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 #: 1922

Measure Title: HBIPS-1 Admission Screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths Completed

Measure Steward: The Joint Commission

Brief Description of Measure: The proportion of patients, age greater than and equal to 1 year, admitted to a hospital-based inpatient psychiatric setting who are screened within the first three days of hospitalization for all of the following: risk of violence to self or others, substance use, psychological trauma history and patient strengths.

Developer Rationale: Evidence exists that there is a high prevalence of co-occurring substance use disorders as well as history of trauma among persons admitted to acute psychiatric settings. Professional literature suggests that these factors are under-identified yet integral to current psychiatric status and should be assessed in order to develop appropriate treatment (Ziedonis, 2004; NASMHPD, 2005). Similarly, persons admitted to inpatient settings require a careful assessment of risk for violence and the use of seclusion and restraint. Careful assessment of risk is critical to safety and treatment. Effective, individualized treatment relies on assessments that explicitly recognize patients´ strengths. These strengths may be characteristics of the individuals themselves, supports provided by families and others, or contributions made by the individuals´ community or cultural environment (Rapp, 1998). In the same way, inpatient environments require assessment for factors that lead to conflict or less than optimal outcomes.

As stated above, recent literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.Evidence exists that there is a high prevalence of co-occurring substance use disorders as well as history of trauma among persons admitted to acute psychiatric settings. Professional literature suggests that these factors are under-identified yet integral to current psychiatric status and should be assessed in order to develop appropriate treatment (Ziedonis, 2004; NASMHPD, 2005). Similarly, persons admitted to inpatient settings require a careful assessment of risk for violence and the use of seclusion and restraint. Careful assessment of risk is critical to safety and treatment.

Effective, individualized treatment relies on assessments that explicitly recognize patients´ strengths. These strengths may be characteristics of the individuals themselves, supports provided by families and others, or contributions made by the individuals´ community or cultural environment (Rapp, 1998). In the same way, inpatient environments require assessment for factors that lead to conflict or less than optimal outcomes.

The literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.As stated above, recent literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.

Numerator Statement: Psychiatric inpatients with admission screening within the first three days of admission for all of the following: risk of violence to self or others; substance use; psychological trauma history; and patient strengths

Denominator Statement: Psychiatric inpatient discharges

Denominator Exclusions: • Patients for whom there is an inability to complete admission screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths within the first three days of admission due to the patient's inability or unwillingness to answer screening questions

• Patients with a Length of Stay = or less than 3 days or = or greater than 365 days

Measure Type: Process

Data Source: Electronic Health Records, Paper Medical Records

Level of Analysis: Facility, Other

IF Endorsement Maintenance – Original Endorsement Date: Mar 04, 2014 **Most Recent Endorsement Date:** Mar 04, 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**

Summary of prior review in 2013

- Evidence from the 2013 submission included recommendations from the American Psychiatric Association (APA) Practice Guideline for the Psychiatric Evaluation of Adults (2006). The Standing Committee agreed that the measure is important and will have a high impact due to the prevalence and burden of violence, substance use, trauma, and suicide. Some Committee members previously questioned the directness of the evidence to the measure focus, but ultimately agreed the correlation was sufficient.

Changes to evidence from last review

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

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

Updates:

- The developer provides a logic model and updates guidelines – APA Practice Guidelines for the Psychiatric Evaluation of Adults Third Edition (2016).
- The logic model shows screening forming the basis of a treatment plan which will decrease the chance of psychiatric relapse, lead to improved treatment and medication adherence, and reduce the cost of ongoing recovery.
- Guideline statements related to each of the specific screening elements are provided.
    - Of the measure's five screening elements "evaluation of patient strengths" is less directly linked to the recommendation statements, yet still cited in the Guideline as a core activity of an initial psychiatric evaluation.
    - All guidelines statements are either graded 1C or 2C (C representing the strength of the supporting research evidence; C represents a low rating.), though difficulties studying assessment approaches in RCTs are noted.
- While the QCC of the literature is generally provided and importance of including these elements in the initial psychiatric screening is reflected by inclusion in guidelines, the effect of these screening components on patient outcomes is not specifically referenced in the submission.

Questions for the Committee:

- The evidence provided by the developer is updated and directionally the same compared to that for the previous NQF review. The Committee can accept the previous Evidence rating or choose to revote on Evidence.
- Does the evidence provided support the link between these screening elements and desired health outcomes? Is there a strong enough relationship between this measure and the desired outcomes?
- Are the relatively low grades of the evidence that support each of the recommendations concerning?

Guidance from the Evidence Algorithm

- Process measure based on systematic review (Box 3) → QQC presented (Box 4) → Quantity: moderate; Quality: moderate; Consistency: moderate (Box 5) → Moderate (Box 5b) → Moderate
- The highest possible rating is high.

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

RATIONALE:

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.

- Performance data from hospitals that was directly submitted to the Joint Commission from 2009-2018 is provided. The number of hospitals reporting data ranged from 300 to 1082 per year.
    - Mean performance for 2018 was 93.7% (Std. dev. 0.13). Scores were lowest in 2009 at 87% and peaked at 95.2% in 2012. Deciles also provided. In 2018, 20th percentile was 92.38% and 70th percentile was 99.85.
- Data shows relatively narrow performance gap – high performance rates across hospitals with some variation.

Disparities

- The literature does not mention disparities related to admission screening.
- The literature supports certain groups as the most vulnerable groups at higher risk of suicide, more likely to be violent, and have SUD.
- Performance rates by gender, ethnicity, race, and age categories did not show strong evidence of disparities for this measure. Rates were slightly lower for those 18-64 years and 65 years and above versus 1-12 and 13-17 years and American Indians and Pacific Islanders versus other race groups.

Questions for the Committee:

- Does the performance data demonstrate a gap in care that warrants a national performance measure?

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

RATIONALE:

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

1a. Evidence
Comments:
**The Measure is important and has the potential for high impact. The evidence appears directional. The developer submitted the updated APA guidelines (2016) and relate to each of the 5 screening elements,

although evaluation of patient strengths seems less clearly defined to me. The consensusgrading of 1C or 2C, low in “strength of supporting research evidence” is attributed to "the difficulties in studying psychiatric assessment" but for the most part the benefits are viewed as out waying any harms.sthe harms approaches in controlled studies

1b. Performance Gap

Comments:

**There is a narrow performance gap with some variation across facilities. There is only modest differences based on demographics. even so, I recommend a rating of moderate given the long standing clinical importance of the screening questions in this measure.

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**

Evaluators: NQF Staff

Evaluation of Reliability and Validity:

- Reliability testing results from a sample of 191 patient records indicated a high agreement rate (all >97.9%) for each data element in the numerator and denominator.
- Validity testing at the score level indicated a slight positive correlation (0.13857, p=0.0002) between the measure and HBIPS-5: Multiple Antipsychotic Medications at Discharge with Appropriate Justification.
- An analysis of meaningful differences showed improvements over time, however, with a relatively narrow performance gap.
- Exclusions analysis indicated that in 2017, 16.8% of patients met the Length of Stay exclusion.

- The measure is not risk adjusted, but results are stratified by age group.

Questions for the Committee regarding reliability:

- Is the data element testing provided (e.g., elements tested, methods) convincing of the measure's reliability?
- The staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions)?
- Are the results of the score-level analysis strong enough to support the measure's validity?

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: 1922

Measure Title: HBIPS-1 Admission Screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths Completed

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.**

RELIABILITY: TESTING

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

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

   ☐ **Yes** ☐ **No**
6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   - Data element reliability testing was performed.
   - Ten out of 487 hospitals in the population were randomly sampled, using a stratified sampling methodology to represent the three bed size and three ownership categories.
   - Data from 191 patient cases from a sample of facilities was re-abstracted by trained Joint Commission Staff and agreement rates were calculated for each data element.
7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

Data Elements with a Mismatch*	Total Numerator	Total Denominator	Agreement Rate
Numerator Data Elements			
Patient Strengths	187	191	97.9%
Psychological Trauma History	189	191	98.9%
Substance Use	188	191	98.4%
Violence Risk to Others	187	191	97.9%
Violence Risk to Self	188	191	98.4%
Denominator Data Elements			
Admission Date	191	191	100%
Birthdate	191	191	100%
Discharge Date	191	191	100%
ICD-9-CM Other Diagnosis Code**	191	191	100%
ICD-9-CM Principal Diagnosis Code**	191	191	100%
Psychiatric Care Setting	191	191	100%

** No cases were excluded for the reliability testing.*
*** 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.*

   - Exclusion data elements (1. Unable to complete admission screen due to patient's inability/unwillingness OR previous admission to psychiatric unit during the hospitalization and 2. Length of stay <3 days or >365 days) are not specifically referenced in the testing results.
   - The developer suggests the inability to complete admission screening is captured by an allowable value in the individual screening data elements and the length of stay exclusion is captured by the data elements Discharge Date minus Admission Date.
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.**

    - Does the Committee agree that the data element testing method and results, including the description of how exclusions were tested, support the measure's reliability?
    - Additional detail regarding the thoroughness of each of the five screening assessments would be useful to present.
        - For example, did providers get enough functional information about patients' strengths, or their substance use and violence history, or did they quickly ask a question and mark a checkbox?
        - The developer does provide descriptors of the documentation required for a "yes" to be abstracted for each of the five numerator data elements.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    - Since the measure allows for 3 days to complete the admission screening, patients with a LOS <=3 days are excluded. Is 3 days the appropriate amount of time to allow for admission screening (for all of the required elements) to occur? Approximately 17% of patients are excluded due to a LOS <=3 days or >=365 days.

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

    Submission document: Testing attachment, section 2b4.

    - Target analysis and performance scores by percentile are provided. Results generally indicate that most providers achieve high performance on this measure (2018 mean > 0.93, std deviation 0.13)
    - The developer also states that "Employing a longitudinal logistic regression model with the hospital as a random effect yields a significant improvement of rates over time ($p<0.0001$)".
        - Results seem to suggest that between 2009 and 2018 the odds of screening increased by 18.3%, after adjusting for hospital as a random effect. This suggests an annual increase of

1-2% per year, changes that roughly correspond to the annualized data presented in section 1b.2.

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

15. **Please describe any concerns you have regarding missing data.**
Submission document: Testing attachment, section 2b6.
None.

16. **Risk Adjustment**

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

16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**
☐ Yes ☐ No ☐ Not applicable

16c. **Social risk adjustment:**

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

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

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

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**

- The measure is stratified by age groups: Children (1 through 12 years), Adolescents (13 through 17 years) , Adults (18 through 64 years), Older Adults (65 years or greater).
- Additional information regarding the stratification approach is not provided and the developer does not refer to this as risk-adjustment.

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

- A correlation test between this measure and the other HBIPS measures (*0640: HBIPS-2: Hours of physical restraint use, 0641: HBIPS-3: Hours of seclusion use, 0560: HBIPS 5: Patients discharged on*

multiple antipsychotic medications with appropriate justification) performed using 2017 data from all hospitals submitting HBIPS measure data to the Joint Commission.

- The developer hypothesized that this measure *and 0560: HBIPS 5: Patients discharged on multiple antipsychotic medications with appropriate justification* would be correlated the most closely, with weaker correlations with *0640: HBIPS-2: Hours of physical restraint use* and *0641: HBIPS-3: Hours of seclusion use*.
- Face validity was previously assessed via survey and focus group of pilot hospitals. Details of the process for establishing face validity and voting results from the +-are not provided. The developer notes feedback from users was previously used to update the specifications and that queries from users have significantly decreased over the past 3 years.

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

    Submission document: Testing attachment, section 2b2.3

    - Tests for correlations between this measure and the remaining HBIPS measures (HBIPS-2, HBIPS-3, HBIPS-5) are -0.00313 (p=0.9328), -0.00875 (p=0.8144), and 0.13857 (p=0.0002), respectively.
    - Results indicate a slight positive correlation between this measure and HBIPS-5 and no statistically significant correlations between this measure and HBIPS-2 and HBIPS-3.

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 <u>are</u> threats to validity and/or relevant threats to validity were <u>not assessed OR</u> 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 <u>is required</u>; 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.**

    - Score-level testing indicated a weak but significant association between the measure and HBIPS-5.
    - *0560: HBIPS 5: Patients discharged on multiple antipsychotic medications with appropriate justification* is an external standard of quality, but not one that represents a strong and 'proximal-to-recovery' indicator.

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.**
    - Is the Committee comfortable with the reliability testing, even though exclusion data elements were not tested? Are the exclusions appropriate?
    - Is the strength of the correlation of the measure with *0560: HBIPS 5: Patients discharged on multiple antipsychotic medications with appropriate justification* convincing of the measure's validity?
    - Do analyses provided in section 2b4. demonstrate the ability to identify meaningful differences in performance between facilities?

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

2a1. Reliability – Specifications
Comments:
**The elements and logic are adequately defined and it appears the measure has been able to be consistently implemented.

2a2. Reliability – Testing
Comments:
**No.

2b1. Validity –Testing
Comments:
**No.

2b2-3. Meaningful Differences
Comments:
**The exclusions do not seem inappropriate forscreening in the acute inpatient setting

2b4-7. Threats to Validity
Comments:
**There is adequate face validity. Significant improvement of rates over time ($p<0.0001$) as well as ability to measure meaningful differences in quality between hospitals indicating performance gaps..

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.

- Data elements are generated during the care process and abstracted by another individual.
    - Some data elements are in defined fields in electronic sources.
- Developer reports most hospitals collect measure data via manual review of the paper medical record.

- Specifications are freely available for use, and user feedback indicate the specifications are generally easy to understand.

Questions for the Committee:

- Are there any concerns regarding the measure's feasibility or burden of capturing the necessary data?
- Is there any concern about whether screening completion and results are recorded honestly and accurately?

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

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility
Comments:
**I think all required elements would be routinely generated but whether as part of a paper chart vs EHR not sure. Both are used and seems to be operational feasible.

Criterion 4: Usability and Use

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

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

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

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

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

OR

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

Accountability program details

- Publicly reported in ORYX Performance Measurement Reporting Program and used in the Joint Commission's Hospital Accreditation Program.

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

- Measure rates are provided to hospitals quarterly and publicly available through the Joint Commission Quality Check website. Information includes the top 10 percentiles at the state and national level.
- A feedback system allows users to ask questions or provide feedback. Queries have significantly decreased the past few years. Feedback trends are analyzed and discussed with a Technical Advisory Panel as needed or during the twice-yearly maintenance process.

Additional Feedback: N/A

Questions for the Committee:

- Do performance results further the goal of high-quality, efficient healthcare?

Preliminary rating for Use: ☒ **Pass** ☐ **No Pass**

RATIONALE:

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

- Results of a binomial random effects model show statistically significant improvement from 2009 to 2018 ($P<0.001$) and an odds ratio estimate of time to be 1.183.
    - This suggests that rates of screening increased about 2% per year, after adjusting for "random effects of healthcare organization," (i.e., pure time effects are implied absent any within provider effects including quality differences). These findings are consistent with cited work (Ralinski et al., 2018).

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

- Results from a recent study that compared results on psychiatric performance measures among cohorts of hospitals with different characteristics demonstrated that all cohorts significantly improved across quarters for admission screening (Ralinski et al., 2018)

Potential harms

- None identified.

Additional Feedback:

- N/A

Questions for the Committee:

- Can the measure be used to improve the quality of comprehensive admission screening for patients admitted to the inpatient psychiatric setting?

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

RATIONALE:

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

4a1. Use - Accountability and Transparency
Comments:
**Yes, publicly reported in ORYX and is part of Joint Commission's Accreditation. User feedback and a TAP are utilized.

4b1. Usability – Improvement
Comments:
**Consistent performance of this screening set and findings incorporated into the treatment plan can clearly further the goal. There is no evidence of harm.

Criterion 5: Related and Competing Measures

Related or competing measures
Developer lists the following measure as related:
0104: Adult Major Depressive Disorder (MDD): Suicide Risk Assessment
0110*: Bipolar Disorder and Major Depression: Appraisal for alcohol or chemical substance use
0111*: Bipolar Disorder: Appraisal for risk of suicide
1365: Child and Adolescent Major Depressive Disorder (MDD): Suicide Risk Assessment
2152: Preventive Care and Screening: Unhealthy Alcohol Use: Screening & Brief Counseling
2599: Alcohol Screening and Follow-up for People with Serious Mental Illness
SUB-1* Alcohol Use Screening STEWARD: The Joint Commission
**not NQF-endorsed*

Other related measure identified by NQF:
2806: Pediatric Psychosis: Screening for Drugs of Abuse in the Emergency Department

Harmonization

- Developer notes the measures are not harmonized to the extent possible and points out differences in the measures: level of analysis (for 2599), screening focus area(s), and population diagnoses and age.
- This measure is not directly competing with other measures, but the Committee will discuss potential harmonization opportunities.

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

5. Related and Competing
Comments:
**Developer states the other NQF endorsed measures are not hospital based.

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 #: 1922

Corresponding Measures:

De.2. Measure Title: HBIPS-1 Admission Screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths Completed

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, admitted to a hospital-based inpatient psychiatric setting who are screened within the first three days of hospitalization for all of the following: risk of violence to self or others, substance use, psychological trauma history and patient strengths.

1b.1. Developer Rationale: Evidence exists that there is a high prevalence of co-occurring substance use disorders as well as history of trauma among persons admitted to acute psychiatric settings. Professional literature suggests that these factors are under-identified yet integral to current psychiatric status and should be assessed in order to develop appropriate treatment (Ziedonis, 2004; NASMHPD, 2005). Similarly, persons admitted to inpatient settings require a careful assessment of risk for violence and the use of seclusion and restraint. Careful assessment of risk is critical to safety and treatment. Effective, individualized treatment relies on assessments that explicitly recognize patients´ strengths. These strengths may be characteristics of the individuals themselves, supports provided by families and others, or contributions made by the individuals´ community or cultural environment (Rapp, 1998). In the same way, inpatient environments require assessment for factors that lead to conflict or less than optimal outcomes.

As stated above, recent literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.Evidence exists that there is a high prevalence of co-occurring substance use disorders as well as history of trauma among persons admitted to acute psychiatric settings. Professional literature suggests that these factors are under-identified yet integral to current psychiatric status and should be assessed in order to develop appropriate treatment (Ziedonis, 2004; NASMHPD, 2005). Similarly, persons admitted to inpatient settings require a careful assessment of risk for violence and the use of seclusion and restraint. Careful assessment of risk is critical to safety and treatment. Effective, individualized treatment relies on assessments that explicitly recognize patients´ strengths. These strengths may be characteristics of the individuals themselves, supports provided by families and others, or contributions made by the individuals´ community or cultural environment (Rapp, 1998). In the same way, inpatient environments require assessment for factors that lead to conflict or less than optimal outcomes.

The literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn

decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.As stated above, recent literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.

S.4. Numerator Statement: Psychiatric inpatients with admission screening within the first three days of admission for all of the following: risk of violence to self or others; substance use; psychological trauma history; and patient strengths

S.6. Denominator Statement: Psychiatric inpatient discharges

S.8. Denominator Exclusions: • Patients for whom there is an inability to complete admission screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths within the first three days of admission due to the patient's inability or unwillingness to answer screening questions

• Patients with a Length of Stay = or less than 3 days or = or greater than 365 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: Mar 04, 2014 **Most Recent Endorsement Date:** Mar 04, 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

1922_evidence_attachment_7.1_HBIPS1.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.

Yes

1a. Evidence (subcriterion 1a)

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 1922

Measure Title: Admission Screening

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 to determine whether admission screenings are completed for violence risk to self and others, substance use, psychological trauma history and patient strengths to help to develop an initial treatment plan that will incorporate these

findings and thereby promote treatment adherence and reduce the likelihood of psychiatric relapse and improve medication compliance, thus reducing the overall cost of ongoing recovery.

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

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**Title: Practice Guidelines for the Psychiatric Evaluation of Adults Third Edition** **Author: American Psychiatric Association Work Group on Psychiatric Evaluation** **Date: 2016** **Citation including page number: American Psychiatric Association (2016). Practice Guidelines for the Psychiatric Evaluation of Adults. Third edition. Arlington (VA): American Psychiatric Association. 164 pages.**

	URL: https://psychiatryonline.org/doi/pdf/10.1176/appi.books.9780890426760
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.	**Guideline I.** Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History **Guideline Statements** **Statement 1.** APA recommends (1C) that the initial psychiatric evaluation of a patient include review of the patient's mood, level of anxiety, thought content and process, and perception and cognition. **Statement 2.** APA recommends (1C) that the initial psychiatric evaluation of a patient include review of the patient's trauma history. **Statement 3.** APA recommends (1C) that the initial psychiatric evaluation of a patient include review of the following aspects of the patient's psychiatric treatment history: • Past and current psychiatric diagnoses • Past psychiatric treatments (type, duration, and, where applicable, doses) • Adherence to past and current pharmacological and nonpharmacological psychiatric treatments • Response to past psychiatric treatments • History of psychiatric hospitalization and emergency department visits for psychiatric issues[1] Assessment of psychiatric symptoms and psychiatric treatment history is by definition a core activity of an initial psychiatric evaluation. Other core activities include identifying the reason that the patient is presenting for evaluation and understanding the patient's background, relationships, life circumstances, and strengths and vulnerabilities. Each of these elements can be affected if a patient has been exposed to trauma. **Guideline II.** Substance Use Assessment **Guideline Statements** APA recommends (1C) that the initial psychiatric evaluation of a patient include assessment of the patient's use of tobacco, alcohol, and other substances (e.g., marijuana, cocaine, heroin, hallucinogens) and any misuse of prescribed or over-the-counter medications or supplements. **Guideline III.** Assessment of Suicide Risk **Guideline Statements**

	Statement 1. APA recommends (1C) that the initial psychiatric evaluation of a patient include assessment of the following: • Current suicidal ideas, suicide plans, and suicide intent, including active or passive thoughts of suicide or death • Prior suicidal ideas, suicide plans, and suicide attempts, including attempts that were aborted or interrupted • Prior intentional self-injury in which there was no suicide intent • Anxiety symptoms, including panic attacks • Hopelessness • Impulsivity • History of psychiatric hospitalization and emergency department visits for psychiatric issues • Current or recent substance use disorder or change in use of alcohol or other substances • Presence of psychosocial stressors (e.g., financial, housing, legal, school/occupational or interpersonal/relationship problems; lack of social support; painful, disfiguring, or terminal medical illness) • Current aggressive or psychotic ideas, including thoughts of physical or sexual aggression or homicide[2] • Mood, level of anxiety, thought content and process, and perception and cognition[3] • Past and current psychiatric diagnoses[3] • Trauma history[3] **Statement 2.** APA recommends (1C) that the initial psychiatric evaluation of a patient *who reports current suicidal ideas* include assessment of the following: • Patient's intended course of action if current symptoms worsen • Access to suicide methods, including firearms • Patient's possible motivations for suicide (e.g., attention or reaction from others; revenge, shame, humiliation, delusional guilt, command hallucinations) • Reasons for living (e.g., sense of responsibility to children or others, religious beliefs) • Quality and strength of the therapeutic alliance • History of suicidal behaviors in biological relatives **Statement 3.** APA recommends (1C) that the initial psychiatric evaluation of a patient *who reports prior suicide attempts* includes assessment

	of the details of each attempt (e.g., context, method, damage, potential lethality, intent). **Statement 4.** APA recommends (1C) that the clinician who conducts the initial psychiatric evaluation document an estimation of the patient's suicide risk, including factors influencing risk. **Guideline IV.** Assessment of Risk for Aggressive Behaviors **Guideline Statements** **Statement 1.** APA recommends (1C) that the initial psychiatric evaluation of a patient include assessment of the following: • Current aggressive or psychotic ideas, including thoughts of physical or sexual aggression or homicide • Prior aggressive or psychotic ideas, including thoughts of physical or sexual aggression or homicide • Past aggressive behaviors (e.g., homicide, domestic or workplace violence, other physically or sexually aggressive threats or acts) • Legal or disciplinary consequences of past aggressive behaviors • History of psychiatric hospitalization and emergency department visits for psychiatric issues • Current or recent substance use disorder or change in use of alcohol or other substances • Presence of psychosocial stressors • Exposure to violence or aggressive behavior, including combat exposure or childhood abuse • Past or current neurological or neurocognitive disorders or symptoms **Statement 2.** When it is determined during an initial psychiatric evaluation that the patient has aggressive ideas, APA recommends (1C) assessment of the following: • Impulsivity, including anger management issues • Access to firearms • Specific individuals or groups toward whom homicidal or aggressive ideas or behaviors have been directed in the past or at present • History of violent behaviors in biological relatives **Statement 3.** APA suggests (2C) that the clinician who conducts the initial psychiatric evaluation should document an estimation of risk of aggressive behavior (including homicide), including factors influencing risk.

	Guideline V. Assessment of Cultural Factors **Guideline Statements** **Statement 1.** APA recommends (1C) that the initial psychiatric evaluation of a patient include assessment of the patient's need for an interpreter. **Statement 2.** APA recommends (1C) that the initial psychiatric evaluation of a patient include assessment of cultural factors related to the patient's social environment. **Statement 3.** APA suggests (2C) that the initial psychiatric evaluation of a patient include assessment of the patient's personal/cultural beliefs and cultural explanations of psychiatric illness.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Each guideline statement is separately rated to indicate strength of recommendation and strength of supporting research evidence. Grades varied per guideline statements but as noted above in each statement, were either 1C or 2C. Evidence is graded with an alphabetical letter. C: rating for the "strength of supporting research evidence." C represents a low rating. The strength of supporting research evidence for these recommendations and suggestions is given rating C (low) because of the difficulties in studying psychiatric assessment approaches in controlled studies.
Provide all other grades and definitions from the evidence grading system	• High (denoted by the letter *A*) = High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect. • Moderate (denoted by the letter *B*) = Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate. • Low denoted by the letter *C*) = Low confidence that the evidence reflects the true effect. Further research is likely to change our confidence in the estimate of effect and is likely to change the estimate.

Grade assigned to the **recommendation** with definition of the grade	Grades varied per guideline statements but as noted above in each statement, were either 1C or 2C. Recommendations are graded using a number. 1: a "recommendation" that indicates confidence that the benefits of the intervention clearly outweigh harms. 2: a "suggestion" indicates uncertainty (i.e., the balance of benefits and harms is difficult to judge, or either the benefits or the harms are unclear).
Provide all other grades and definitions from the recommendation grading system	In accordance with the "Guideline Development Process," each final rating is a consensus judgment of the authors of the guidelines and is endorsed by the APA Board of Trustees. A "recommendation" (denoted by the numeral *1* after the guideline statement) indicates confidence that the benefits of the intervention clearly outweigh harms. A "suggestion" (denoted by the numeral *2* after the guideline statement) indicates uncertainty (i.e., the balance of benefits and harms is difficult to judge, or either the benefits or the harms are unclear). Each guideline statement also has an associated rating for the "strength of supporting research evidence."
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	**Quantity:** An initial search of MEDLINE was conducted in October 2010. This search yielded 250,981 articles. A second set of searches was conducted in October 2011. These searches yielded 32,895 articles in MEDLINE, 7,052 articles in PsycINFO, and 5,986 articles in the Cochrane database. All searches were done for the years from 1900 to the time of the search. A total of 5,073 articles met the broad inclusion criteria. The total number of studies that were agreed to have relevance to the PICOTS question for each guideline topic is as follows: 0 studies for Review of Psychiatric Symptoms, Trauma History, and Psychiatric Treatment History; 4 studies for Substance Use Assessment; 1 study for Assessment of Suicide Risk; 2 studies for Assessment of Risk for Aggressive Behaviors; 0 studies for Assessment of Cultural Factors; 3 studies for Assessment of Medical Health; 2 studies for Quantitative Assessment; 17 studies for Involvement of the Patient in Treatment Decision Making; and 0 studies for Documentation of the Psychiatric Evaluation.

	Quality: Included articles were those that pertained to a clinical trial (including a controlled or randomized trial), observational study, meta-analysis, or systematic review and were clinically relevant to psychiatric evaluation (i.e., relevant to any possible clinical question that might be addressed by potential APA practice guidelines). Excluded references included articles on nosology of psychiatric disorders, risk factors or associated features of specific disorders, potential etiologies of specific disorders, and course and prognosis of specific disorders.
Estimates of benefit and consistency across studies	**Benefits:** The focus of the measure is to evaluate all psychiatric inpatients to determine whether admission screenings are completed for violence risk to self and others, substance use, psychological trauma history and patient strengths. The benefit is the improvement in the identification of psychiatric signs and symptoms, psychiatric disorders (including substance use disorders), other medical conditions (that could affect the accuracy of a psychiatric diagnosis), and patients who are at increased risk for suicidal or aggressive behaviors. This information assessed during the initial psychiatric evaluation contributes in determination of an appropriate treatment plan. **Consistency:** Research on psychiatric assessment is complicated by multiple confounding factors, such as the interaction between the clinician and the patient or the patient's unique circumstances and experiences. For these and other reasons, the vast majority of topics related to psychiatric evaluation have relied on forms of evidence such as consensus opinions of experienced clinicians or indirect findings from observational studies rather than being based on research from randomized trials. Despite the difficulties in obtaining quantitative evidence from randomized trials, the body of evidence supports the initial screening of the psychiatric patient.
What harms were identified?	Potential harms of assessment were not a focus of the study but were determined to be minimal. Potential harms considered could include spending too much time on one assessment domain which may result in reducing time available to assess/document other, potentially more important findings of an evaluation.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	An update of the literature search was conducted in September 2014 using the same databases and search strategies used for the October 2011 search. These searches in September 2014 yielded 8,521 additional articles in MEDLINE, 1,980 additional articles in PsycINFO, and 1,310 additional articles in the Cochrane database. After duplicates were eliminated, 11,644 abstracts were screened for relevance by two individuals (L.F., J.Y.). A total of 65 additional references met the broad inclusion criteria, and of these, 1 study was relevant to Quantitative Assessment.

From previous submission:

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**Title: Practice Guideline for the Psychiatric Evaluation of Adults Second Edition** **Author: American Psychiatric Association Work Group on Psychiatric Evaluation** **Date: June 2006** **Citation, including page number: American Psychiatric Association (APA). Practice guideline for the psychiatric evaluation of adults. 2nd ed. Washington (DC): American Psychiatric Association (APA); 2006 Jun. 62 p. [302 references]** **URL:** **https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/psychevaladults.pdf** National Guideline Clearinghouse: http://www.guideline.gov/content.aspx?id=9317&search=psychiatric+evaluation 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.	Domains of the Clinical Evaluation General psychiatric evaluations involve a systematic consideration of the broad domains described in this guideline and vary in scope and intensity. Table 1 summarizes the domains. The intensity with which each domain is assessed depends on the purpose of the evaluation and the clinical situation. An evaluation of lesser scope may be appropriate when its purpose is to answer a circumscribed question. Such an evaluation may involve a particularly intense assessment of one or more domains especially relevant to the reason for the evaluation. Across all domains, evaluations are generally based on three sources of information: 1) observation and interview of the patient; 2) information from others (e.g., family, significant others, case managers, other clinicians [including the patient´s primary care physician]) that corroborates, refutes, or elaborates on the patient´s report; and 3) medical records. An awareness of how people report current symptoms and events is important to the clinical assessment process. In considering the information obtained, the patient´s current mental state is relevant. Mistakes in comprehension, recall, and expression may also lead to erroneous reporting of information (F). A. Reason for the Evaluation The purpose of the evaluation influences the focus of the examination and the form of documentation. The reason for the evaluation usually includes (but may not be limited to) the chief complaint of the patient. It should be elicited in sufficient detail, including the patient´s words, to permit an understanding of the duration of the complaint and the patient´s specific goals for the evaluation. If the symptoms are of long

	standing, the reason for seeking treatment at this specific time is relevant; if the evaluation was occasioned by a hospitalization, the reason for the hospitalization is also relevant. If the patient did not initiate the evaluation, the reason another individual or entity may have requested or required it should be noted. The opinions of other parties, including family, can also assist in establishing a reason for evaluation. Under some circumstances (e.g., with psychotic or uncommunicative patients), input from others may be crucial. B. History of the Present Illness The history of the present problem or illness is a chronologically organized history of recent exacerbations or remissions and current symptoms or syndromes. These may involve descriptions of worries, changes in mood, suspicions, preoccupations, delusions, or hallucinatory experiences as well as recent changes in sleep, appetite, libido, concentration, memory, or behavior, including suicidal or aggressive behaviors. Information gathered on the pertinent positive and pertinent negative features of the history of present illness will vary with the patient´s presenting symptoms or syndrome. Temporal features relating to the onset or exacerbation of symptoms may also be relevant (e.g., onset after use of exogenous hormones, herbal products, or licit or illicit substances; variation in symptoms with the menstrual cycle; postpartum onset). Also pertinent are factors that the patient and other informants believe to be precipitating, aggravating, or otherwise modifying the illness. Available details of previous treatments and the patient´s response to those treatments will be delineated as part of the history of present illness. If the patient was or is in treatment with another clinician, the effects of that relationship on the current illness, including transference and countertransference issues, are considered. Input from members of a clinical team who care for the patient can be very helpful (Section IV.A.6). For patients seen on medical-surgical units, it is important to consider the history of both the present medical-surgical illness and the present psychiatric illness (G). C. Past Psychiatric History The past psychiatric history includes a chronological summary of all past episodes of mental illness, including substance use disorders, and treatment. The summary includes prior hospitalizations; suicide attempts, aborted suicide attempts, or other self-destructive behavior; psychiatric syndromes not formally diagnosed at the

	time; previously established diagnoses; treatments offered; and responses to and satisfaction with treatment. With respect to psychotherapy, it is important to ascertain the type (e.g., psychodynamic, cognitive, behavioral, supportive), format (e.g., group, individual, couple), frequency, duration, patient´s perception of the alliance, and adherence. With respect to medications, the dose, efficacy, side effects, treatment duration, and adherence are important to ascertain while understanding that reporting errors are more likely to occur when treatment involved more than one medication (G). With respect to other somatic therapies such as electroconvulsive therapy, information on the number of treatment sessions, treatment course duration, technical parameters, efficacy, and side effects is similarly useful to obtain. When past medical records are available and readily accessible, it is important that they be consulted for ancillary information. The chronological summary also delineates the most recent periods of stability as well as episodes when the patient was functionally impaired or seriously distressed by mental or behavioral symptoms, even if no formal treatment occurred. Such episodes frequently can be identified by asking the patient about the past use of psychotropic medications prescribed by other clinicians and otherwise unexplained episodes of social or occupational disability. D. History of Substance Use The psychoactive substance use history includes past and present use of both licit and illicit psychoactive substances, including but not limited to alcohol, caffeine, nicotine, marijuana, cocaine, opiates, sedative-hypnotic agents, stimulants, solvents, MDMA (methylenedioxymethamphetamine), androgenic steroids, and hallucinogens (G). Relevant information includes the quantity and frequency of use and route of administration; the pattern of use (e.g., episodic versus continual, solitary versus social); functional, interpersonal, or legal consequences of use; tolerance and withdrawal phenomena; any temporal association between substance use and other present psychiatric illnesses; and any self-perceived benefits of use. It is also important to inquire about prior treatments for substance use disorders as well as about periods of abstinence, including their duration, recentness, and factors that aided in sobriety or contributed to relapse. Obtaining an accurate substance use history often involves a gradual, nonconfrontational approach to inquiry that involves asking multiple questions to seek

	the same information in different ways and using slang terms for drugs, patterns of use, and drug effects. Patients are particularly likely to underestimate their level of substance abuse and their related functional impairments; corroboration by other family members is useful when possible. It is also helpful to inquire about patterns of substance use by others within the family or living constellation. For more extensive discussion of the assessment of substance use, abuse, and dependence, the reader is referred to the Center for Substance Abuse Treatment´s Assessment and Treatment of Patients With Coexisting Mental Illness and Alcohol and Other Drug Abuse (G) and APA´s Practice Guideline for the Treatment of Patients With Substance Use Disorders (G). E. General Medical History The general medical history includes available information on known general medical illnesses (e.g., hospitalizations, procedures, treatments, and medications), allergies or drug sensitivities, and undiagnosed health problems that have caused the patient major distress or functional impairment. This includes history of any episodes of important physical injury or trauma; sexual and reproductive history; and any history of endocrinological, infectious (including but not limited to HIV, tuberculosis, and hepatitis C) (G), neurological disorders, sleep disorders (including sleep apnea), and conditions causing pain and discomfort. Of particular importance is a specific history regarding diseases and symptoms of diseases that have a high prevalence among individuals with the patient´s demographic characteristics and background—for example, infectious diseases in users of intravenous drugs or pulmonary and cardiovascular disease in people who smoke. Information regarding all current and recent medications, including hormones (e.g., birth control pills, androgens), over-the-counter medications, herbal supplements, vitamins, complementary and alternative medical treatments, and medication side effects, is part of the general medical history. With all aspects of the general medical history, obtaining corroborating information (e.g., from medical records, treating clinicians, family) can be helpful, since ordinary errors in comprehension, recall, and expression can lead to errors in patient reports (F). F. Developmental, Psychosocial, and Sociocultural History The personal history reviews the stages of the patient´s life, with special attention to perinatal events, delays in physical or psychological development, formal

	educational history, academic performance, and patterns of response to normal life transitions and major life events, including parental loss or divorce; physical, emotional, or sexual abuse; and other trauma such as exposure to political repression, war, or a natural disaster (G). The childhood and adolescent history of risk factors for later psychiatric disorders (Table 2) may also be relevant. History of adaptive skills and strengths to overcome challenges is also relevant. The patient´s capacity to maintain stable and gratifying interpersonal relationships should be noted, including the patient´s capacities for attachment, trust, and intimacy. A sexual history is obtained and includes consideration of sexual orientation and practices, past sexual experiences (including unwanted experiences), and cultural beliefs about sex (G). The psychosocial history also determines the patient´s past and present levels of interpersonal functioning in family and social roles (e.g., marriage, parenting) (C, F and C). This includes a delineation of the patient´s history of marital and other significant relationships. For patients with children (including biological, foster, adopted, or stepchildren), the psychosocial history will include information about these individuals and their relationship to the patient. As part of the psychosocial history, past or current stressors are assessed and include the categories on axis IV of DSM-IV-TR: primary support group, social environment (e.g., discrimination and acculturation), education, occupation, housing, economic status, and access to health care. Specific information obtained in evaluating psychosocial stressors may include details about patients´ living arrangements, access to transportation, sources of income, insurance or prescription coverage, and past or current involvement with social agencies. Assessment of the patient´s self-care functioning may also include consideration of exercise behavior and money management skills, including gambling behavior. The sociocultural history delineates the patient´s migration history and past and current sociocultural context of supports and stressors as well as other important cultural and religious influences on the patient´s life (G). Emphasis is given to relationships, both familial and nonfamilial, and to religion and spirituality that may give meaning and purpose to the patient´s life and provide support, as described in the DSM-IV-TR Outline for Cultural Formulation (described in more detail in Section IV.B.1.a).

	Patients present for a psychiatric evaluation with their own interests, preferences, and value systems pertaining to health care practice, and these are another important part of the sociocultural history. They may involve cultural factors and explanatory models of illness that affect attitudes, expectations, and preferences for professional and popular treatments, as described in the DSM-IV-TR Outline for Cultural Formulation and the 2004 Core Competencies of the American Board of Psychiatry and Neurology (G). Also important to the assessment and treatment process are other domains such as existential, moral, and interpersonal values and social influences such as school, church, work, and community or other agencies. Attending to these factors plays a crucial role in developing a therapeutic alliance, negotiating a treatment plan, determining the outcome criteria for successful treatment, and enhancing treatment adherence. Belief systems may also influence the handling of privacy and confidentiality during the evaluation as well as influence the type and amount of information disclosed as part of any informed consent process. In addition, patients´ value systems are relevant to clinical considerations at important life transitions (e.g., job and career transitions, marital transitions, genetic counseling before or during pregnancy, end-of-life planning). G. Occupational and Military History The occupational history describes the sequence and duration of jobs held by the patient, reasons for job changes, and the patient´s current or most recent employment, including quality of work relationships and whether current or recent jobs have involved shift work, a noxious or perilous environment, exposure to hazardous materials, unusual physical or psychological stress, or injury or exposure to trauma while in the military or hazardous occupations (e.g., fire and rescue, law enforcement). Work skills and strengths are noted, as well as the quality of the patient´s relationships with co-workers and work supervisors. Past or current experience with the workers´ compensation system and patterns of recovery or disability following episodes of illness are also determined (G, F, B and. B). When appropriate, a history of preparation for and adjustment to retirement is included. Relevant data about military experience include volunteer, recruited, or draftee status; reasons for rejection at time of enlistment (if relevant); combat exposure (if any); awards; disciplinary actions; and discharge status.

	H. Legal History The legal history includes a description of past or current involvement with the legal system (G). This may include interactions with the police without formal arrest as well as involvement with the juvenile or criminal justice system (e.g., arrests, detentions including jail or prison confinement). Individuals may be on probation or parole or may have pending court appearances or active warrants for arrest that will influence treatment planning. A history of legal problems relating to aggressive behaviors or occurring in the context of substance intoxication is similarly relevant. Other past or current interactions with the court system (e.g., family court, workers´ compensation, civil litigation, court-ordered psychiatric treatment) may serve as significant stressors for the patient and are also important to address (G).
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	**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. **Grade Assigned to the Body of Evidence:** Not Applicable
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:** Varies by domain
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 evidence base for these practice guidelines is derived from two sources: research studies

	and clinical consensus. Where gaps exist in the research data, evidence is derived from clinical consensus, obtained through extensive review of multiple drafts of each guideline. In addition, each reference at the end of the original guideline document is followed by a letter code in brackets that indicates the nature of the supporting evidence, as follows: • [A] Double-blind, randomized clinical trial. A study of an intervention in which subjects are prospectively followed over time; there are treatment and control groups; subjects are randomly assigned to the two groups; both the subjects and the investigators are blind to the assignments. • [A] Randomized clinical trial. Same as above but not double-blind. • [B] Clinical trial. A prospective study in which an intervention is made and the results of that intervention are tracked longitudinally; study does not meet standards for a randomized clinical trial. • [C] Cohort or longitudinal study. A study in which subjects are prospectively followed over time without any specific intervention. • [D] Control study. A study in which a group of patients and a group of control subjects are identified in the present and information about them is pursued retrospectively or backward in time. • [E] Review with secondary analysis. A structured analytic review of existing data, e.g., a meta-analysis or a decision analysis. • [F] Review. A qualitative review and discussion of previously published literature without a quantitative synthesis of the data. • [G] Other. Textbooks, expert opinion, case reports, and other reports not included above.
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) to include screening for violence risk to self and others, substance use, psychological trauma history and patient strengths during the psychiatric evaluation. These risk factors have been identified to have a high degree of co-occurrence in psychiatric inpatients. The focus of both the performance measure and the body of evidence supports the need for admission screening for violence risk to self and others, substance use, psychological trauma history and patient strengths.

	Quantity: Relevant literature supporting psychiatric screening was identified by the American Psychiatric Association through a computerized search of MEDLINE, using PubMed, for the period from 1994 to 2005. The search strategy (psychiatric assessment OR psychiatric assessments OR psychiatric emergencies OR psychiatric emergency OR psychiatric evaluation OR psychiatric evaluations OR psychiatric histories OR psychiatric history OR psychiatric interview OR psychiatric interviewing OR psychiatric interviews OR psychological assessment OR psychological assessments OR psychological evaluation OR psychological interview OR mental status examination OR mental status examinations OR psychiatric rating) OR (mental disorders/diagnosis AND [laboratory findings OR laboratory techniques OR laboratory test OR laboratory tests OR radiograph OR radiographic OR radiography OR x ray OR imaging OR MRI OR tomography OR physical exam OR physical examination OR interview OR interviewing OR history taking OR evaluation OR assessment]) yielded 19,429 references, of which 7,894 were published between 1994 and 2005 in English and had associated abstracts. An additional search on history taking AND (psychiatric OR sexual OR occupational OR social OR psychosocial) yielded 1,927 references, with 731 of these published with abstracts in English between the years 1994 and 2005. **Quality:** The quality of evidence supporting the value of screening patients for violence risk to self and others, substance use, psychological trauma history and patient strengths completed is moderate. It is noteworthy to examine the fact that randomized control trials cannot be conducted, as one cannot randomly select patients who receive the screening and those who do not given the serious consequences of under detection and treatment of violence risk to self and others, substance use, psychological trauma history and patient strengths. As noted above, the American Psychiatric Association has had guidelines in place since 1995 addressing the key aspects of a psychiatric evaluation which include screening for violence risk to self and others, substance use, psychological trauma history and patient strengths. In spite of the fact that all studies reviewed were either

	retrospective or prospective cohort studies, no study design flaws were noted. **Summary of Controversy/Contradictory Evidence:** There is no documented evidence regarding controversy related to admission screening for violence risk to self and others, substance use, psychological trauma history and patient strengths. A review of recent studies also supports the use of quality improvement interventions to further increase compliance with the admission screening for violence risk to self and others, substance use, psychological trauma history and patient strengths. No position against the importance to screen for risk of violence to self or others, substance use, psychological trauma history and patient strengths was identified in the literature. **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?** **1c.25** Quantity: High **1c.26** Quality: Moderate **1c.27** Consistency**:** High
Estimates of benefit and consistency across studies	**Benefit:** The initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths is to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The resultant reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately result in substantial savings in health care costs. Focusing on patient strengths instead of problems during the screening process will also empower the patient to embrace the ongoing recovery model of treatment. No harms to the patient receiving an initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths was found during the literature review. **Consistency:** The body of evidence consistently supports the benefit of routine initial screening for risk of violence to self or others, substance use, psychological trauma history and

	patient strengths. Initial admission screening is consistently mentioned as an important step that is necessary to perform in order to develop the appropriate treatment plan in studies on risk of violence to self or others, substance use, psychological trauma history and patient strengths. No position against the importance to screen for risk of violence to self or others, substance use, psychological trauma history and patient strengths was identified in the literature.
What harms were identified?	No harms to the patient receiving an initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths was found during the literature review.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	An updated MEDLINE search, using PubMed, of all literature related to psychiatric evaluation topics was done for the years 2005 to 2010. In addition, a search of MEDLINE and PsycInfo for randomized controlled trials and meta-analyses for the years 1966 to September 2011 was done, using the EBSCO Host database, again on topics related to psychiatric evaluation. Search results (approximately 95,000 references) were screened for relevance by a single researcher, and a subset of the results (approximately 5,000 articles) were screened by a second researcher, demonstrating >90% concordance on ratings.

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. (2004) Practice guidelines for the treatment of patients with substance use disorders. Arlington (VA): American Psychiatric Association. http://www.guideline.gov/summary/summary.aspx?doc_id=9316
- Arseneault, L., Moffit, T., Caspt, A. et al. (2000). Mental disorders and violence in a total birth cohort: results from the Dunedin Study. Arch Gen Psychiatry. 57:979-986.
- Baldessarini, R.J. & Tondo, L. (2001). Suicide: causes and clinical management. New York, NY: McGraw-Hill Press.
- Beck, J., & Van der Kolk, B. (1987). Reports of childhood incest and current behavior of chronically hospitalized women. Am J Psychiatry, 144, 1474.
- Centers for Disease Control and Prevention. (2009). Deaths: Final data for 2006. National Vital Statistics Report, 57.1.
- Cusack, K. J., Frueh, B. C., Bray, KT. (2004). Trauma history screening in a community mental health center, Psychiatric Services, 155:157-162.
- Friedman, R.A. (2006). Violence and mental illness. N Engl J Med. 335, 20. 2064-2066.
- Goldsmith, S.K., Pellmar, T.C., Kleinman, A.M., & Bunney, W.E. (Eds.) (2002). Reducing suicide: a national imperative. Washington DC: The National Academies Press.
- Goodman, L.A., Salyers, M.P., Mueser, K.T., Rosenberg, S.D., Swartz, M.S., Essock, S.M., Oscher, F.C., Butterfield, M.I., Swanson, J. (2001). Recent victimization in women and men with severe mental illness: Prevalence and correlates. Journal of Traumatic Stress, 14, 615-632.
- Hanson, T.C., Hesselbrock, M., Tworkowski, S.H., Swan, S. (2002). The prevalence and

  management of trauma in the public domain: An agency and clinician perspective. The
- Journal of Behavioral Health Services and Research, 13:40, 365-380.
- Harwood, H.J., (2000). Updating Estimates of the Economic Costs of Alcohol Abuse in the United States. National Institute on Alcohol Abuse and Alcoholism. Office of National Drug Control Policy. The Economic Costs of Drug Abuse in the United States: 1992–2002. Washington, DC: Executive Office of the President (Publication No. 207303), 2004 Retrieved on December 30, 2011 at: http://pubs.niaaa.nih.gov/publications/economic-2000.
- Hiday, V.A., Swartz, M., Swanson, J., Borum, R., & Wagner, H.R. (1999). Criminal Victimization of Persons with Severe Mental Illness. Psychiatric Services, 50:1.
- Hirschfeld, R & Russell, JM. (1997).Assessment and treatment of suicidal patients. N Engl J Med. 337:910-915.
- Johnson, ME. (2004). Violence on inpatient psychiatric units: State of the science. Journal of the American Psychiatric Nurses Association, 10(3). 113-121.
- Kessler, R. C., Sonnega, A., Bromet, E., et al. (1995). Posttraumatic Stress Disorder in the national co-morbidity survey. Archives of General Psychiatry, 52:1048-1060.
- Linehan, MM, Comtois, KA & Ward-Ciesielski, EF. (2011). Assessing and managing risk with suicidal individuals, Cognitive and Behavioral Practice, doi:10.1016/j.ebpra. 2010.11.008.
- Lyons, J.S., Uziel-Miller, N.D., Reyes, F., Sokol, P.T. (2000). Strengths of children and adolescents in residential settings: Prevalence and associations with psychopathology and discharge placement. Journal of the American Academy of Child & Adolescent Psychiatry, 39(2): 176-181.
- Mallin, R., Slott, K., Tumblin, M. & Hunter, M. (2002). Detection of substance use disorders in patients presenting with depression. Substance Abuse. 115-120.
- McNiel, DE, Chamberlain, JR, Weaver, CM, Hall, SE, Fordwood SR & Binder RL. (2008). Impact of clinical training on violence risk assessment. Am J Psychiatry. 165:2, 195-200.
- Mokdad, A.H., Marks, J.S., Stroup, D.S., & Geberding, J.L. (2004). Actual causes of death in the United States, 2000. JAMA.291:128-1245.

- Morgan, H.G. & Stanton, R. (1997). Suicide among psychiatric inpatients in a changing clinical scene. Suicide ideation as a paramount index of short-term risk. British Journal of Psychiatry, 171, 561-563.
- Mullen, P.E., Burgess, P., Wallace, C. et al. (2000). Community care and criminal offending in schizophrenia. Lancet. 355:614-617.
- Mueser, K. T., Goodman, L. B., Trumbetta, S. L., Rosenberg, S. D., Osher, F. C., Vidaver, R., Auciello, P., & Foy, D. W. (1998). Trauma and posttraumatic stress disorder in severe mental illness. Journal of Consulting and Clinical Psychology, 66, 493-499.
- National Association of State Mental Health Program Directors. (2005). Position Statement on Services and Supports to Trauma Survivors. Alexandria, VA: NASMHPD.
- National Institute on Drug Abuse. (2011). Topics in Brief: Comorbid drug abuse and mental illness. Retrieved December 28, 2011 at: http://www.drugabuse.gov/DrugPages/comorbidity.html.
- Rangan, A & Sekar, K. (2006). Strengths perspective in mental health (evidence based case study. Brisbane Institute of Strengths-Based Practice. Retrieved December 29, 2011 at: http://www.strengthbasedstrategies.com/papers.htm.
- Rapp, C.A. (1998). The strengths model: Case management with people suffering from severe and persistent mental illness. London: Oxford University Press.
- Ruiz, P. (2004). Addressing Culture, Race, & Ethnicity in Psychiatric Practice. Psychiatric Annals, 34:7, 527-532.
- Saleebey, D. (1996). The strengths perspective in social work practice: extension and cautions. Social Work. 41:3, 296-305.
- Substance Abuse and Mental Health Services Administration. (2005). Results from the 2004 national survey on drug use and health: national findings. NSDUH Series H-28. DHHS Publication No. (SMA) 05-4062. Rockville, MD: Office of Applied Studies.
- Swanson, J.W., (1994). Mental disorder, substance abuse, and community violence: an epidemiological approach. In: Monahan, J. & Steadman, H.J. (Eds.) Violence and mental disorder: developments in risk assessment. Chicago: University of Chicago Press, 101-136.
- Swanson, J.W., Borum, R., Swartz, M. et al. (1996). Psychotic symptoms and disorders and the risk of violent behavior in the community. Criminal Behav Ment Hlth. 6:317-338.
- Swartz, M.S., Swanson, J.W., Hiday, V.A. et al. (1998). Violence and severe mental illness: the effects of substance abuse and nonadherence to medication. Am J Psychiatry. 155:226-231.
- The Joint Commission. (2010). Sentinel Event Alert: Preventing violence in the health care setting. Retrieved December 28, 2011 at: http://www.jointcommission.org/sentinel_event_alert_issue_45_preventing_violence_in_the_health_care_setting_/
- The Joint Commission. (2011). Sentinel Event Data - Root Causes by Event Type. Retrieved December 28, 2011 at: http://www.jointcommission.org/Sentinel_Event_Statistics/
- The National Quality Forum, (2007). National Voluntary Consensus Standards for the Treatment of Substance Use Conditions: Evidence-Based Treatment Practices; A Consensus Report.
- Tucker, W.M. (2002). How to include the trauma history in the diagnosis and treatment of psychiatric inpatients. Psychiatric Quarterly. 73:2,135-145.
- US Department of Health and Human Services. (2001). Mental health: culture, race and ethnicity-a supplement to mental health: report of the surgeon general. Rockville MD: US Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- Witness Justice. (2006).Trauma-the “common denominator”. Retrieved on December 29, 2011 at: http://www.witnessjustice.org/health/trauma.cfm.
- Ziedonis, D.M. (2004). Integrated treatment of co-occurring mental illness and addiction: Clinical intervention, program, and system perspectives. CNS Spectrums 9(12): 892,894-904,925.

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.

Evidence exists that there is a high prevalence of co-occurring substance use disorders as well as history of trauma among persons admitted to acute psychiatric settings. Professional literature suggests that these factors are under-identified yet integral to current psychiatric status and should be assessed in order to develop appropriate treatment (Ziedonis, 2004; NASMHPD, 2005). Similarly, persons admitted to inpatient settings require a careful assessment of risk for violence and the use of seclusion and restraint. Careful assessment of risk is critical to safety and treatment. Effective, individualized treatment relies on assessments that explicitly recognize patients´ strengths. These strengths may be characteristics of the individuals themselves, supports provided by families and others, or contributions made by the individuals´ community or cultural environment (Rapp, 1998). In the same way, inpatient environments require assessment for factors that lead to conflict or less than optimal outcomes.

As stated above, recent literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.Evidence exists that there is a high prevalence of co-occurring substance use disorders as well as history of trauma among persons admitted to acute psychiatric settings. Professional literature suggests that these factors are under-identified yet integral to current psychiatric status and should be assessed in order to develop appropriate treatment (Ziedonis, 2004; NASMHPD, 2005). Similarly, persons admitted to inpatient settings require a careful assessment of risk for violence and the use of seclusion and restraint. Careful assessment of risk is critical to safety and treatment. Effective, individualized treatment relies on assessments that explicitly recognize patients´ strengths. These strengths may be characteristics of the individuals themselves, supports provided by families and others, or contributions made by the individuals´ community or cultural environment (Rapp, 1998). In the same way, inpatient environments require assessment for factors that lead to conflict or less than optimal outcomes.

The literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery

model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.As stated above, recent literature supports the routine initial screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths to assist the clinician in determining which patients require a more in depth assessment based on findings which will ultimately form the basis for an appropriate treatment plan. The reduction in the under-detection of violence risk, SUD and trauma history will in turn decrease the chance of psychiatric relapse and lead to improved medication compliance which will ultimately reduce the ongoing costs of psychiatric treatment. And finally, by focusing on patient strengths instead of problems during the screening process, the patient will become empowered to embrace the ongoing recovery model of treatment thereby reducing the need for readmission to more restrictive levels of treatment such as inpatient care.

The measure will assist health care organizations (HCOs) to track admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.

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	300	323	475	484	522	673	1054	1082	742	726
Mean	0.87073	0.918	0.91833	0.95211	0.94764	0.88469	0.90865	0.91536	0.93886	0.93725
Std. Dev.	0.19	0.1429	0.1449	0.0878	0.1242	0.2209	0.1665	0.1629	0.1214	0.1346
Max	1	1	1	1	1	1	1	1	1	1
Q3	0.99096	0.99868	0.99736	0.99928	0.9991	0.99769	0.99824	0.99846	0.99905	1
Median	0.94479	0.97829	0.9769	0.98447	0.98894	0.9809	0.97998	0.98408	0.98262	0.98406
Q1	0.82917	0.90942	0.91373	0.9466	0.95755	0.9078	0.90643	0.91808	0.93452	0.9446
Min	0.0082	0.18258	0	0.26987	0.00836	0.00357	0	0	0.05068	0
10th Pctl	0.64043	0.75051	0.76257	0.87302	0.87561	0.57728	0.71264	0.75	0.85333	0.83071
20th Pctl	0.80694	0.87957	0.89394	0.92754	0.94467	0.86036	0.87195	0.88166	0.91406	0.92381
30th Pctl	0.86897	0.92727	0.92912	0.95816	0.96724	0.93333	0.93155	0.93931	0.94788	0.95599
40th Pctl	0.91918	0.95975	0.95978	0.97521	0.98028	0.96553	0.96382	0.9673	0.96999	0.97187
60th Pctl	0.9722	0.98884	0.98745	0.99257	0.99461	0.98942	0.99184	0.99238	0.99145	0.99237
70th Pctl	0.98656	0.99644	0.99528	0.99679	0.9976	0.99563	0.99664	0.99694	0.99733	0.9985
80th Pctl	0.99346	1	1	1	1	0.99932	1	1	1	1
90th Pctl	1	1	1	1	1	1	1	1	1	1
# Patients	206992	224505	311947	333425	337484	386345	518224	538940	425128	217531

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.*

There is a great deal of literature supporting admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths. There is no mention of disparities related to race or socioeconomic status regarding admission screening for risk of violence to self or others, substance use, psychological trauma history and patient strengths.

The literature supports male veterans, adolescents, the elderly, Native Americans, those with mood disorders, co morbid substance abuse disorders and a history of physical and sexual abuse as the most vulnerable group of patients at higher risk for suicide (U.S. Preventive Services Task Force, 2004). Swanson, et al. (1990) and Mullen (2000) noted that those with mental illness who were violent were more likely to have a lower socioeconomic status.

Young male veterans are also more likely to experience SUD especially when combined with trauma history (Substance Abuse and Mental Health Services Administration, 2007). Patients with co-occurring SMI and SUD were more likely to be unemployed, white and female according to the Substance Abuse and Mental Health Services Administration (2003).

For data source see data in 1b.2

Rates by Population Group

Gender	2013	2014	2015	2016	2017
Male	0.961	0.960	0.927	0.931	0.950
Female	0.973	0.917	0.941	0.950	0.955

Hispanic Ethnicity	2013	2014	2015	2016	2017
Hispanic	0.952	0.933	0.919	0.938	0.949
Non-Hispanic	0.967	0.937	0.935	0.940	0.952

Race	2013	2014	2015	2016	2017
White	0.969	0.945	0.934	0.941	0.952
African American	0.964	0.935	0.928	0.933	0.950
American Indian	0.941	0.937	0.893	0.904	0.927
Asian	0.961	0.950	0.933	0.934	0.949
Pacific Islander	0.951	0.943	0.924	0.909	0.929

Age Category	2013	2014	2015	2016	2017
1-12 years	0.981	0.981	0.961	0.954	0.968
13-17 years	0.984	0.980	0.963	0.962	0.965
18-64 years	0.961	0.932	0.930	0.937	0.950
65 years and above	0.953	0.876	0.910	0.924	0.941

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):*

Behavioral Health : Alcohol, Substance Use/Abuse, Behavioral Health : Suicide

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

Health and Functional Status : Change, Person-and Family-Centered Care, Safety

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-636794265723952869.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.

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 with admission screening within the first three days of admission for all of the following: risk of violence to self or others; substance use; psychological trauma history; and patient strengths

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).

Five data elements are used to calculate the numerator:

1. Patient Strengths - Documentation in the medical record that an admission screening for a minimum of two patient strengths was performed within the first three days of admission. Allowable values: Yes, No/UTD, or X unable to complete admission screening.

2. Psychological Trauma History - Documentation in the medical record that an admission screening for a psychological trauma history was performed within the first three days of admission. Allowable values: Yes, No/UTD, or X unable to complete admission screening.

3. Substance Use - Documentation in the medical record that an admission screening for substance use and alcohol use which occurred over the past twelve (12) months was performed within the first three days of admission. The screening must include: the type, amount, frequency of use and any problems due to past use. Allowable values: Yes, No/UTD, or X unable to complete admission screening.

4. Violence Risk to Others - Documentation in the medical record that an admission screening for violence risk to others over the past six months was performed within the first three days of admission. Violence Risk to Others includes: threats of violence and/or actual commission of violence toward others. Documentation should include violence risk within the 6 months prior to admission AND any lifetime risk of violence to others beyond the 6 months prior to admission. Allowable values: Yes, No/UTD, or X unable to complete admission screening.

5. Violence Risk to Self - Documentation in the medical record that an admission screening for violence risk to self over the past six months was performed within the first three days of admission. Violence Risk to Self includes: ideation, plans/preparation and/or intent to act if ideation present, past suicidal behavior and risk/protective factors within the 6 months prior to admission. Allowable values: Yes, No/UTD, or X unable to complete admission screening.

Patients are eligible for the numerator population when the allowable value equals “yes” for all five data elements: Patient Strengths, Psychological Trauma History, Substance Use, Violence Risk to Others and Violence Risk to Self as defined above.

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.)*

<u>IF an OUTCOME MEASURE</u>, 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.)

(See S.2b for attached code table)

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

5. 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.

6. 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.

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

• Patients for whom there is an inability to complete admission screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths within the first three days of admission due to the patient's inability or unwillingness to answer screening questions

• Patients with a Length of Stay = or less than 3 days or = or greater than 365 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 for whom screening cannot be completed due to the patient's inability or unwillingness to answer assessment questions within the first three days of admission OR patients with a previous admission to the psychiatric unit during a single hospitalization.

• Length of stay (LOS) in days is equal to the Discharge Date minus the Admission Date. If the LOS is less than 3 days or greater than 365 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 all cases that are included in the Initial Patient Population for HBIPS Discharge 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 to 3 days or greater than or equal to 365 days, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop processing.

b. If Length of Stay is greater than 3 days and less than 365 days, continue processing and proceed to Psychiatric Care Setting.

4. Check Psychiatric Care Setting

a. If Psychiatric Care Setting equals No, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop processing.

b. If Psychiatric Care Setting equals Yes, continue processing.

5. Initialize Missing Counter to equal zero. Initialize No Screening Counter to equal zero, Initialize Incomplete Screening Counter to equal zero. Continue processing and proceed to Patient Strengths.

6. Check Patient Strengths

a. If Patient Strengths equals No, add one to No Screening Counter. Continue processing and proceed to Psychological Trauma History.

b. If Patient Strengths is missing, add one to Missing Counter. Continue processing and proceed to Psychological Trauma History.

c. If Patient Strengths equals Yes or X, Continue processing and proceed to check Patient Strengths.

7. Check Patient Strengths

a. If Patient Strengths equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Psychological Trauma History.

b. If Patient Strengths equals Yes, Continue processing and proceed to Psychological Trauma History.

8. Check Psychological Trauma History

a. If Psychological Trauma History equals No, add one to No Screening Counter. Continue processing and proceed to Substance Use.

b. If Psychological Trauma History is missing, add one to Missing Counter. Continue processing and proceed to Substance Use.

c. If Psychological Trauma History equals Yes or X, Continue processing and proceed to check Psychological Trauma History.

9. Check Psychological Trauma History

a. If Psychological Trauma History equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Substance Use.

b. If Psychological Trauma History equal Yes, Continue processing and proceed to Substance Use.

10. Check Substance Use

a. If Substance Use equals No, add one to No Screening Counter. Continue processing and proceed to Violence Risk to Others.

b. If Substance Use is missing, add one to Missing Counter. Continue processing and proceed to Violence Risk to Others.

c. If Substance Use equals Yes or X, Continue processing and proceed to check Substance Use.

11. Check Substance Use

a. If Substance Use equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Violence Risk to Others.

b. If Substance Use equal Yes, Continue processing and proceed to Violence Risk to Others.

12. Check Violence Risk to Others

a. If Violence Risk to Others equals No, add one to No Screening Counter. Continue processing and proceed to Violence Risk to Self.

b. If Violence Risk to Others is missing, add one to Missing Counter. Continue processing and proceed to Violence Risk to Self.

c. If Violence Risk to Others equals Yes or X, Continue processing and proceed to check Violence Risk to Others.

13. Check Violence Risk to Others

a. If Violence Risk to Others equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Violence Risk to Self.

b. If Violence Risk to Others equal Yes, Continue processing and proceed to Violence Risk to Self.

14. Check Violence Risk to Self

a. If Violence Risk to Self equals No, add one to No Screening Counter. Continue processing and proceed to Incomplete Screening Counter.

b. If Violence Risk to Self is missing, add one to Missing Counter. Continue processing and proceed to Incomplete Screening Counter.

c. If Violence Risk to Self equals Yes or X, Continue processing and proceed to check Violence Self.

15. Check Violence Risk to Self

a. If Violence Risk to Self equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Incomplete Screening Counter.

b. If Violence Risk to Self equal Yes, Continue processing and proceed to Incomplete Screening Counter.

16. Check Incomplete Screening Counter

a. If Incomplete Screening Counter equals 5, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Continue processing and proceed to initialize the Measure Category Assignment for each strata measure.

b. If Incomplete Screening Counter is less than five, continue processing and proceed to Missing Counter.

17. Check Missing Counter

a. If Missing Counter is more than zero, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-1a) and will be rejected. Proceed to step initialize the Measure Category Assignment for each strata measure.

b. If Missing Counter equals zero, continue processing and proceed to No Screening Counter.

18. Check No Screening Counter

a. If No Screening Counter is greater than zero, the case will proceed to a Measure Category Assignment of D for Overall Rate (HBIPS-1a) and will be in the measure population. Continue processing and proceed to step 19 and initialize the Measure Category Assignment for each strata measure.

b. If No Screening Counter equals zero, the case will proceed to a Measure Category Assignment of E and will be in the measure population. Continue processing and proceed to step 19 and initialize the Measure Category Assignment for each strata measure.

19. Initialize the Measure Category Assignment for each strata measure (b-e) equal ´B´. Do not change the Measure Category Assignment that was already calculated for the overall rate (HBIPS-1a). The rest of the algorithm will reset the appropriate Measure Category Assignment to be equal to the overall rate´s (HBIPS-1a) Measure Category Assignment. Continue processing and proceed to Overall Rate Category Assignment.

20. Check Overall Rate Category Assignment

a. If Overall Rate Category Assignment equals B, retain the Measure Category Assignment for the strata measures (HBIPS-1b through HBIPS-1e) equals B. Stop processing.

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

21. Check Patient Age at Discharge

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

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

22. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal to 13 years and less than 18 years, set the Measure Category Assignment for the measure HBIP-1c equal to Measure Category Assignment for measure HBIP-1a. Stop processing.

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

23. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal to 18 years and less than 65 years, set the Measure Category Assignment for the measure HBIP-1d equal to Measure Category Assignment for measure HBIP-1a. Stop processing.

b. If Patient Age at Discharge is greater than or equal to 65 years, set the Measure Category Assignment for the measure HBIP-1e equal to Measure Category Assignment for measure HBIP-1a. Stop processing.1. Run all cases that are included in the Initial Patient Population for HBIPS Discharge 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 to 3 days or greater than or equal to 365 days, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop processing.

b. If Length of Stay is greater than 3 days and less than 365 days, continue processing and proceed to Psychiatric Care Setting.

4. Check Psychiatric Care Setting

a. If Psychiatric Care Setting equals No, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Stop processing.

b. If Psychiatric Care Setting equals Yes, continue processing.

5. Initialize Missing Counter to equal zero. Initialize No Screening Counter to equal zero, Initialize Incomplete Screening Counter to equal zero. Continue processing and proceed to Patient Strengths.

6. Check Patient Strengths

a. If Patient Strengths equals No, add one to No Screening Counter. Continue processing and proceed to Psychological Trauma History.

b. If Patient Strengths is missing, add one to Missing Counter. Continue processing and proceed to Psychological Trauma History.

c. If Patient Strengths equals Yes or X, Continue processing and proceed to check Patient Strengths.

7. Check Patient Strengths

a. If Patient Strengths equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Psychological Trauma History.

b. If Patient Strengths equals Yes, Continue processing and proceed to Psychological Trauma History.

8. Check Psychological Trauma History

a. If Psychological Trauma History equals No, add one to No Screening Counter. Continue processing and proceed to Substance Use.

b. If Psychological Trauma History is missing, add one to Missing Counter. Continue processing and proceed to Substance Use.

c. If Psychological Trauma History equals Yes or X, Continue processing and proceed to check Psychological Trauma History.

9. Check Psychological Trauma History

a. If Psychological Trauma History equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Substance Use.

b. If Psychological Trauma History equal Yes, Continue processing and proceed to Substance Use.

10. Check Substance Use

a. If Substance Use equals No, add one to No Screening Counter. Continue processing and proceed to Violence Risk to Others.

b. If Substance Use is missing, add one to Missing Counter. Continue processing and proceed to Violence Risk to Others.

c. If Substance Use equals Yes or X, Continue processing and proceed to check Substance Use.

11. Check Substance Use

a. If Substance Use equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Violence Risk to Others.

b. If Substance Use equal Yes, Continue processing and proceed to Violence Risk to Others.

12. Check Violence Risk to Others

a. If Violence Risk to Others equals No, add one to No Screening Counter. Continue processing and proceed to Violence Risk to Self.

b. If Violence Risk to Others is missing, add one to Missing Counter. Continue processing and proceed to Violence Risk to Self.

c. If Violence Risk to Others equals Yes or X, Continue processing and proceed to check Violence Risk to Others.

13. Check Violence Risk to Others

a. If Violence Risk to Others equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Violence Risk to Self.

b. If Violence Risk to Others equal Yes, Continue processing and proceed to Violence Risk to Self.

14. Check Violence Risk to Self

a. If Violence Risk to Self equals No, add one to No Screening Counter. Continue processing and proceed to Incomplete Screening Counter.

b. If Violence Risk to Self is missing, add one to Missing Counter. Continue processing and proceed to Incomplete Screening Counter.

c. If Violence Risk to Self equals Yes or X, Continue processing and proceed to check Violence Self.

15. Check Violence Risk to Self

a. If Violence Risk to Self equals X, add one to Incomplete Screening Counter. Continue processing and proceed to Incomplete Screening Counter.

b. If Violence Risk to Self equal Yes, Continue processing and proceed to Incomplete Screening Counter.

16. Check Incomplete Screening Counter

a. If Incomplete Screening Counter equals 5, the case will proceed to a Measure Category Assignment of B and will not be in the measure population. Continue processing and proceed to initialize the Measure Category Assignment for each strata measure.

b. If Incomplete Screening Counter is less than five, continue processing and proceed to Missing Counter.

17. Check Missing Counter

a. If Missing Counter is more than zero, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-1a) and will be rejected. Proceed to step initialize the Measure Category Assignment for each strata measure.

b. If Missing Counter equals zero, continue processing and proceed to No Screening Counter.

18. Check No Screening Counter

a. If No Screening Counter is greater than zero, the case will proceed to a Measure Category Assignment of D for Overall Rate (HBIPS-1a) and will be in the measure population. Continue processing and proceed to step 19 and initialize the Measure Category Assignment for each strata measure.

b. If No Screening Counter equals zero, the case will proceed to a Measure Category Assignment of E and will be in the measure population. Continue processing and proceed to step 19 and initialize the Measure Category Assignment for each strata measure.

19. Initialize the Measure Category Assignment for each strata measure (b-e) equal ´B´. Do not change the Measure Category Assignment that was already calculated for the overall rate (HBIPS-1a). The rest of the algorithm will reset the appropriate Measure Category Assignment to be equal to the overall rate´s (HBIPS-1a) Measure Category Assignment. Continue processing and proceed to Overall Rate Category Assignment.

20. Check Overall Rate Category Assignment

a. If Overall Rate Category Assignment equals B, retain the Measure Category Assignment for the strata measures (HBIPS-1b through HBIPS-1e) equals B. Stop processing.

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

21. Check Patient Age at Discharge

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

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

22. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal to 13 years and less than 18 years, set the Measure Category Assignment for the measure HBIP-1c equal to Measure Category Assignment for measure HBIP-1a. Stop processing.

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

23. Check Patient Age at Discharge

a. If Patient Age at Discharge is greater than or equal to 18 years and less than 65 years, set the Measure Category Assignment for the measure HBIP-1d equal to Measure Category Assignment for measure HBIP-1a. Stop processing.

b. If Patient Age at Discharge is greater than or equal to 65 years, set the Measure Category Assignment for the measure HBIP-1e equal to Measure Category Assignment for measure HBIP-1a. 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

221-977 20% of the 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 strataum 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 Minimum Required

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

Not applicable

2. Validity – See attached Measure Testing Submission Form

1922_MeasureTesting_7.1_HBIPS1-636898056580786636.docx

2.1 For maintenance of endorsement

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

No

2.2 For maintenance of endorsement

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

Yes

2.3 For maintenance of endorsement

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

No - This measure is not risk-adjusted

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

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

Measure Number (*if previously endorsed*)**:** 1922
Measure Title: Admission Screening
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:	Measure Tested with Data From:

(*must be consistent with data sources entered in S.17*)	
☒ 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 <u>patients</u> 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

***<u>Note</u>**: 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 tests used were the calculated agreement rates for individual data elements that are used to compute measure rates for HBIPS-1.

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 with a Mismatch*	Total Numerator	Total Denominator	Agreement Rate
Numerator Data Elements			
Patient Strengths	187	191	97.9%
Psychological Trauma History	189	191	98.9%
Substance Use	188	191	98.4%
Violence Risk to Others	187	191	97.9%
Violence Risk to Self	188	191	98.4%
Denominator Data Elements			
Admission Date	191	191	100%
Birthdate	191	191	100%
Discharge Date	191	191	100%
ICD-9-CM Other Diagnosis Code**	191	191	100%
ICD-9-CM Principal Diagnosis Code**	191	191	100%
Psychiatric Care Setting	191	191	100%

* No cases were excluded for the reliability testing.

** 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.

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?*)

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)*

At the time this measure was originally tested measure validity was assessed via survey and focus groups of hospitals participating in the pilot test. All measure specifications, including population identification, numerator and denominator statements, and data elements and their definitions were found to be understandable, retrievable, and relevant.

Since the measure has been in national use, continued face validity of the measure has been determined through analysis of feedback from measure users. The Joint Commission provides a web-based application with which measure users can provide feedback regarding appropriateness of measure specifications, request clarification of specifications, and/or provide other comments pertinent to the measure. This feedback is systematically, continually, reviewed in order to identify trends and to identify areas of the measure specifications that require clarification or revision. Additionally, Joint Commission staff continually monitors the national literature and environment in order to assess continued validity of this measure. And finally, the crosswalk from ICD-9-CM diagnosis codes to ICD-10-CM diagnosis codes has been completed and reviewed by the Technical Advisory Panel for face validity. The panel has determined that the intent of the measure has not changed as a result of the conversion.

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

Tests for correlations between HBIPS-1 and the remaining HBIPS measures (HBIPS-2, HBIPS-3, HBIPS-5) are -0.00313(p=0.9328), -0.00875(p=0.8144), and 0.13857(p=0.0002), respectively. This indicates that there are no statistically significant correlations between HBIPS-1 and HBIPS-2 and HBIPS-3. There is a slight positive correlation between HBIPS-1 and HBIPS-5. 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 issues with the data elements for this measure and no updates needed to the data element specifications.

Analysis of feedback obtained via our automated feedback system reveals around 100 submissions regarding specifications for this measure over the past three years. Predominant themes of these submissions involved questions regarding clarification of the data elements Violence Risk to Others, Violence Risk to Self and Substance Use with respect to both definitions and the length of time to inquire about past history. For all of the numerator data elements, the allowable value for unable to complete admission screening was revised to include patients with a previous admission to the psychiatric unit during a single hospitalization, since the admission screening only needs to be completed once during the initial admission to the psychiatric unit. Additional notes for abstractors were also added to the data elements for clarification. Finally, for patients with a length of stay less than 3 days a denominator population exclusion was added, since the measure allows up to 3 days for completion of the admission screening.

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-1 and HBIPS-5 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

☐ **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 were 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.

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= 519,664

Exclusions:

- Patients who have a length of stay (LOS) less than 3 days or greater than 365 days =16.8%
- Patients for whom there is an inability to complete admission screening for Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths within the first three days of admission =0.3%

Rationale for exclusions:

- Patients with a Length of Stay ≤ 3 days or ≥ 365 days

  Rationale:
  - Screening must take place within 3 days of admission. Patient stay of less than 3 days will inappropriately fail these cases.
  - In the initial testing it was discovered that patients discharged after a long-term hospitalization (several years or longer) were lacking some items in the initial assessment.
- Patients for whom there is an inability to complete admission screening for *Violence Risk, Substance Use, Psychological Trauma History and Patient Strengths* within the first three days of admission

  Rationale: It is recognized that for some patients their condition may be such that they are unable to participate in the admission screening process within the allotted time frame. As this is out of provider control, these cases are excluded from the 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 large 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.

Not Applicable

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

Not Applicable

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*)

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

Not Applicable

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-1 Distribution of the Measure Results

2018 2nd Quarter Data:
Scores on this measure: N=722, Mean 93.7%, SD 0.1405
10th Percentile= 83.1%
25th Percentile= 94.5%
50th Percentile= 98.7%
75th Percentile= 100%
90th Percentile= 100%

365 Favorable (50.7%) – results statistically significantly higher than the national rate
260 (36.1%) Neutral – results not significantly different from target range
97 (13.4%) 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 the rates for patient screening.

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/ 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: 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 1.183.

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.

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 admission screening.

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.

Yes

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

0104 : Adult Major Depressive Disorder (MDD): Suicide Risk Assessment

0110 : Bipolar Disorder and Major Depression: Appraisal for alcohol or chemical substance use

0111 : Bipolar Disorder: Appraisal for risk of suicide

1365 : Child and Adolescent Major Depressive Disorder (MDD): Suicide Risk Assessment

2152 : Preventive Care and Screening: Unhealthy Alcohol Use: Screening & Brief Counseling

2599 : Alcohol Screening and Follow-up for People with Serious Mental Illness

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

SUB-1 Alcohol Use Screening STEWARD: The Joint Commission

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?

No

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

Five of the six NQF endorsed measures are provider level measures, 2599 is a health plan measure. All pertain to the ambulatory setting for patients. All (except 2152) are specific to the diagnoses of major depression and/or bipolar disorder. The measures only evaluate one aspect of screening: either suicide risk or alcohol or substance use. Measures 0104, 0110, 0111, 2159, and 2599 only evaluate patients age 18 years and older. The SUB-1 measure pertains to all inpatients 18 years and older, with screening limited to substance use. HBIPS-1 addresses inpatient organizational performance for all psychiatric diagnoses and evaluates the care of all patient ages (greater than 1 year). Additionally, HBIPS-1 evaluates several aspects of screening (risk of violence to self or others, substance use, psychological trauma history and patient strengths).

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.

No appendix **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

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.