



Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

Brief Measure Information

NQF #: 0640

Corresponding Measures:

Measure Title: HBIPS-2 Hours of physical restraint use

Measure Steward: The Joint Commission

sp.02. Brief Description of Measure: The total number of hours that all patients admitted to a hospital-based inpatient psychiatric setting were maintained in physical restraint.

1b.01. Developer Rationale: Mental health providers that value and respect an individual's autonomy, independence and safety seek to avoid the use of dangerous or restrictive interventions at all times (Donat, 2003). The use of seclusion and restraint is limited to situations deemed to meet the threshold of imminent danger and when restraint and seclusion are used; such use is rigorously monitored and analyzed to prevent future use. Providers also seek to prevent violence or aggression from occurring in their treatment environments by focusing their attention on prevention activities that have a growing evidence base (Donat, 2003).

The literature supports a reduction in the use of physical restraint and seclusion. A reduction in the use of physical restraint and seclusion will improve patient safety, reduce overall organizational costs, leading to a decrease in staff and patient injuries and an increase in staff productivity. And finally, focusing on behavioral interventions to defuse aggressive and violent behaviors will result in less retraumatization for patients with trauma histories leading to shortened and improved recovery.

The measure will assist health care organizations (HCOs) to track physical restraint use to determine patterns and trends to aid the organization in efforts to decrease use.

sp.12. Numerator Statement: The total number of hours that all psychiatric inpatients were maintained in physical restraint.

Numerator Basis: The numerator evaluates the number of hours of physical restraint; however, the algorithm calculates the number of minutes to ensure a more accurate calculation of the measure. Convert the minutes to hours when analyzing and reporting this measure.

sp.14. Denominator Statement: Number of psychiatric inpatient days

Denominator basis: per 1,000 hours

To compute this measure rate, a base of 1000 hours has been applied to total patient days in the denominator (i.e., total patient days are divided by 1000). The purpose of this is to create a smaller denominator number, thus providing a more understandable rate. When multiplied by 1000, this rate measures numerator occurrence per total patient days.

sp.16. Denominator Exclusions: Total leave days

Measure Type: Process

sp.28. Data Source:

Electronic Health Records

Paper Medical Records

sp.07. Level of Analysis:

Facility

Other

IF Endorsement Maintenance – Original Endorsement Date: 2010-05-05 12:00 AM

Most Recent Endorsement Date: 10/24/2019 1:38:25 PM

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

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

sp.03. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?:

1. 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

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

[Response Begins]

No

[Response Ends]

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

Current Submission:

Updated evidence information here.

Previous (Year) Submission:

Evidence from the previous submission here.

1a.01. Provide a logic model.

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

If the evidence is not based on a systematic review, skip to the end of the section and do not complete the repeatable question group below. If you wish to include more than one systematic review, add additional tables by clicking "Add" after the final question in the group.

Evidence - Systematic Reviews Table (Repeatable)

Group 1 - Evidence - Systematic Reviews Table

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

1a.15. Detail the process used to identify the evidence.

[Response Begins]

[Response Ends]

1a.16. Provide the citation(s) for the evidence.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

Mental health providers that value and respect an individual's autonomy, independence and safety seek to avoid the use of dangerous or restrictive interventions at all times (Donat, 2003). The use of seclusion and restraint is limited to situations deemed to meet the threshold of imminent danger and when restraint and seclusion are used; such use is rigorously monitored and analyzed to prevent future use. Providers also seek to prevent violence or aggression from occurring in their treatment environments by focusing their attention on prevention activities that have a growing evidence base (Donat, 2003).

The literature supports a reduction in the use of physical restraint and seclusion. A reduction in the use of physical restraint and seclusion will improve patient safety, reduce overall organizational costs, leading to a decrease in staff and patient injuries and an increase in staff productivity. And finally, focusing on behavioral interventions to defuse aggressive and violent behaviors will result in less retraumatization for patients with trauma histories leading to shortened and improved recovery.

The measure will assist health care organizations (HCOs) to track physical restraint use to determine patterns and trends to aid the organization in efforts to decrease use.

[Response Ends]

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

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

[Response Begins]

Below are the data from 2009-2018 adjusted per 1000 patient days. 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. . Also included are the number of cases/occurrences that contributed to the numerator.

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
N	299	318	470	480	517	667	1041	1065	740	725
Mean	0.3036	0.26683	0.25543	0.29883	0.32079	0.33013	1.02502	0.71258	0.37637	0.4013
Std. Dev.	0.7972	0.5377	0.51	0.7643	0.8847	1.1124	10.3758	6.5778	1.3828	1.5454
Max	10.268	5.293	4.455	9.907	9.291	17.09	222.222	145.105	22.242	22.17
Q3	0.28824	0.25384	0.24109	0.23274	0.22987	0.23112	0.2538	0.25329	0.22875	0.21895
Median	0.11371	0.09899	0.08627	0.09293	0.09767	0.08063	0.08016	0.07672	0.07312	0.0701
Q1	0.02875	0.0339	0.02667	0.02702	0.0233	0.01807	0.0139	0.01414	0.01722	0.01609
Min	0	0	0	0	0	0	0	0	0	0
10th Pctl	0.00384	0.0046	0.0021	0.0032	0.00282	0.00135	0	0	0.00016	0
20th Pctl	0.01858	0.0221	0.01935	0.01929	0.01564	0.01373	0.00722	0.0055	0.01156	0.00907
30th Pctl	0.04201	0.04679	0.0379	0.03522	0.03118	0.02778	0.0216	0.02212	0.02344	0.02288
40th Pctl	0.07542	0.06471	0.06149	0.05688	0.0598	0.05738	0.04783	0.04569	0.04226	0.04019
60th Pctl	0.1474	0.14529	0.1366	0.13391	0.14084	0.12482	0.12615	0.11384	0.11793	0.11462
70th Pctl	0.22613	0.19998	0.19373	0.19707	0.19275	0.19435	0.19679	0.18962	0.18237	0.17676
80th Pctl	0.35552	0.3313	0.30858	0.2952	0.27992	0.32529	0.3458	0.33422	0.31546	0.30712
90th Pctl	0.56838	0.58494	0.61379	0.62437	0.62406	0.60094	0.73654	0.76161	0.67191	0.7138
# Cases	26455	30611	51842	55095	57607	58391	69345	73880	68635	36192

[Response Ends]

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

[Response Begins]

See data in 1b.2

[Response Ends]

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

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

According to Mead et al (2008), minority youths are more likely to be restrained than white youths. African American youths are two times more likely and Hispanic youths are 70% more likely to be restrained upon admission to a psychiatric hospital. One study conducted from 2006 through 2010 in mental health facilities in Canada showed that the oldest patients were 46% less likely to be restrained or secluded than younger patients and patients who are unemployed are 22% more likely to be restrained or secluded (Canadian Institute for Health Information [CIHI], 2011). In a case-control study conducted by Knutzen et al (2011), restraint use was also associated with multiple admissions, involuntary admissions and patients with serious mental illness (SMI).

For gender, Hispanic ethnicity and race, data are provided by median restraint time for those with a restraint since the denominator data was not broken down by these demographic categories. For age group the measure rates are provided.

For data source see data in 1b.2

Median Restraint Time (in minutes) for patients that were restrained

Gender 2013 2014 2015 2016 2017

Male 14 16 20 16 14

Female 15 15 15 15 14

Hispanic Ethnicity 2013 2014 2015 2016 2017

Hispanic 36 50 47 45 30

Non-Hispanic 13 15 16 15 13

Race 2013 2014 2015 2016 2017

White 15 17 30 17 15

African American 11 13 16 16 13

American Indian 50 30 11 8 7

Asian 33 14 35 51 21

Pacific Islander 26 107 135 75 75

Median Measure Rate (per 1,000 patient hours) by age group

Age Category 2013 2014 2015 2016 2017

1-12 years 0.17 0.17 0.16 0.15 0.13

13-17 years 0.10 0.10 0.09 0.08 0.08

18-64 years 0.07 0.07 0.07 0.06 0.05

65+ years 0.00 0.00 0.00 0.00 0.00

[Response Ends]

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

[Response Begins]

Not applicable

[Response Ends]

2. 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.

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

[Response Begins]

Yes

[Response Ends]

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

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

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

[Response Begins]

Data Element Minutes of Physical Restraint:

- Updates were made to the table Guidelines for Abstraction to provide clarification of what constitutes physical restraint.
- Notes for Abstraction were updated to clarify the priority for tracking time in restraint/seclusion when a patient is placed in restraint and seclusion at the same time.

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.

[Response Ends]

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see [What Good Looks Like](#)).

[Response Begins]

HBIPS-2 Hours of physical restraint use

[Response Ends]

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

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]

The total number of hours that all patients admitted to a hospital-based inpatient psychiatric setting were maintained in physical restraint.

[Response Ends]

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

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Surgery: General*

[Response Begins]

[Response Ends]

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

[Response Begins]

Person-and Family-Centered Care: Person-and Family-Centered Care

Safety

Safety: Overuse

[Response Ends]

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

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

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Populations at Risk: Populations at Risk*

[Response Begins]

Elderly (Age >= 65)

[Response Ends]

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

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

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

Facility

Other

[Response Ends]

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

Check ONLY the settings for which the measure is SPECIFIED and TESTED.

[Response Begins]

Inpatient/Hospital

[Response Ends]

sp.09. 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. If no URL is available, indicate "none available".

[Response Begins]

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

[Response Ends]

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

Attach an excel or csv file; if this poses an issue, [contact staff](#). Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

[Response Begins]

No data dictionary/code table – all information provided in the submission form

[Response Ends]

sp.13. State the numerator.

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

DO NOT include the rationale for the measure.

[Response Begins]

The total number of hours that all psychiatric inpatients were maintained in physical restraint.

Numerator Basis: The numerator evaluates the number of hours of physical restraint; however, the algorithm calculates the number of minutes to ensure a more accurate calculation of the measure. Convert the minutes to hours when analyzing and reporting this measure.

[Response Ends]

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

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

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

[Response Begins]

Three data elements are used to calculate the numerator:

1. Event Date* - The month, day and year of the event.
2. Event Type* - The measure-related event being identified. Allowable values: 1. Physical Restraint 2. Seclusion
3. Minutes of Physical Restraint - The total minutes recorded in the medical record that a patient was maintained in Event Type 1 (physical restraint(s)) for the associated Event Date. Allowable values 1-1440 minutes

*The data elements Event Date and Event Type are used for both HBIPS-2 (Hours of Physical Restraint Use) and HBIPS-3: Hours of Seclusion Use).

Patients are eligible for the numerator population when a physical restraint event occurs.

A physical restraint is any manual method or physical or mechanical device, material, or equipment that immobilizes or reduces the ability of a patient to move his or her arms, legs, body or head freely when it is used as a restriction to manage a patient's behavior or restrict the patient's freedom of movement and is not a standard treatment for the patient's medical or psychiatric condition. This definition is noted in the data element Minutes of Physical Restraint included with the submission.

[Response Ends]

sp.15. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

Number of psychiatric inpatient days

Denominator basis: per 1,000 hours

To compute this measure rate, a base of 1000 hours has been applied to total patient days in the denominator (i.e., total patient days are divided by 1000). The purpose of this is to create a smaller denominator number, thus providing a more understandable rate. When multiplied by 1000, this rate measures numerator occurrence per total patient days.

[Response Ends]

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

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

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

[Response Begins]

Seven 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. 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.
4. Psychiatric Inpatient Days - Medicare Only* - The sum of the number of days each Medicare patient was included in the psychiatric inpatient census during the month (includes clients on leave status).
5. Psychiatric Inpatient Days – Non-Medicare Only* - The sum of the number of days each non-Medicare patient was included in the psychiatric inpatient census during the month (includes clients on leave status).
6. Total Leave Days - Medicare Only* - The aggregate number of leave days for Medicare patients during the month.
7. Total Leave Days – Non-Medicare Only* - The aggregate number of leave days for non-Medicare patients during the month.

* The distinction between Medicare and Non-Medicare was added to account for the adoption of the HBIPS measures by the Centers for Medicare and Medicaid Services (CMS) Inpatient Psychiatric Facilities Quality Reporting Program

Populations: All psychiatric inpatient days.

[Response Ends]

sp.17. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

Total leave days

[Response Ends]

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

All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

- Patients who are on leave defined as an authorized or unauthorized absence of the patient from a psychiatric care setting, excluding discharges, during which the patient is absent from the psychiatric care setting at the time of the daily census and is not under the direct supervision of psychiatric care setting staff while absent.

[Response Ends]

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

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

[Response Begins]

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

[Response Ends]

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

[Response Begins]

[Response Ends]

sp.21. Select the risk adjustment type.

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

[Response Begins]

[Response Ends]

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

Attachment: If available, please provide a sample report.

[Response Begins]

Ratio

[Response Ends]

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

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

[Response Begins]

Better quality = Lower score

[Response Ends]

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

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

[Response Begins]

1. Run all cases that are included in the Initial Patient Population for HBIPS-2 and 3 and pass the edits defined in the Transmission Data Processing Flow: Clinical Through this measure.
2. Check Event Type
 - a. If Event Type equals 2, the case will proceed to a Measure Category Assignment of U for Overall Rate (HBIPS-2a) and will not be in the measure population. Continue processing and proceed to step 4, Initialize the Measure Category Assignment for each strata measure
 - b. If Event Type equals 1, continue processing and proceed to Minutes of Physical Restraint.
3. Check Minutes of Physical Restraint
 - a. If Minutes of Physical Restraint is missing, the case will proceed to a Measure Category Assignment of X for Overall Rate (HBIPS-2a) and will be rejected. Continue processing and proceed to step 4, Initialize the Measure Category Assignment for each strata measure.
 - b. If Minutes of Physical Restraint equals UTD, the case will proceed to a Measure Category Assignment of Y for Overall Rate (HBIPS-2a) and will be in the measure population. Continue processing and proceed to step 4, Initialize the Measure Category Assignment for each strata measure.
 - c. If Minutes of Physical Restraint equals a Not able to Determine Value, the case will proceed to a Measure Category Assignment of E for Overall Rate (HBIPS-2a) and will be in the measure population. Continue processing and proceed to step 4, Initialize the Measure Category Assignment for each strata measure.
4. Check Overall Rate Category Assignment
 - a. If Overall Rate Category Assignment equals U, Set the Measure Category Assignment for the strata measures (HBIPS-2b through HBIPS-2e) = 'U'. Stop processing.
 - b. If Overall Rate Category Assignment equals E, X or Y, continue processing and proceed to Patient Age at Time of Event.
5. Initialize the Measure Category Assignment for each strata measure (b-e) = 'B'. Do not change the Measure Category Assignment or Total Overall Restraint Minutes that was already calculated for the overall rate (HBIPS-2a). The rest of the algorithm will reset the appropriate Measure Category Assignment to be equal to the overall rate's (HBIPS-2a) Measure Category Assignment.
6. Check Patient Age at Time of Event
 - a. If Patient Age at Time of Event is greater than or equal 1 years and less than 13 years, set the Measure Category Assignment for measure HBIPS-2b = Measure Category Assignment for measure HBIPS-2a. Stop processing.
 - b. If Patient Age at Time of Event is greater than or equal 13 years, continue processing and proceed to Patient Age at Time of Event.
7. Check Patient Age at Time of Event
 - a. If Patient Age at Time of Event is greater than or equal 13 years and less than 18 years, set the Measure Category Assignment for measure HBIPS-2c = Measure b. Category Assignment for measure HBIPS-2a. Stop processing.

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

8. Check Patient Age at Time of Event

a. If Patient Age at Time of Event is greater than or equal 18 years and less than 65 years, set the Measure Category Assignment for measure HBIPS-2d = Measure Category Assignment for measure HBIPS-2a. Stop processing.

b. If Patient Age at Time of Event is greater than or equal 65 years, set the Measure Category Assignment for measure HBIPS-2e = Measure Category Assignment for measure HBIPS-2a. Stop processing

[Response Ends]

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

Examples of samples used for testing:

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

- *The sample should represent the variety of entities whose performance will be measured. The [2010 Measure Testing Task Force](#) recognized that the samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.*

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

- *When possible, units of measurement and patients within units should be randomly selected.*

[Response Begins]

The measure is not eligible for sampling.

[Response Ends]

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

[Response Begins]

Electronic Health Records

Paper Medical Records

[Response Ends]

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

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

[Response Begins]

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.

[Response Ends]

sp.32. Provide the data collection instrument.

[Response Begins]

No data collection instrument provided

[Response Ends]

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

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

No

[Response Ends]

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

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

Current Submission:

Updated testing information here.

Previous Submission:

Testing from the previous submission here.

[Response Begins]

Yes

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

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

[Response Begins]

[Response Ends]

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

Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.

All required sections must be completed.

For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.

If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.

An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.

Contact NQF staff with any questions. Check for resources at the [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 the [2021 Measure Evaluation Criteria and Guidance](#).

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

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

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

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

AND

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

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

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

rationale/data support no risk adjustment/ stratification.

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

OR

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

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

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

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

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

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

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

Definitions

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

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

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

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

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

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

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

Current Submission:

Updated testing information here.

Previous (Year) Submission:

Testing from the previous submission here.

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]

[Response Ends]

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

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

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.

Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]

[Response Ends]

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

Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter “see validity testing section of data elements”; and enter “N/A” for 2a.11 and 2a.12.

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

Choose one or both levels.

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

Examples may include correlations or t-test results.

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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 eQMs). 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.

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

Examples may include correlation, and/or rank order.

[Response Begins]

[Response Ends]

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

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

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

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]

[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]

[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]

[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

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

[Response Begins]

[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]

[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]

[Response Ends]

2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter “N/A” for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]

[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]

[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]

[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]

[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]

[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

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

[Response Begins]

[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

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

[Response Begins]

[Response Ends]

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.

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

[Response Begins]

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

[Response Ends]

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

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.

[Response Begins]

Some data elements are in defined fields in electronic sources

[Response Ends]

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

[Response Begins]

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.

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]

[Response Ends]

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

[Response Begins]

[Response Ends]

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

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

Attach the fee schedule here, if applicable.

[Response Begins]

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

[Response Ends]

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.

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

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 demonstrating performance improvement.

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

Name of program and sponsor

URL

Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]

Public Reporting

Payment Program

Regulatory and Accreditation Programs

Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

Quality Improvement (Internal to the specific organization)

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

[Response Ends]

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

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

[Response Begins]

Not applicable

[Response Ends]

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

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

[Response Begins]

Not applicable

[Response Ends]

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

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

[Response Begins]

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.

[Response Ends]

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

[Response Begins]

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.

[Response Ends]

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

[Response Begins]

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.

[Response Ends]

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

[Response Begins]

Queries submitted via the automated feedback system have decreased significantly for the HBIPS measure set in the past 3 years. (522 in 2016, 288 in 2017, 187 for 2018 YTD). There have been no major issues with the data elements for this measure. Trends in the few questions raised focused on determination of time in restraints and time in restraints vs. time in seclusion.

[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]

Same as above in 4a2.2.2.

[Response Ends]

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

[Response Begins]

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.

Updates were made to the Guidelines for Abstraction of the data element Minutes of Physical Restraint to provide clarification for abstracting time in restraints. Notes for Abstraction were additionally updated to clarify the priority for tracking time in restraint vs. seclusion when a patient is placed in restraint and seclusion at the same time.

[Response Ends]

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

[Response Begins]

Though 2009 to 2nd quarter 2018, a Poisson 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 and the number of patient days as the offset variable. The results of the model show statistical significant increase in rates over time ($P < 0.001$) and an estimate of 0.048 for the fixed effect time variable.

[Response Ends]

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

[Response Begins]

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

[Response Ends]

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

[Response Begins]

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 restraint hours significantly dropped over the initial reporting periods, for the 2009 and 2015 cohorts.

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.

[Response Ends]

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.

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

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

NOTE: If there are no related measures, please select N/A.

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus and target population).

NOTE: If there are no competing measures, please select N/A.

(Can search and select measures.)

[Response Begins]

[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]

Not applicable

[Response Ends]

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

[Response Begins]

No

[Response Ends]

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

[Response Begins]

Measure 203 excludes patients under 18 years of age, only evaluates vest or limb restraint use and is based on a survey conducted at selected time intervals. Measure 0687 evaluates the percent of all patients in a long term care setting who had a vest or limb restraint applied during the reporting period. HBIPS-2 evaluates the total time all patients > 1 year of age in a psychiatric care setting were maintained in all forms of physical restraint for the reporting period.

[Response Ends]

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

Provide analyses when possible.

[Response Begins]

Not Applicable

[Response Ends]

Appendix

Supplemental materials may be provided in an appendix.:

No appendix

Contact Information

Measure Steward (Intellectual Property Owner): The Joint Commission

Measure Steward Point of Contact: Alban, JohnMarc, jalban@jointcommission.org

Measure Developer if different from Measure Steward: The Joint Commission

Measure Developer Point(s) of Contact: Alban, JohnMarc, jalban@jointcommission.org

Additional Information

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

[Response Begins]

No appendix

[Response Ends]

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

Describe the members' role in measure development.

[Response Begins]

Ann Doucette, PhD
Claremont Graduate University

Scott Dziengelski
National Association for Behavioral Healthcare

Frank A Ghinassi, PhD, ABPP (Chair)
President and CEO
Rutgers Health, University Behavioral Health Care

Richard Hermann, MD, MS
Tufts University School of Medicine, Tufts-NEMC

Karen E. Johnson, MSW
Universal Health Services, Inc.

Michael Lambert, PhD
Professor
Brigham Young University

Kathleen McCann, RN, PhD
National Association for Behavioral Healthcare

Dr. John Oldham, MD
Baylor College of Medicine

Lucille M Schacht, PhD, CPHQ
NRI, Inc

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

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

[Response Ends]

3. Indicate the year the measure was first released.

[Response Begins]

[Response Ends]

4. Indicate the month and year of the most recent revision.

[Response Begins]

[Response Ends]

5. Indicate the frequency of review, or an update schedule, for this measure.

[Response Begins]

Biannual

[Response Ends]

6. Indicate the next scheduled update or review of this measure.

[Response Begins]

[Response Ends]

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

[Response Begins]

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.

[Response Ends]

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

[Response Begins]

[Response Ends]

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

[Response Begins]

Recent revision is dated January 1, 2019. This represents the date the specifications go into effect. The specifications were published in October 2018.

[Response Ends]

