**

	of representation and any disclosures regarding bias: American Psychiatric Association **Grade Assigned to the Recommendation:** I to II For definition, see below
Provide all other grades and definitions from the recommendation grading system	**System Used for Grading the Strength of Guideline Recommendation:** Other **If other, identify and describe the grading scale with definitions:** The system for grading the strength of the guidelines recommendations is as follows: • [I] Recommended with substantial clinical confidence. • [II] Recommended with moderate clinical confidence. • [III] May be recommended on the basis of individual circumstances.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Directness of Evidence to the Specified Measure This measure is consistent with the guidelines recommended by the American Psychiatric Association (APA) when assessing treatment resistance or partial response in patients prescribed antipsychotic medications. It is important to carefully evaluate whether the patient has had an adequate trial of antipsychotic medications, including an adequate dose and length of time for the trial. The focus of both the performance measure and the body of evidence supports the need for monitoring antipsychotic prescribing practice, and whether there is an appropriate justification for prescribing more than one antipsychotic medication. **Quantity:** During a meta-analysis of studies conducted from 1966 through December 2007 the following were identified: • Six RCTs comparing antipsychotic polypharmacy to monotherapy in samples with established treatment resistance to trials of a single antipsychotic were identified.

	• Three RCTs that compared antipsychotic polypharmacy to monotherapy in samples without established treatment resistance to a single antipsychotic were identified. • Nine noncontrolled observational trials comparing antipsychotic polypharmacy to monotherapy in samples with established treatment resistance to trials of a single antipsychotic were identified. • Six nonrandomized controlled trials that compared antipsychotic polypharmacy to monotherapy in samples without established treatment resistance to a single antipsychotic were identified. • Six noncontrolled observational studies that examined the relationship between antipsychotic polypharmacy and clinical outcomes in samples without documented treatment resistance to monotherapy were identified. **Quality:** The quality of evidence supporting the reduction in antipsychotic polypharmacy without an appropriate justification is high. RCTs have consistently reported no clear benefit in antipsychotic polypharmacy versus monotherapy for controlling symptoms without established treatment resistance to a single antipsychotic medication. As stated previously, the increased risk of sudden cardiac death has been noted with increased doses of antipsychotic medications. Pediatric exposure to multiple antipsychotic medications increased the odds of developing obesity/excessive weight gain (odds ratio [OR], 2.28), Type II diabetes (OR, 2.36) and dyslipidemia (OR, 5.26), cardiovascular conditions (OR, 2.70), digestive/urogenital problems and neurological/sensory symptoms. As noted above, the American Psychiatric Association has had guidelines in place since 1997 addressing appropriate antipsychotic use. One antipsychotic medication should be prescribed at a time for patients with psychotic disorders. For patients who do not respond to an adequate dose and duration of different trials of monotherapy, antipsychotic combination treatment may be considered with close clinical monitoring. Future trials evaluating long term safety and tolerability

	trials and comparisons of specific antipsychotic medication combinations are still required. **Summary of Controversy/Contradictory Evidence:** There is no documented evidence regarding controversy related to the three appropriate justifications for prescribing multiple antipsychotic medications: previous failed trials of monotherapy, cross-tapering to monotherapy and augmentation of clozapine. There is no empiric evidence supporting other justifications, i.e. addition of a second antipsychotic medication for sleep. **Based on the NQF descriptions for rating the evidence, what was the <u>developer's assessment</u> of the quantity, quality, and consistency of the body of evidence?** Quantity: High Quality: High Consistency**:** High
Estimates of benefit and consistency across studies	**Benefit:** The purpose of this measure is to evaluate the number of patients discharged on multiple antipsychotic medications to assist the clinician in determining whether there is an appropriate justification supporting the practice. The evidence shows that monitoring the justifications will lead to a change in prescribing practice leading to a reduction in the number of multiple antipsychotic medications prescribed, which will in turn decrease the chance of developing serious side effects and will ultimately result in substantial savings in health care costs. **Consistency:** The body of evidence consistently supports a reduction in antipsychotic polypharmacy without an appropriate justification. A minimum number of three trials of monotherapy at adequate doses and duration should be completed prior to initiation of more than one antipsychotic medication. Additionally, the evidence supports the use of a second antipsychotic medication to augment clozapine and a tapering plan to monotherapy as appropriate justifications for multiple antipsychotic medications. No position against the importance to

	reduce antipsychotic polypharmacy without an appropriate justification was identified in the literature.
What harms were identified?	No harms to the patient receiving justified multiple antipsychotic medications were found during the literature review.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	A review of recent studies also supports the use of quality improvement interventions to educate staff on these appropriate justifications which may further reduce antipsychotic polypharmacy. No position against reducing the number of antipsychotic medications prescribed without an appropriate justification was identified in the literature.

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.

Not applicable for this submission

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

Not applicable for this submission

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

Not applicable for this submission

From previous submission - Citations for Evidence other than Guidelines:

- American Psychiatric Association (APA). (2004). Steering Committee on Practice Guidelines. Practice guideline for the treatment of patients with schizophrenia, second edition. Am J Psychiatry. 161(2 Suppl):1-56
- Ananth, J., Parameswaran, S., & Gunatilake, S. (2004). Antipsychotic polypharmacy comparing monotherapy with polypharmacy and augmentation. Curr Med Chem. 11(3):313-327 Curr Pharm Des. 10(18):2231-2238.

- Baandrup, L., Sorensen, J., Lublin, H., Nordentoft, M. & Glenthoj, B. (2011). Association of antipsychotic polypharmacy with health service cost: a register-based cost analysis. Eur J Health Econ. Retrieved March 27, 2012 at: http://www.ncbi.nlm.nih.gov/pubmed.
- Barnes, R.E. & Paton, C. (2011). Antipsychotic polypharmacy in schizophrenia. CNS Drugs. 25(5):383-399.
- Busch, A.B., Lehman, A.F., Goldman, H. & Frank, R.G. (2009). Changes over time and disparities in schizophrenia treatment quality. Med Care. 47(2):199-207.
- Centorrino, F., Gören, J.L., Hennen, J., Salvatore, P., Kelleher, J.P., & Baldessarini, R.J. (2004) Multiple versus single antipsychotic agents for hospitalized psychiatric patients: a case control study of risk versus benefit. Am J Psychiatry. 161 (4):700-706.
- Covell, N.H., Jackson, C.T., Evans, A.C., & Essock, S.M. (2002). Antipsychotic prescribing practices in Connecticut's public mental health system: rates of changing medication prescribing styles. Schiz Bull. 28(1):17-29.
- Dolder, C.R. & McKinsey, J. (2010). Antipsychotic polypharmacy among patients admitted to a geriatric psychiatry unit. J Psychiatr Pract. 17(5):368-74.
- Finnerty, M.T., Kealey, E., Leckman-Westin, E., Gupta,N., White, G.M., Engel, G.M. & Opler, L.A. (2011). Best practices: long-term impact of web-based tools, leadership feedback, and policies on inpatient antipsychotic polypharmacy. Psychiatric Services. 62(10):1124-6.
- Ganguly, R., Kotzan, J.A., Miller, L.S., Kennedy, K., & Martin, B.C. (2004). Prevalence, trends, and factors associated with antipsychotic polypharmacy among Medicaid-eligible schizophrenia patients, 1998-2000. J Clin Psychiatry. 65(10):1377-88.
- Gilmer, T.P., Dolder, C.R., Folsom, D.P., Mastin, W., & Jeste, D.V. (2007), Antipsychotic polypharmacy trends among Medicaid beneficiaries with schizophrenia in San Diego County, 1999 - 2004. Psychiatric Serv. 59(7):1007-1010.
- Goren, J.L., Beck, S.E., Mills, B.J., Shtasel, D.L. & Dufresne, R.L. (2010). Development and delivery of a quality improvement program to reduce polytherapy. J Manag Care Pharm. 16(6):393-401.
- Jaffe, A.B. & Levine, J. (2003). Antipsychotic medication co-prescribing in a large state hospital system. Pharmacoepidemiol Drug Saf.12:41-48.
- Jerrell, J.M. & McIntyre, R.S. (2008). Metabolic and cardiovascular adverse events associated with antipsychotic treatment in children and adolescents. Pediatr Adolesc Med. 162(10):929-35.
- Joukamaa, M., Heliovaara, M., Knekt, P., Aromaa, A., Raitasalo, R. & Lehtinen, V. (2006). Schizophrenia, neuroleptic medication and mortality. Br J Psychiatry. 188:122-7.
- Kreyenbuhl, J., Valenstein, M., McCarthy, J.F., Ganocyz, D., & Blow, F.C. (2006). Long-term combination antipsychotic treatment in VA patients with schizophrenia. Schiz Res.84:90-99.
- National Association of State Mental Health Program Directors (NASMHPD). (2001).Technical report on psychiatric polypharmacy. Alexandria, VA.
- Ogbru, E.O. (2009). Patient provider and payer characteristics associated with antipsychotic drug polypharmacy in ambulatory care. Dissertation Abstracts International: section B: The Sciences and Engineering. 70(6-B):3359.
- Pandurangi, A.K. & Dalkilic, A. (2008). Polypharmacy with second-generation antipsychotics: a review of evidence. J Psychiatr Pract. 14(6):345-67.
- Potkin, S.G., Thyrum, P.T., Alva, G., Bera, R., Yeh, C., & Arvanitis, L.A. (2002). The safety and pharmacokinetics of quetiapine when coadministered with haloperidol, risperidone or thioridazine. J Clin Psychopharmacol. 22:121-130.

• Shim, J.C., Shin, J.G., Kelly, D.L., Jung, D.U., Seo, Y.S., Liu, K.H., et al. (2007). Adjunctive treatment with a dopamine partial agonist aripiprazole, for treatment of antipsychotic-induced hyperprolactinemia: A placebo controlled trial. Am J Psych.164:1404-1410.

• Stahl, S.M. & Grady, M.M. (2004). A critical review of atypical antipsychotic utilization: comparing monotherapy with polypharmacy augmentation. Curr Med Chem.11:313-327.

• Tandon, R. (2012). Antipsychotic polypharmacy: update and guidelines for practice. Retrieved on March 27, 2012 at: http://medicaidmentalhealth.org

• Thompson, A., Sullivan, S.A., Barley, M., Strange, S.O., Moore, L., Rogers, P., Sipos, A.& Harrison, G. (2008). The DEBIT trial: an intervention to reduce antipsychotic polypharmacy prescribing in adult psychiatry wards- a randomized controlled trial. Psychological Medicine. 38(5);705-15.

• Tranulis, C., Skalli, L., Lalonde, P., & Nicole, L. (2008). Benefits and risks of antipsychotic polypharmacy. An evidence based review of the literature. _Drug Saf. 31_(1):7-20

• University HealthSystem Consortium. (2006). Mental health performance measures field brief. Oakbrook, IL.

• Zhu, B., Ascher-Svanum, H., Faries, D.E., Correll, C.U., Kane, & J.M. (2008). Cost of antipsychotic polypharmacy in the treatment of schizophrenia. BMC Psychiatry. 8, ArtID 19.

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)*

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

Research studies have found that 4-35% of outpatients and 30-50% of inpatients treated with an antipsychotic medication concurrently received 2 or more antipsychotics (Covell, Jackson, Evans, & Essock, 2002; Ganguly, Kotzan, Miller, Kennedy, & Martin, 2004; Gilmer, Dolder, Folsom, Mastin, & Jeste, 2007; Kreyenbuhl, Valenstein, McCarthy, Ganocyz, & Blow, 2006; Stahl & Grady, 2004). One study reported 4.6% of patients concurrently received 3 or more antipsychotics (Jaffe & Levine, 2003). These findings are seen across diverse sectors: state mental health authorities, the Veterans Health System and Medicaid-financed care. Antipsychotic polypharmacy can lead to greater side effects, often without improving clinical outcomes (Ananth, Parameswaran, & Gunatilake, 2004; Stahl & Grady, 2004). As a result, a range of stakeholders have called for efforts to reduce unnecessary use of multiple antipsychotics (Centorrino, Gören, Hennen, Salvatore, Kelleher, & Baldessarini, 2004; Gilmer, Dolder, Folsom, Mastin, & Jeste, 2007; National Association of State Mental Health Program Directors, 2001; University HealthSystem Consortium, 2006). Practice guidelines recommend the use of a second antipsychotic only after multiple trials of a single antipsychotic have proven inadequate (American Psychiatric Association [APA] Practice Guidelines, 2004). Randomized controlled trials (RCTs) provide some evidence to support augmentation with a second antipsychotic in treatment resistant patients. Most of these studies were limited to augmentation of clozapine with another second-generation antipsychotic (Tranulis, Skalli, Lalonde, & Nicole, 2008). Among patients without a documented history of previous treatment failures of antipsychotic monotherapy, multiple RCTs and other controlled trials failed to show a benefit of antipsychotic polypharmacy over monotherapy (Ananth, Parameswaran, & Gunatilake, 2004; Centorrino, Gören, Hennen, Salvatore, Kelleher, & Baldessarini, 2004; Potkin, Thyrum, Alva, Bera, Yeh, & Arvanitis, 2002; Shim et al., 2007; Stahl,& Grady, 2004). Clinical circumstances, such as shorter inpatient stays, may require hospitals to discharge a patient on multiple antipsychotics with an aftercare plan to transition to monotherapy.

In such cases, effective communication between the inpatient and aftercare clinician is an essential element of care.

As stated above, recent literature supports three appropriate justifications for prescribing multiple antipsychotic medications: previous failed trials of monotherapy, cross-tapering to monotherapy and augmentation of clozapine. A review of the justifications for prescribing more than one antipsychotic medication will help hospitals determine if their practice is supported by the evidence-base.

The measure will assist health care organizations (HCOs) to track the number of patients prescribed two or more antipsychotic medications at the time of discharge with appropriate justification.

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.*

Below are the data from 2009-2018. The Year of data submission is the first row followed by N,the number of Hospitals that have directly submitted data to the Joint Commission. Descriptive statistics include mean, std. dev, min, max, median, first and 3rd quartiles (Q1 and Q3) along the deciles listed at the 10 percentile (10th pctl), etc.

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
N	296	316	463	477	505	651	1015	1035	729	688
Mean	0.29737	0.40139	0.40022	0.48456	0.50948	0.53775	0.57072	0.58785	0.60979	0.61938
Std. Dev.	0.2617	0.3377	0.3285	0.3387	0.3309	0.3344	0.332	0.3252	0.3185	0.3389
Max	1	1	1	1	1	1	1	1	1	1
Q3	0.43651	0.66667	0.68182	0.8125	0.80357	0.84848	0.88	0.87879	0.89474	0.94444
Median	0.25	0.3144	0.33333	0.46809	0.51685	0.54545	0.60119	0.64706	0.66667	0.68661
Q1	0.06594	0.10286	0.10345	0.16667	0.2	0.25	0.3	0.32258	0.375	0.36932
Min	0	0	0	0	0	0	0	0	0	0
10th Pctl	0	0	0	0	0.03704	0.03333	0.06667	0.07143	0.08511	0.0101
20th Pctl	0.04545	0.05556	0.07246	0.11538	0.14286	0.19048	0.2381	0.25	0.28571	0.25
30th Pctl	0.10526	0.13953	0.14815	0.22472	0.28571	0.3125	0.35185	0.3913	0.44444	0.4375
40th Pctl	0.19737	0.22581	0.21622	0.36	0.4	0.44828	0.5	0.53145	0.55856	0.57143
60th Pctl	0.3125	0.44737	0.45732	0.58621	0.65109	0.66667	0.73221	0.75	0.75	0.78261
70th Pctl	0.375	0.59292	0.6	0.71875	0.75	0.78571	0.82759	0.83333	0.85714	0.9
80th Pctl	0.5	0.77419	0.76256	0.85714	0.85714	0.9	0.92206	0.91667	0.93902	1

90th Pctl 0.69444 0.96774 0.91429 0.97959 0.98824 1 1
1 1 1

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.

See data in 1b.2

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.*

No disparities were noted in the literature.

For data source see data in 1b.2

The following data are measure rates by population group

Gender	2013	2014	2015	2016	2017
Male	0.541	0.577	0.620	0.617	0.618
Female	0.540	0.591	0.621	0.601	0.623

Hispanic Ethnicity	2013	2014	2015	2016	2017
Hispanic	0.647	0.585	0.658	0.625	0.643
Non-Hispanic	0.531	0.583	0.618	0.610	0.619

Race	2013	2014	2015	2016	2017
White	0.522	0.565	0.608	0.602	0.618
African American			0.541	0.581	0.634	0.636	0.647
American Indian			0.434	0.506	0.534	0.441	0.433
Asian	0.536	0.604	0.683	0.626	0.621
Pacific Islander	0.539	0.390	0.611	0.530	0.644

Age Category	2013	2014	2015	2016	2017
1-12 years	0.575	0.562	0.584	0.628	0.434
13-17 years	0.505	0.522	0.590	0.588	0.509
18-64 years	0.537	0.569	0.631	0.620	0.632
65+ years	0.463	0.512	0.563	0.561	0.594

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):*

Safety, Safety : Medication, Safety : Overuse

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

Elderly

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.)*

https://manual.jointcommission.org/releases/TJC2018B1/HospitalBasedInpatientPsychiatricServices.html

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 **Attachment:** HBIPS_Code_Tables_Med_-636794264289743033.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.

Denominator statement changed from: Psychiatric inpatients discharged on two or more routinely scheduled antipsychotic medications to: Psychiatric inpatient discharges.

Reason for change: patients discharged on on two or more antipsychotic medications were previously identified in the HBIPS-4 measure - Patients Discharged on Multiple Antipsychotic Medications. HBIPS-4 was retired effective January 1, 2016.

Data element: Patient Referral to Next Level of Care Provider removed January 1, 2016. Replaced with data element: Patient Status at Discharge

Reason for change: HBIPS-6 - Post Discharge Continuing Care Plan Created and HBIPS-7 - Continuing Care Plan Transmitted to the Next Level of Care Provider upon Discharge were retired January 1, 2016. New data element Patient Status at Discharge takes into account circumstances in which patient is discharged abruptly or outside the plan of care.

Appendix C Medication Table 10.0 (Antipsychotic Medications) was updated July 1, 2016 and January 1, 2017 with the addition of new antipsychotic medications.

The ICD-10-CM code table for Mental Disorders was revised to reflect the ICD-10 code updates for Fiscal Year (FY) 2019, effective for discharges October 1, 2018.

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).

Psychiatric inpatients discharged on two or more routinely scheduled antipsychotic medications with appropriate justification.

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).

One data element is used to calculate the numerator:

Appropriate Justification for Multiple Antipsychotic Medications - Documentation in the medical record of appropriate justification for discharging the patient on two or more routine antipsychotic medications. Allowable values: 1. The medical record contains documentation of a history of a minimum of three failed multiple trials of monotherapy. 2. The medical record contains documentation of a recommended plan to taper to monotherapy due to previous use of multiple antipsychotic medications OR documentation of a cross-taper in progress at the time of discharge. 3. The medical record contains documentation of augmentation of Clozapine. 4. The medical record contains documentation of a justification other than those listed in Allowable Values 1-3. 5. The medical record does not contain documentation supporting the reason for being discharged on two or more antipsychotic medications OR unable to determine from medical record documentation.

Patients are eligible for the numerator population when they are discharged on two or more routinely scheduled antipsychotic medications with appropriate justification.

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

Psychiatric inpatient discharges

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).

Included populations:

Patients with ICD-10-CM Principal or Other Diagnosis Codes for Mental Disorders as defined in Appendix A, Table 10.01 (See S.2b.) discharged on two or more routinely scheduled antipsychotic medications (refer to Appendix C, Table 10.0- Antipsychotic Medications).

Nine data elements are used to calculate the denominator:

1. Admission Date - The month, day and year of admission to acute inpatient care.

2. Birthdate - The month, day and year the patient was born.

3. Discharge Date – The month day and year the patient was discharged from acute care, left against medical advice or expired during the stay.

4. Discharge Disposition- The patient's discharge disposition. Allowable values: 1. Home, 2. Hospice – Home, 3. Hospice – Health Care Facility, 4. Acute Care Facility, 5. Other Health Care Facility, 6. Expired, 7. Left Against Medical Advice/AMA, 8 Not Documented or Unable to Determine (UTD).

5. ICD-10-CM Other Diagnosis Codes- The other or secondary (ICD-10-CM) codes associated with the diagnosis for this hospitalization.

6. ICD-10-CM Principal Diagnosis Code- The ICD-10-CM diagnosis code that is primarily responsible for the admission of the patient to the hospital for care during this hospitalization.

7. Number of Antipsychotic Medications Prescribed at Discharge- The number of routinely scheduled antipsychotic medications prescribed to the patient at discharge as documented in the medical record. Allowable values: 0-99, UTD (Unable to determine)

8. Patient Status at Discharge - Documentation in the medical record of the patient´s status at the time the patient left the hospital-based inpatient psychiatric care setting. Allowable values: 1 The medical record contains documentation that the patient was discharged from the inpatient psychiatric care setting under these circumstances:

• Patient is leaving the psychiatric unit within the acute care hospital AND the hospital facility completely.

• Patient is leaving the freestanding inpatient psychiatric facility completely.

2 The medical record contains documentation of one of the following:

• The patient eloped and was discharged

• The patient failed to return from leave and was discharged

• The patient has not yet been discharged from the hospital

• The patient was transferred/discharged from the inpatient psychiatric unit in an acute care setting to another level of care, (i.e. medical unit), and subsequently discharged from that level of care

3 Unable to determine from medical record documentation.

9. Psychiatric Care Setting - Documentation in the medical record that the patient was receiving care primarily for a psychiatric diagnosis in an inpatient psychiatric setting, i.e., a psychiatric unit of an acute care hospital or a free-standing psychiatric hospital. Allowable values: Yes, No.

Populations: Discharges with Table 10.01 Mental Disorders in the Psychiatric Care Setting who were discharged on two or more routinely scheduled antipsychotic medications on Table 10.0.

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

• Patients who expired

• Patients with an unplanned departure resulting in discharge due to elopement

• Patients with an unplanned departure resulting in discharge due to failing to return from leave

• Patients with a length of stay less than or equal to 3 days

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.)*

• Patients who expired are identified by the data element Discharge Disposition.

• Patients with an unplanned departure resulting in discharge due to elopement, failing to return from leave are identified by the data element Patient Status at Discharge.

• Length of stay (LOS) in days is equal to the Discharge Date minus the Admission Date. If the LOS is equal to or less than 3 days the patient is excluded.

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.)*

The measure is stratified by the following age groups:

• Children (1 through 12 years) — A Patient Age at Discharge (Discharge Date minus Birthdate) greater than or = 1 year and less than 13 years

• Adolescent (13 through 17 years) — A Patient Age at Discharge (Discharge Date minus Birthdate) greater than or = 13 years and less than 18 years

• Adult (18 through 64 years) - A Patient Age at Discharge (Discharge Date minus Birthdate) greater than or = 18 years and less than 65 years

• Older Adult (65 years or greater) - A Patient Age at Discharge (Discharge Date minus Birthdate) greater than or = 65 years

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.)*

1. Run cases that are included in the Initial Patient Population for HBIPS-1,5 and pass the edits defined in the Transmission Data Processing Flow: Clinical through this measure.

2. Calculate Length of Stay. Length of Stay, in days, is equal to the Discharge Date minus the Admission Date.

3. Check Length of Stay

a. If Length of Stay is less than or equal 3 days, the case will proceed to a Measure Category Assignment of B for Overall Rate (HBIPS-5a) and will not be in the measure population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

b. If Length of Stay is greater than 3 days, continue processing and proceed to Discharge Status.

4. Check Discharge Disposition

a. If Discharge Disposition is missing, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-5a) and will be rejected. Continue processing and proceed to step 10 and Initialize the Measure Category Assignment for each strata measure.

b. If Discharge Disposition equals 6, the case will proceed to a Measure Category Assignment of B for Overall Rate (HBIPS-5a) and will not be in the measure population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

C. Discharge Disposition equals 1, 2, 3, 4, 5, 7, or 8, continue processing and proceed to Psychiatric Care Setting.

5. Check Psychiatric Care Setting

a. If Psychiatric Care Setting equals No, the case will proceed to a Measure Cat Category Assignment of B for Overall Rate (HBIPS-5a) and will not be in the measure population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

b. If Psychiatric Care Setting is missing, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-5a) and will be rejected. Continue processing and proceed to step 10 and Initialize the Measure Category Assignment for each strata measure.

c. If Psychiatric Care Setting equals Yes, the case will proceed to Patient Status at Discharge.

6. Check Patient Status at Discharge

a. If Patient Status at Discharge is missing, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-5a) and will be rejected. Continue processing and proceed to step 10 and Initialize the Measure Category Assignment for each strata measure.

b. If Patient Status at Discharge equals 2, the case will proceed to a Measure Category Assignment of B for Overall Rate (HBIPS-5a) and will not be in the measure population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

c. If Patient Status at Discharge equals 1 or 3, the case will continue processing and proceed to Number of Antipsychotic Medications Prescribed at Discharge.

7. Check Number of Antipsychotic Medications Prescribed at Discharge

a. If Number of Antipsychotic Medications Prescribed at Discharge is missing, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-5a) and will be rejected. Continue processing and proceed to step 10 and Initialize the Measure Category Assignment for each strata measure.

b. If Number of Antipsychotic Medications Prescribed at Discharge is less than or equal 1, the case will proceed to a Measure Category Assignment of B for Overall Rate (HBIPS-5a) and will not be in the measure population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

c. If Number of Antipsychotic Medications Prescribed at Discharge is greater than or equal 2 or equal UTD, the case will continue processing and proceed to Number of Antipsychotic Medications Prescribed at Discharge.

8. Check Number of Antipsychotic Medications Prescribed at Discharge

a. If Number of Antipsychotic Medications Prescribed at Discharge equals UTD, the case will proceed to a Measure Category Assignment of D for Overall Rate (HBIPS-5a) and will be in the measure population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

b. If Number of Antipsychotic Medications Prescribed at Discharge is greater than or equal 2, the case will proceed to Appropriate Justification for Multiple Antipsychotic Medications.

9. Check Appropriate Justification for Multiple Antipsychotic Medications

a. If Appropriate Justification for Multiple Antipsychotic Medications is missing, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-5a) and will be rejected. Continue processing and proceed to step 10 and Initialize the Measure Category Assignment for each strata measure.

b. If Appropriate Justification for Multiple Antipsychotic Medications equals 4 or 5, the case will proceed to a Measure Category Assignment of D for Overall Rate (HBIPS-5a) and will be in the measure population.

Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

c. If Appropriate Justification for Multiple Antipsychotic Medications equals 1, 2 or 3, the case will proceed to a Measure Category Assignment of E for Overall Rate (HBIPS-5a) and will be in the numerator population. Continue processing and proceed to step 10 and initialize the Measure Category Assignment for each strata measure.

10. Initialize the Measure Category Assignment for each strata measure (b-e) = ´B´. Do not change the Measure Category Assignment that was already calculated for the overall rate (HBIPS-5a). The rest of the algorithm will reset the appropriate Measure Category Assignment to be equal to the overall rate´s (HBIPS-5a) Measure Category Assignment.

11. Check Overall Rate Category Assignment

a. If Overall Rate Category Assignment equals B, Set the Measure Category Assignment for the strata measures (HBIPS-5b through HBIPS-5e) = ´B´. Stop processing.

b. If Overall Rate Category Assignment equals D or E or X, continue processing and proceed to Patient Age at Discharge.

12. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal 1 years and less than 13 years, set the Measure Category Assignment for measure HBIPS-5b = Measure Category Assignment for measure HBIPS-5a. Stop processing.

b. If is greater than or equal 13 years, continue processing and proceed to Patient Age at Discharge.

13. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal 13 years and less than 18 years, set the Measure Category Assignment for measure HBIPS-5c = Measure Category Assignment for measure HBIPS-5a. Stop processing.

b. If Patient Age at Discharge is greater than or equal 18 years, continue processing and proceed to Patient Age at Discharge.

14. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal 18 years and less than 65 years, set the Measure Category Assignment for measure HBIPS-5d = Measure Category Assignment for measure HBIPS-5a. Stop processing.

b. If Patient Age at Discharge is greater than or equal 65 years, set the Measure Category Assignment for measure HBIPS-5e = Measure Category Assignment for measure HBIPS-5a. Stop processing.

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.

Hospitals that choose to sample have the option of sampling quarterly or sampling monthly. A hospital may choose to use a larger sample size than is required. Hospitals whose Initial Patient Population size is less than the minimum number of cases per quarter/month for the stratum cannot sample that stratum.

Regardless of the option used, hospital samples must be monitored to ensure that sampling procedures consistently produce statistically valid and useful data. Due to exclusions, hospitals selecting sample cases MUST submit AT LEAST the minimum required sample size.

Quarterly Sampling

For hospitals selecting sample cases for the HBIPS discharge measures, a modified sampling procedure is required. Hospitals selecting sample cases for this set must ensure that each individual stratum's population and effective quarterly sample size meet the following conditions:

- Select within each of the four individual measure strata. The effective quarterly sample size within a stratum is at least 44 cases per quarter. Cases are placed into the appropriate stratum based upon the patient's age.
- The required quarterly sample size is at least 20% of the stratum population for the quarter.

Quarterly Sample Size

Based on Initial Patient Population for the HBIPS Discharge Measures

Average Quarterly Stratum Initial Patient Population Size	Minimum Required Stratum Sample Size
> 877	176
221-877	20% of Initial Patient Population size
44-220	44
< 44	No sampling; 100% Initial Patient Population required

Monthly Sampling

Hospitals selecting sample cases for this set must ensure that each individual stratum population and effective monthly sample size meet the following conditions:

- Select within each of the four individual measure strata. The effective monthly sample size within a stratum is at least 15 cases per month. Cases are placed into the appropriate stratum based upon the patient's age.
- The required monthly sample size is at least 20% of the stratum population for the month.

Monthly Sample Size

Based on Initial Patient Population Size for the HBIPS Measure Set

Average Monthly Stratum Initial Patient Population Size	Minimum Required Stratum Sample Size
> 295	60
76-295	20% of Initial Patient Population size
15-75	15
< 15	No sampling; 100% Initial Patient Population required

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.

Not applicable

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

If other, please describe in S.18.

Electronic Health Records, Paper Medical Records

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.

Each data element in the data dictionary includes suggested data sources. The data are collected using contracted Performance Measurement Systems (vendors) that develop data collection tools based on the measure specifications. The tools are verified and tested by Joint Commission staff to confirm the accuracy and conformance of the data collection tool with the measure specifications. The vendor may not offer the measure set to hospitals until verification has been passed.

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, Other

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

Inpatient/Hospital

If other:

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

Not applicable

2. Validity – See attached Measure Testing Submission Form

0560_MeasureTesting_7.1_HBIPS5-636898058369558963.docx

2.1 <u>For maintenance of endorsement</u>

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.

No

2.2 <u>For maintenance of endorsement</u>

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 <u>For maintenance of endorsement</u>

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.

No - This measure is not risk-adjusted

2.3 <u>For maintenance of endorsement</u>

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.

Measure Testing (subcriteria 2a2, 2b1-2b6)

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

Measure Number (*if previously endorsed*)**:** 0560
Measure Title: Patients discharged on multiple antipsychotic medications with appropriate justification
Date of Submission: 12/20/2018

Type of Measure:

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

Instructions

- **Measures must be tested for all the data sources and levels of analyses that are specified. *If there is more than one set of data specifications or more than one level of analysis, contact NQF staff* about how to present all the testing information in one form.**
- **For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures, section 2b3 also must be completed.**
- **If specified for multiple data sources/sets of specificaitons (e.g., claims and EHRs), section 2b5 also must be completed.**
- **Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2 a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.**
- **If you are unable to check a box, please highlight or shade the box for your response.**
- **Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). *Contact NQF staff if more pages are needed.***
- **Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.**
- **For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.**

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

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument-based measures

(including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

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

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

OR

- **rationale/data support no risk adjustment/ stratification.**

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

OR

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

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

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

Notes

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

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

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

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

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

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

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: Click here to describe	☐ other: Click here to describe

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*).

Not applicable

1.3. What are the dates of the data used in testing? 4/1/2007 – 7/1/2007

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: Click here to describe	☐ other: Click here to describe

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*)

Description of the population characteristics

This measure has been in national use since the 4th quarter of 2008. Demographics of organizations collecting and reporting data on these measures are as follows:

487 Health care organizations representing various types, locations and sizes:

408 Free-Standing Psychiatric Hospitals, 79 Acute-Care Hospitals with Psychiatric Units

103 For Profit, 120 Not for Profit, 184 Government

103 >=300 beds; 217 100-299 beds; 67 <100 beds

States represented in this data collection effort include: AK, AL, AR, AZ, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS, KY, LA, MA, MD, ME, MI, MN, MO, MS, NC, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, PR, RI, SC, SD, TN, TX, UT, VA, VT, WA, WI, WV, WY

27 performance measurement systems are used for data transmission to The Joint Commission.

Description of sampling method

Ten hospitals were randomly sampled from the 487 hospitals in the population, using a stratified sampling methodology to represent the three bed size and three ownership categories.

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*)

Patients were randomly sampled from each of the ten hospitals in the sample, using a stratified sampling methodology so that measure numerator and denominator cases identified in the original abstraction were represented in the sample and an equal number of cases were sampled for each hospital. There were 191 patients sampled in all.

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.

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.

Not applicable, not required at the time this testing was done.

_____

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*)

All sampled cases were re-abstracted by trained Joint Commission staff. Re-abstracted data are compared with originally abstracted data on a data element by data element basis. The test used were the calculated agreement rates for individual data elements that are used to compute measure rates for the measure.

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*)

Data Elements	Total Numerator	Total Denominator	Agreement Rate
Numerator Data Element			
Appropriate justification for multiple antipsychotic medications	5	5	100%
Denominator Data Elements			
Admission Date	191	191	100%
Birthdate	191	191	100%
Discharge Date	191	191	100%
Discharge Disposition	191	191	100%
ICD-9-CM Other Diagnosis Codes*	191	191	100%
ICD-9-CM Principal Diagnosis Code*	191	191	100%
Number of Antipsychotic Medications Prescribed at Discharge	191	191	100%
Patient Status at Discharge	191	191	100%
Psychiatric Care Setting	191	191	100%

* The mesure was tested with ICD-9-CM codes. A crosswalk from ICD-9-CM diagnosis codes to ICD-10-CM diagnosis codes was done and reviewed by the Technical Advisory Panel. The panel determined that the intent of the measure was not changed as a result of the conversion.

The above numerator and denominator data elements were assessed for reliability: appropriate justification for multiple antipsychotic medications. There was a 100% match for the calculated agreement rate for the data element used to compute measure rates for HBIPS-5.

Additionally, re-abstraction data analysis containing the health care organization's Category Assignment Agreement Rate (CAAR) which represents assignment to the numerator or denominator was performed on data from sample hospitals resulting in an agreement rate of 100%.

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?*)

A perfect agreement rate between originally abstracted data and re-abstracted data equals 100%, and an agreement rate below 75% is considered failing. These agreement rates are considered to be well within acceptable levels.

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)*

Face validity was tested by a total of 40 hospitals during May and June 2006. Measure information was sent to the test hospitals for review. In addition, three site visits with focus interviews were conducted. One site visit had a total of nine state hospitals represented. Criterion validity was evaluated during the focus group interviews conducted during the reliability site visits mentioned above as well as through an online survey that all pilot hospitals were invited to complete.

The measure information form and the data dictionary were evaluated for face validity. The following parts of the measure information form were evaluated: numerator statement, numerator inclusions, numerator exclusions, denominator statement, denominator inclusions, denominator exclusions and an overall understanding of the measure information form. Each area was scored utilizing a five-point likert scale. For each data element, the hospitals were asked to comment on the clarity and understanding of the abstraction guidelines and data definitions. And finally, the data dictionary was reviewed for overall understanding, usefulness and overall clarity utilizing a five-point likert scale. Qualitative analysis was performed on measure feedback received during the focus group interviews and from the online surveys.

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

Tests for correlations between HBIPS-5 and the remaining HBIPS measures (HBIPS-1, HBIPS-2, HBIPS-3) are 0.13857(p=0.0002), -0.04720 (p=0.2068), and -0.04642 (p=0.2164), respectively. This indicates that there are no statistically significant correlations between HBIPS-5 and HBIPS-2 and HBIPS-3. There is a slight positive correlation between HBIPS-5 and HBIPS-1. Employing a longitudinal logistic regression model with the hospital as a random effect yields a significant improvement of rates over time (p<0.0001).

Queries submitted via the automated feedback system have decreased significantly for the HBIPS measure set in the past 3 years. (522 in 2016, 288 in 2017, 187 for 2018 YTD). There have been no major issues with the data elements for this measure.

A total of 36 hospitals completed the face validity evaluation and rated the overall understanding of the measure as follows: very good n=7, good n=17, average n=7, poor n=4 and very poor n=1. Modifications to improve the understanding and clarity of the measure specifications were made prior to pilot testing based on feedback received from the hospitals during the face validity evaluation. Analysis of the focus group discussions and the online survey revealed a majority of the pilot hospitals recommended moving the measure forward in the final measure set with suggested modifications. Since that time continual feedback from customers does not indicate a change in their perception of the measure. Also, this measure has been evaluated for validity and adopted for use in a national reimbursement program (CMS).

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 positive correlation between HBIPS-5 and HBIPS-1 validates the use of these 2 measures for evaluating quality of care in the behavioral health setting.

The measure has considerable face validity which has been improved over time.

_____

2b2. EXCLUSIONS ANALYSIS

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

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*)

Data from reporting hospitals was analyzed to determine the incidence of the measure exclusions based on 2017 HBIPS data.

Measure exclusions that were not derived directly from the evidence are presented below. Please note that these are population exclusions that are necessary to ensure consistency in all measures in this measure set.

These denominator exclusions were analyzed for frequency of occurrence. An issue that is of great concern to users of this measure is that due to the presence of exceptions to the measure, attainment of a 100% measure rate is not possible. Because of the role of this measure in the current Joint Commission accreditation process this is especially troubling to end users. This concern is the basis for a number of the non-evidence-based exclusions to these measures. The following measure exclusions that were not derived directly from the evidence are as follows:

- Patients who expired

- Patients with an unplanned departure resulting in discharge due to elopement or failing to return from leave
- Patients with a length of stay equal to or less than 3 days.
- Number of antipsychotic medications prescribed at discharge ≤ 1

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*)

2017 Discharges N= 520,778

- Patients who expired= 0.07%
- Patients with an unplanned departure resulting in discharge due to elopement or failing to return from leave= 2.2%
- Patients with a length of stay equal to or less than 3 days= 16.0%
- Number of antipsychotic medications prescribed at discharge ≤ 1 = 76.6%

Rationale for exclusions:

- Patients who expired

  Rationale: There are no discharge medications
- Patients with an unplanned departure resulting in discharge due to elopement or failing to return from leave

  Rationale: There is no opportunity for discharge planning
- Patients with a length of stay ≤ 3 days

  Rationale: Some patients are admitted that are stable on more than one antipsychotic medication. A short length of stay does not allow time for medication changes and patient evaluation.
- Antipsychotics prescribed ≤ 1

  Rationale: Only those cases where the number of antipsychotics prescribed is ≥ 2 are in this measure.

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*)

The rationale indicates that based on the exclusions, these cases would not be eligible for the measure.

The incidence of these exclusion is frequent enough to continue to include in the measure specifications.

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.

Not applicable

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

☐ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

Not applicable

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

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.

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?

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)**

Not applicable

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

Not applicable

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

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

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

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

Not applicable

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

Not applicable

2b3.9. Results of Risk Stratification Analysis:

Not applicable

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

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

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)*

The method used to analyze meaningful differences in performance at The Joint Commission is Target Analysis. The object of target analysis is to compare a health care organization's data against a comparative norm for the purpose of evaluating performance improvement opportunities. When an organization's performance level is statistically significantly different from a comparative norm, it is considered a statistical deviation. A statistical deviation may be desirable or undesirable depending on the "direction of improvement" of the measure.

There are two components to the target analysis methodology used at The Joint Commission. Given the national average for a performance measure, a target range is constructed. Using generalized linear mixed models methodology (also known as hierarchical models), a predicted estimate of an HCO's performance, with a corresponding 95% confidence interval, is generated. This confidence interval is compared to the target range, to determine the HCO's rating. The estimate of the organization's true performance is based on both the data from that organization and on data from the entire set of reporting organizations.

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*)

HBIPS-5 Distribution of Measure Results

2018 2nd Quarter Data:
Scores on this measure: N=646, Mean 62.8%, SD 0.3640
10th Percentile= 0%
25th Percentile= 33.3%
50th Percentile= 73.8%
75th Percentile= 100%
90th Percentile= 100%

199 (30.8%) Favorable – results statistically significantly higher than the national rate
352 (54.5%) Neutral – results not significantly different from target range
95 (14.7%) Unfavorable - results statistically significantly lower than the national rate

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?*)

Employing a longitudinal logistic regression model with the hospital as a random effect yields a significant improvement of rates over time ($p<0.0001$). Although there were improvements over time, measure results continue to demonstrate a gap in care. This measure is important to continue improvement in decreasing the rates for prescribing multiple antipsychotics at discharge.

An appreciable number of hospitals were identified with substandard performance for this measure, with performance significantly below the national average.

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

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

Not applicable

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

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

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

_____

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

Not applicable. The measure has been collected since 2008 and hospitals transmitting data with missing data on any of the critical data elements are not accepted.

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 nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable.

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*)

Not applicable

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 nonresponders) 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)

Not applicable

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 by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition, Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

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**.

Some data elements are in defined fields in electronic sources

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).

Although The Joint Commission had intended to pursue the process to convert this measure to an electronic quality measure (eCQM), this has not occurred for the following reasons:

- The adoption of eCQMs may be difficult for free-standing psychiatric facilities because the electronic medical record (EMR) has not been consistently integrated across these facilities.
- It has been the experience of The Joint Commission that it can be difficult and resource intensive to successfully re-engineer a chart-based measure to an eCQM as opposed to new eCQM development.

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.

Hospitals using this performance measure generally collect measure data via manual review of the paper medical record. Collected data are submitted to The Joint Commission on a quarterly basis, by way of contracted performance measurement system vendors, as described previously. Specifications for this measure are freely available to anyone who wishes to use the measure. Feedback from hospitals using this measure indicates that required data elements are generally available in the medical record, and measure specifications are robust and easy to understand. If feedback from measure users has indicated the need for clarification or revision of measure specifications, this has taken place.

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)*.

Not applicable, there are no fees, licensing, or other requirements.

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 ORYX Performance Measurement Reporting Program https://www.qualitycheck.org/ Payment Program Inpatient Psychiatric Facility Quality Reporting (IPFQR) Program https://www.medicare.gov/hospitalcompare/search.html Regulatory and Accreditation Programs Hospital Accreditation Program http://jointcommission.org Quality Improvement (external benchmarking to organizations) America's Hospitals: Improving Quality and Safety – The Joint Commission's Annual Report 2017 https://www.jointcommission.org/annualreport.aspx Quality Improvement (Internal to the specific organization) ORYX Performance Measurement Report Not available to public; only accessible to the organization

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

• Name of program and sponsor: Inpatient Psychiatric Facility Quality Reporting (IPFQR) Program/Centers for Medicare & Medicaid Services
• Purpose: The IPFQR Program gives consumers care quality information to help them make more informed decisions about their healthcare options. This includes providing consumers with data about quality measures that aim to assess and foster improvement in the quality of care provided to patients with mental illness. The IPFQR Program encourages facilities and clinicians to improve the quality of inpatient care. The program helps by making sure providers know about and report on the best practices for their facilities and type of care they give by submitting quality data to CMS annually.
• Geographic area and number and percentage of accountable entities and patients included: United States All IPFs paid under the Inpatient Psychiatric Facilities Prospective Payment System (IPF PPS) have to meet IPFQR Program requirements. As of 12/1/2018, there are 1,635 participating providers in the IPFQR Program.
• Level of measurement and setting: The IPF PPS applies to inpatient psychiatric services given by psychiatric hospitals or psychiatric units (also known as mental health or behavioral health units) in Acute Care Hospitals (ACHs) or Critical Access Hospitals (CAHs) in the United States that participate in Medicare.
• Name of program and sponsor: ORYX Performance Measurement Reporting Program/The Joint Commission
• Purpose: The Joint Commission's ORYX initiative integrates performance measurement data into the accreditation process. ORYX measurement requirements support Joint Commission-accredited organizations in their quality improvement efforts
• Geographic area and number and percentage of accountable entities and patients included: Nationwide; 726 free-standing psychiatric hospitals and hospitals with psychiatric units accredited by The Joint Commission
• Level of measurement and setting: Level of measurement and setting: facility level of measurement, inpatient setting
• Name of program and sponsor: America's Hospitals: Improving Quality and Safety – The Joint Commission's Annual Report 2017/The Joint Commission
• Purpose: The Joint Commission's ORYX initiative integrates performance measurement data into the accreditation process. ORYX measurement requirements support Joint Commission-accredited organizations in their quality improvement efforts

• Geographic area and number and percentage of accountable entities and patients included: Nationwide; 726 free-standing psychiatric hospitals and hospitals with psychiatric units accredited by The Joint Commission
• Level of measurement and setting: Level of measurement and setting: facility level of measurement, inpatient setting
• Name of program and sponsor: ORYX Performance Measurement Report/The Joint Commission
• Purpose: The Joint Commission's ORYX initiative integrates performance measurement data into the accreditation process. ORYX measurement requirements support Joint Commission-accredited organizations in their quality improvement efforts
• Geographic area and number and percentage of accountable entities and patients included: Nationwide; 726 free-standing psychiatric hospitals and hospitals with psychiatric units accredited by The Joint Commission
• Level of measurement and setting: Level of measurement and setting: facility level of measurement, inpatient setting
• Name of program and sponsor: Hospital Accreditation Program/The Joint Commission
• Purpose: The Joint Commission's ORYX initiative integrates performance measurement data into the accreditation process. ORYX measurement requirements support Joint Commission-accredited organizations in their quality improvement efforts
• Geographic area and number and percentage of accountable entities and patients included: Nationwide; 726 free-standing psychiatric hospitals and hospitals with psychiatric units accredited by The Joint Commission
• Level of measurement and setting: Level of measurement and setting: facility level of measurement, inpatient setting

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.

Measure rates are provided to the hospital via a quarterly ORYX Performance Measure Report. This applies to all entities reporting the measure.

The Joint Commission utilizes an email process for hospital contact related to their measure rates and analysis. Response is provided in a timely manner either by email or directly by phone. Additionally, the data is available publicly through The Joint Commission Quality Check website. Individual hospital data for each rolling yearly time period are viewable and can be downloaded from this website.

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.

Patient level data is aggregated at the hospital level quarterly. The hospital Performance Measure Report and Quality Check website are updated. A users guide to the Performance Measure Report is posted on the Joint Commission website. Quality Check includes yearly and quarterly hospital rates, state and national averages, and the top 10 percentile at the national and state level.

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 Joint Commission utilizes an automated feedback system with access available to the measured entities and the vendors contracted by measured entities. A clinical lead is responsible for each individual measure set. The system is monitored on a daily basis and response is provided typically within 8 business hours. If queries cannot be managed via written response, arrangements are made to address any issues or concerns via phone.

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

Queries submitted via the automated feedback system have decreased significantly for the HBIPS measure set in the past 3 years. (522 in 2016, 288 in 2017, 187 for 2018 YTD). There have been no issues with the data elements for this measure and no updates needed to the data element specifications based upon feedback received.

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

Same as above in 4a2.2.2.

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.

Note: all feedback is tracked and considered. If upon analysis there are trends noted giving cause for updates, this is reviewed by the measure work-group to confirm the need for revision. Additionally, The Joint Commission engages a Technical Advisory Panel (TAP) that is consulted on an as needed basis for approval of updates that may require their additional expertise. All measure specifications are reviewed twice a year and updates are made as needed based on feedback from the measure users, input from the TAP, or changes in the guidelines.

Modifications to this measure have not been required based upon feedback received.

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.

Though 2009 to 2nd quarter 2018, a binomial random effects model was used to determine if there was a change in rates over time with time as a fixed effect and healthcare organization as a random effect. The results of the model show statistical significant over time (P<0.001) and an odd ratio estimate of time to be 1.302.

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.

Since implementation, the Notes for Abstraction section and the Allowable Value section of the data element for Appropriate Justification for Multiple Antipsychotic Medications has been updated to clarify issues that have been identified after review of the feedback received from measure users. To the best of our knowledge,

there have been no reports of unintended consequences.To the best of our knowledge, there have been no unexpected findings and no reports of unintended consequences.

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

A study published in July 2018, compared results on psychiatric performance measures among cohorts of hospitals with different characteristics that elected to begin reporting on the HBIPS measures at various points in time.

Quarterly reporting of Hospital-Based Inpatient Psychiatric Services (HBIPS) measures to the Joint Commission was used to examine trends in performance among four hospital cohorts that began reporting in 2009 (N=243), 2011 (N=139), 2014 (N=137), or 2015 (N=372).

Results demonstrated that all cohorts significantly improved across quarters for justification of multiple antipsychotic medications.

Citation:

Rasinksi, K.A., Schmaltz, S.P., Williams, S.C., & Baker, D.W. (2018). Trends in results of HBIPS National Performance Measures and association with year of adoption. Psychiatric Services, 69(7):784-790.

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.

No

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

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

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

Not applicable

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.)

Not applicable

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.

Attachment:

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): The Joint Commission

Co.2 Point of Contact: JohnMarc, Alban, jalban@jointcommission.org, 630-792-5304-

Co.3 Measure Developer if different from Measure Steward: The Joint Commission

Co.4 Point of Contact: JohnMarc, Alban, jalban@jointcommission.org, 630-792-5304-

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.

Ann Doucette, PhD

Claremont Graduate University

Scott Dziengelski

National Association for Behavioral Healthcare

Frank A Ghinassi, PhD, ABPP (Chair)

President and CEO

Rutgers Health, University Behavioral Health Care

Richard Hermann, MD, MS

Tufts University School of Medicine, Tufts-NEMC

Karen E. Johnson, MSW

Universal Health Services, Inc.

Michael Lambert, PhD

Professor

Brigham Young University

Kathleen McCann, RN, PhD

National Association for Behavioral Healthcare

Dr. John Oldham, MD

Baylor College of Medicine

Lucille M Schacht, PhD, CPHQ

NRI, Inc

The Technical Advisory Panel (TAP) met and identified domains for measurement, endorsed the measurement framework and identified extant measures. After measures were received and evaluated by Joint Commission staff, the TAP met to review the measures and recommend candidate measures to move forward for public comment. Following public comment, the TAP reviewed the comment and recommended a set of measures to move forward for pilot testing. After pilot testing was completed, the TAP reviewed the pilot test results and recommended revisions to the measures for the final measure set.

The TAP remains engaged with The Joint Commission and meets on an as needed basis to offer consultation or to suggest updates relative to guideline changes/recommendations.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2008

Ad.3 Month and Year of most recent revision: 01, 2019

Ad.4 What is your frequency for review/update of this measure? Biannual

Ad.5 When is the next scheduled review/update for this measure? 07, 2019

Ad.6 Copyright statement: No royalty or use fee is required for copying or reprinting this manual, but the following are required as a condition of usage: 1) disclosure that the Specifications Manual is periodically updated, and that the version being copied or reprinted may not be up-to-date when used unless the copier or printer has verified the version to be up-to-date and affirms that, and 2) users participating in Joint Commission accreditation, including ORYX® vendors, are required to update their software and associated documentation based on the published manual production timelines.

Ad.7 Disclaimers:

Ad.8 Additional Information/Comments: Recent revision is dated January 1, 2019. This represents the date the specifications go into effect. The specifications were published in October 2018.