

## NATIONAL QUALITY FORUM

### Measure Submission and Evaluation Worksheet 5.0

This form contains the information submitted by measure developers/stewards, organized according to NQF's measure evaluation criteria and process. The evaluation criteria, evaluation guidance documents, and a blank online submission form are available on the [submitting standards web page](#).

<b>NQF #:</b> 1938	<b>NQF Project:</b> Behavioral Health
(for Endorsement Maintenance Review)	
<b>Original Endorsement Date:</b>	<b>Most Recent Endorsement Date:</b> Last Updated Date: Oct 01, 2014
<b>BRIEF MEASURE INFORMATION</b>	
<b>De.1 Measure Title:</b> Emergency department utilization for mental health conditions by people with schizophrenia	
<b>Co.1.1 Measure Steward:</b> National Committee for Quality Assurance	
<b>De.2 Brief Description of Measure:</b> The percentage of individuals 25 – 64 years of age with a schizophrenia diagnosis who had an emergency department admission for mental health.	
<b>2a1.1 Numerator Statement:</b> An admission to the ED with a mental health diagnosis.	
<b>2a1.4 Denominator Statement:</b> Adults 25 – 64 years of age as of December 31 of the measurement year with a schizophrenia diagnosis.	
<b>2a1.8 Denominator Exclusions:</b> Not applicable.	
<b>1.1 Measure Type:</b> Process <b>2a1. 25-26 Data Source:</b> Administrative claims <b>2a1.33 Level of Analysis:</b> Population : State	
<b>1.2-1.4 Is this measure paired with another measure?</b> No	
<b>De.3 If included in a composite, please identify the composite measure (title and NQF number if endorsed):</b> Not applicable.	

<b>STAFF NOTES</b> (issues or questions regarding any criteria)
<b>Comments on Conditions for Consideration:</b>
<b>Is the measure untested?</b> Yes <input checked="" type="radio"/> No <input checked="" type="radio"/> If untested, explain how it meets criteria for consideration for time-limited endorsement:
<b>1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (check De.5):</b> <b>5. Similar/related <a href="#">endorsed</a> or submitted measures (check 5.1):</b> <b>Other Criteria:</b>
<b>Staff Reviewer Name(s):</b>

<b>1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT</b>
Importance to Measure and Report is a threshold criterion that must be met in order to recommend a

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable  
Created on: 10/02/2014 at 02:04 AM

measure for endorsement. All three subcriteria must be met to pass this criterion. See [guidance on evidence](#).

**Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)**

**1a. High Impact: H ● M ● L ● I ●**

(The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact aspect of healthcare.)

**De.4 Subject/Topic Areas** (Check all the areas that apply): [Mental Health, Mental Health : Serious Mental Illness, Prevention](#)

**De.5 Cross Cutting Areas** (Check all the areas that apply): [Care Coordination](#)

**1a.1 Demonstrated High Impact Aspect of Healthcare:** [Affects large numbers, High resource use, Patient/societal consequences of poor quality](#)

**1a.2 If "Other," please describe:**

**1a.3 Summary of Evidence of High Impact** (Provide epidemiologic or resource use data):

[Emergency department use for people with schizophrenia is not well understood compared to other vulnerable populations. Higher rates of medical comorbidities and earlier mortality for people with schizophrenia makes identifying rates of ED use for people with schizophrenia that much more important. ED use is also associated with higher costs and patients are less likely to get a care plan, which is also important for people with schizophrenia. Several studies show higher ED use in their sample populations for people with schizophrenia.](#)

[In a systematic review of the quality of medical care for people with and without comorbid mental illness and substance misuse, the review found that people with a mental health diagnosis are more likely to receive treatment in the ED, use the ED for their routine medical needs.](#)

**1a.4 Citations for Evidence of High Impact cited in 1a.3:** [Hackman AL, Goldberg RW, Brown CH, Fang LJ, Dickerson FB, Wohlheiter K, et al. Use of emergency department services for somatic reasons by people with serious mental illness. Psychiatr Serv 2006 ; 57: 563 -6.](#)

[Mitchell AJ, Malone D, Doebbeling CC. Quality of medical care for people with and without comorbid mental illness and substance misuse: systematic review of comparative studies. Br J Psychiatry. 2009;194:491–499.](#)

[Nossel, IR, Calmes, CA, et al. Patterns of Emergency Department Use for Medical Conditions Among Persons with Serious Mental Illness. Psychiatr Serv. 2010 Dec;61\(12\):1251-4.](#)

**1b. Opportunity for Improvement: H ● M ● L ● I ●**

(There is a demonstrated performance gap - variability or overall less than optimal performance)

**1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:**

[As patients with schizophrenia are at an increased risk for bad health outcomes, these patients have higher rates of emergency department use. This measure will allow state Medicaid programs to identify high patterns of ED use for patients with schizophrenia and target interventions to reduce ED use at those sites.](#)

**1b.2 Summary of Data Demonstrating Performance Gap** (Variation or overall less than optimal performance across providers): **[For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.]**  
[Not applicable.](#)

**1b.3 Citations for Data on Performance Gap: [For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]**

Not applicable.

**1b.4 Summary of Data on Disparities by Population Group: [For Maintenance –Descriptive statistics for performance results for this measure by population group]**

Not applicable.

**1b.5 Citations for Data on Disparities Cited in 1b.4: [For Maintenance – Description of the data or sample for measure results reported in 1b.4 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included]**

Not applicable.

**1c. Evidence (Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.)**

**Is the measure focus a health outcome? Yes ☐ No ☐ If not a health outcome, rate the body of evidence.**

**Quantity: H ☐ M ☐ L ☐ I ☐ Quality: H ☐ M ☐ L ☐ I ☐ Consistency: H ☐ M ☐ L ☐ I ☐**

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?
M-H	M-H	M-H	Yes <input type="radio"/>
L	M-H	M	Yes <input type="radio"/> IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No <input type="radio"/>
M-H	L	M-H	Yes <input type="radio"/> IF potential benefits to patients clearly outweigh potential harms: otherwise No <input type="radio"/>
L-M-H	L-M-H	L	No <input type="radio"/>

**Health outcome** – rationale supports relationship to at least one healthcare structure, process, intervention, or service

**Does the measure pass subcriterion1c?**  
Yes ☐ IF rationale supports relationship

**1c.1 Structure-Process-Outcome Relationship (Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome):**

This is a process measure (process-health outcome) that captures ED utilization for individuals with schizophrenia.

**1c.2-3 Type of Evidence (Check all that apply):**

Selected individual studies (rather than entire body of evidence)

**1c.4 Directness of Evidence to the Specified Measure (State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population):**

The evidence suggests that the ED visits are more common for people with schizophrenia than with the general population. This variance in the type of care suggests that people with schizophrenia are not receiving quality care in the most efficient and effective manner, which is driving their worse health outcomes.

**1c.5 Quantity of Studies in the Body of Evidence** (Total number of studies, not articles): **3**

**1c.6 Quality of Body of Evidence** (Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events): This measure concept is supported by a systematic literature review and individual studies.

**1c.7 Consistency of Results across Studies** (Summarize the consistency of the magnitude and direction of the effect): The studies consistently show that individuals with schizophrenia have higher rates of ED use.

**1c.8 Net Benefit** (Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms):

Benefit: This measure will allow state Medicaid programs to identify high patterns of ED use for patients with schizophrenia and target interventions to reduce ED use at those sites.

**1c.9 Grading of Strength/Quality of the Body of Evidence.** Has the body of evidence been graded? **No**

**1c.10 If body of evidence graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias:** Not applicable.

**1c.11 System Used for Grading the Body of Evidence:** Other

**1c.12 If other, identify and describe the grading scale with definitions:** The body of evidence was not graded.

**1c.13 Grade Assigned to the Body of Evidence:** Not applicable.

**1c.14 Summary of Controversy/Contradictory Evidence:** Not applicable.

**1c.15 Citations for Evidence other than Guidelines(Guidelines addressed below):**  
Not applicable.

**1c.16 Quote verbatim, the specific guideline recommendation** (Including guideline # and/or page #):  
Not applicable.

**1c.17 Clinical Practice Guideline Citation:** Not applicable.

**1c.18 National Guideline Clearinghouse or other URL:** Not applicable.

**1c.19 Grading of Strength of Guideline Recommendation.** Has the recommendation been graded? **No**

**1c.20 If guideline recommendation graded, identify the entity that graded the evidence including balance of representation and any disclosures regarding bias:**

**1c.21 System Used for Grading the Strength of Guideline Recommendation:** Other

**1c.22 If other, identify and describe the grading scale with definitions:** The body of was not graded

**1c.23 Grade Assigned to the Recommendation:** Not applicable.

**1c.24 Rationale for Using this Guideline Over Others:** [Not applicable.](#)

**Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?**

**1c.25** Quantity: [Moderate](#) **1c.26** Quality: [Moderate](#) **1c.27** Consistency: [Moderate](#)

**1c.28** Attach evidence submission form:

**1c.29** Attach appendix for supplemental materials:

**Was the threshold criterion, *Importance to Measure and Report*, met?**

**(1a & 1b must be rated moderate or high and 1c yes)** Yes ☐ No ☒

**Provide rationale based on specific subcriteria:**

**For a new measure if the Committee votes NO, then STOP.**

**For a measure undergoing endorsement maintenance, if the Committee votes NO because of 1b. (no opportunity for improvement), it may be considered for continued endorsement and all criteria need to be evaluated.**

## 2. RELIABILITY & 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. (**evaluation criteria**)

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 field. Supplemental materials may be referenced or attached in item 2.1. See [guidance on measure testing](#).

**S.1 Measure Web Page** (*In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained*). Do you have a web page where current detailed specifications for this measure can be obtained? [No](#)

**S.2** If yes, provide web page URL:

**2a. RELIABILITY. Precise Specifications and Reliability Testing:** ☒ H ☒ M ☒ L ☒ I ☒ NA

**2a1. Precise Measure Specifications.** (*The measure specifications precise and unambiguous.*)

**2a1.1 Numerator Statement** (*Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome*):

[An admission to the ED with a mental health diagnosis.](#)

**2a1.2 Numerator Time Window** (*The time period in which the target process, condition, event, or outcome is eligible for inclusion*):

[The measurement year.](#)

**2a1.3 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, codes with descriptors, and/or specific data collection items/responses*):

[Codes to identify visit type:](#)

[Any Emergency Department visit type:](#)

[CPT code: 99281-99285 with visit related mental health ICD-9 diagnosis code: 290, 293, 295-302, 306-316](#)

**2a1.4 Denominator Statement** (*Brief, narrative description of the target population being measured*):

[Adults 25 – 64 years of age as of December 31 of the measurement year with a schizophrenia diagnosis.](#)

**2a1.5 Target Population Category** (Check all the populations for which the measure is specified and tested if any): [Senior Care](#)

**2a1.6 Denominator Time Window** (The time period in which cases are eligible for inclusion):  
[The measurement year.](#)

**2a1.7 Denominator Details** (All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):

- Medicaid beneficiaries 25 – 64 years of age as of December 31 of the measurement year
- Two separate claims with schizophrenia as a primary diagnosis or one inpatient claim with schizophrenia as a primary diagnosis in the measurement year
- 10 months continuous enrollment during the measurement year

Codes to identify schizophrenia diagnosis:

ICD-9-CM Diagnosis: 295

ICD-10-CM Diagnosis: F20, F25.9

Codes to identify visit type:

Acute inpatient

UB Revenue: 010x, 0110-0114, 0119, 0120-0124, 0129, 0130-0134, 0139, 0140-0144, 0149, 0150-0154, 0159, 016x, 020x, 021x, 072x, 0987

CPT: 90801, 90802, 90816-90819, 90821-90824, 90826-90829, 90845, 90847, 90849, 90853, 90857, 90862, 90870, 90875, 90876, 99221-99223, 99231-99233, 99238, 99239, 99251-99255, 99291

WITH

POS: 21, 51

Outpatient, intensive outpatient and partial hospitalization

CPT: 90804-90815, 98960-98962, 99078, 99201-99205, 99211-99215, 99217-99220, 99241-99245, 99341-99345, 99347-99350, 99385-99387, 99395-99397, 99401-99404, 99411, 99412, 99510

HCPCS: G0155, G0176, G0177, G0409-G0411, H0002, H0004, H0031, H0034-H0037, H0039, H0040, H2000, H2001, H2010-H2020, M0064, S0201, S9480, S9484, S9485

UB Revenue: 0510, 0513, 0515-0517, 0519-0523, 0526-0529, 0900, 0901, 0902-0905, 0907, 0911-0917, 0919, 0982, 0983

CPT: 90801, 90802, 90816-90819, 90821-90824, 90826-90829, 90845, 90847, 90849, 90853, 90857, 90862, 90870, 90875, 90876, 99221-99223, 99231-99233, 99238, 99239, 99251-99255

WITH

POS: 03, 05, 07, 09, 11, 12, 13, 14, 15, 20, 22, 24, 33, 49, 50, 52, 53, 71, 72

ED

CPT: 99281-99285

UB Revenue: 045x, 0981

CPT: 90801, 90802, 90845, 90847, 90849, 90853, 90857, 90862, 90870, 90875, 90876, WITH

POS: 23



Nonacute inpatient

CPT: 99304-99310, 99315, 99316, 99318, 99324-99328, 99334-99337

HCPCS: H0017-H0019, T2048

UB Revenue: 0118, 0128, 0138, 0148, 0158, 019x, 0524, 0525, 055x, 066x, 1000, 1001, 1003-1005

CPT: 90801, 90802, 90816-90819, 90821-90824, 90826-90829, 90845, 90847, 90849, 90853, 90857, 90862, 90870, 90875, 90876

WITH

POS: 31, 32, 56

**2a1.8 Denominator Exclusions** (Brief narrative description of exclusions from the target population):  
Not applicable.

**2a1.9 Denominator Exclusion Details** (All information required to identify and calculate exclusions from the denominator such as definitions, codes with descriptors, and/or specific data collection items/responses):  
Not applicable.

**2a1.10 Stratification Details/Variables** (All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses):  
Not applicable.

**2a1.11 Risk Adjustment Type** (Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13): No risk adjustment or risk stratification    **2a1.12 If "Other," please describe:**

**2a1.13 Statistical Risk Model and Variables** (Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.):  
Not applicable.

**2a1.14-16 Detailed Risk Model Available at Web page URL** (or attachment). Include coefficients, equations, codes with descriptors, definitions, and/or specific data collection items/responses. Attach documents only if they are not available on a webpage and keep attached file to 5 MB or less. NQF strongly prefers you make documents available at a Web page URL. Please supply login/password if needed:

**2a1.17-18. Type of Score:** Rate/proportion

**2a1.19 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 = Lower score

**2a1.20 Calculation Algorithm/Measure Logic** (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; aggregating data; risk adjustment; etc.):  
1. Determine the eligible population. The eligible population is all individuals who satisfy all specified criteria, including any age, continuous enrollment, benefit, event, or anchor date enrollment requirement.

2. Search administrative systems to identify numerator events for all individuals in the eligible population.
3. Calculate the rate.

**2a1.21-23 Calculation Algorithm/Measure Logic Diagram URL or attachment:**

**2a1.24 Sampling (Survey) Methodology.** If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):  
[Not applicable.](#)

**2a1.25 Data Source** (*Check all the sources for which the measure is specified and tested*). If other, please describe:  
[Administrative claims](#)

**2a1.26 Data Source/Data Collection Instrument** (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*): [Not applicable.](#)

**2a1.27-29 Data Source/data Collection Instrument Reference Web Page URL or Attachment:**

**2a1.30-32 Data Dictionary/Code Table Web Page URL or Attachment:**

**2a1.33 Level of Analysis** (*Check the levels of analysis for which the measure is specified and tested*):  
[Population : State](#)

**2a1.34-35 Care Setting** (*Check all the settings for which the measure is specified and tested*): [Other:Any outpatient setting represented with Medicaid claims data](#)

**2a2. Reliability Testing.** (*Reliability testing was conducted with appropriate method, scope, and adequate demonstration of reliability.*)

**2a2.1 Data/Sample** (*Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*):  
[Using Medicaid Analytic Extract \(MAX\) claims data from 2007 we included beneficiaries from 22 states who met the following criteria \(1\) enrolled in fee-for-service plans\\* \(2\) disability as the basis of eligibility; and \(3\) continuously enrolled in Medicaid for 10 months. From these beneficiaries we drew two analytic samples. Beneficiaries who had a primary diagnosis of schizophrenia on either one inpatient or two outpatient claims on different days were included in our schizophrenia sample. Overall, there were 98,412 beneficiaries in the schizophrenia sample.](#)

[Data from the following states were included in both analytic samples: Alabama, Alaska, California, Connecticut, DC, Georgia, Idaho, Illinois, Indiana, Iowa, Louisiana, Maryland, Missouri, Mississippi, Nevada, New Hampshire, North Carolina, North Dakota, Oklahoma, South Dakota, West Virginia and Wyoming.](#)

[Beneficiaries ranged in age from 25 – 64 years. Just under half of the schizophrenia population was female](#)



(49.2%). About 7% and 34% of the sample was Hispanic and African-American, respectively.

(\*Beneficiaries enrolled in managed care plans (e.g. BHO or HMO plans) that provided usable claims records were included. About 1% of the schizophrenia sample was enrolled in a BHO (1.4%) and 11.5% were enrolled in an HMO).

#### **2a2.2 Analytic Method** *(Describe method of reliability testing & rationale):*

The relevant unit of analysis for the proposed measures is aggregated state-level performance. Therefore, we conducted an analysis of test-retest reliability for state results to assess the reliability of state-level performance. To assess stability of state-level performance over time, we computed quartiles of performance based on the state distribution for each measure and assigned each state a score reflecting each state's performance relative to other states in the distribution during the measurement year. For example, a state in the top quartile of all states in 2007 for a given measure would be assigned a performance quartile score of '1' for 2007. This method was replicated for each measure. Next, we repeated this method using 2008 claims data and examined stability of performance quartile between 2007 and 2008.

We also report Pearson correlations measuring the association between 2007 and 2008 measure performance for the 16 states with data.

#### **2a2.3 Testing Results** *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted):*

In general, the measure showed good test-retest reliability. Overall, 7 of 16 states (44%) had no change in performance quartile between 2007 and 2008. State performance was correlated at  $r=0.42$ , indicating that 2007 performance on this measure accounted for 18% of the variance in 2008 scores.

### **2b. VALIDITY. Validity, Testing, including all Threats to Validity: H M L I NA**

#### **2b1.1 Describe how the measure specifications (measure focus, target population, and exclusions) are consistent with the evidence cited in support of the measure focus (criterion 1c) and identify any differences from the evidence:**

The evidence suggests that the ED visits are more common for people with schizophrenia than with the general population. This variance in the type of care suggests that people with schizophrenia are not receiving quality care in the most efficient and effective manner, which is driving their worse health outcomes.

#### **2b2. Validity Testing.** *(Validity testing was conducted with appropriate method, scope, and adequate demonstration of validity.)*

##### **2b2.1 Data/Sample** *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*

Using Medicaid Analytic Extract (MAX) claims data from 2007 we included beneficiaries from 22 states who met the following criteria (1) enrolled in fee-for-service plans\* (2) disability as the basis of eligibility; and (3) continuously enrolled in Medicaid for 10 months. From these beneficiaries we drew two analytic samples. Beneficiaries who had a primary diagnosis of schizophrenia on either one inpatient or two outpatient claims on different days were included in our schizophrenia sample. Overall, there were 98,412 beneficiaries in the schizophrenia sample.

Data from the following states were included in both analytic samples: Alabama, Alaska, California, Connecticut, DC, Georgia, Idaho, Illinois, Indiana, Iowa, Louisiana, Maryland, Missouri, Mississippi, Nevada, New Hampshire, North Carolina, North Dakota, Oklahoma, South Dakota, West Virginia and Wyoming.

Beneficiaries ranged in age from 25 – 64 years. Just under half of the schizophrenia population was female

(49.2%). About 7% and 34% of the sample was Hispanic and African-American, respectively.

(\*Beneficiaries enrolled in managed care plans (e.g. BHO or HMO plans) that provided usable claims records were included. About 1% of the schizophrenia sample was enrolled in a BHO (1.4%) and 11.5% were enrolled in an HMO).

**2b2.2 Analytic Method** *(Describe method of validity testing and rationale; if face validity, describe systematic assessment):*

Validity was assessed using several complementary methods.

Face validity was assessed through a multistakeholder Technical Advisory Group responsible for overseeing measure development. Additionally, face validity was captured through a public comment period and a series of focus groups involving the Medicaid Medical Directors Learning Network, Managed Behavioral Health Care Organizations, and State Mental Health Commissioners and Medical Directors. The panelists assessed the usability and feasibility of the measures.

Concurrent validity was assessed via Medicaid resource utilization from the Medicaid claims data. We examined rates of schizophrenia-related hospital and emergency room utilization as well as total Medicaid costs comparing beneficiaries in the highest and lowest performance quartiles for each measure.

Convergent and discriminant validity were assessed using the Medicaid Analytic Extract (MAX) from Medicaid claims in using 2007 data. Pearson correlation coefficients were used to assess measure correlations. We hypothesized similar measures (e.g. screening and monitoring) would be correlated and (b) process measures would have negative correlations with measures of adverse events (e.g. mental health emergency room utilization).

**2b2.3 Testing Results** *(Statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment):*

Face validity:

The measures were deemed important, usable, and feasible to collect by the Technical Advisory Group overseeing the measure development, as well as focus groups with the Medicaid Medical Directors Learning Network, Managed Behavioral Healthcare Organizations, and State Mental Health Commissioners and Medical Directors.

Among 22 states, the measure had a minimum value of 22.3%, mean=31.0%, 25th percentile=26.8%, median=32.5%, 75th percentile=34.4% and a maximum value of 36.8%.

Concurrent validity:

Beneficiaries in the lowest performing states for the measure had higher rates of ED use than highest performing states (27.0% versus 14.9%, respectively).

**POTENTIAL THREATS TO VALIDITY.** *(All potential threats to validity were appropriately tested with adequate results.)*

**2b3. Measure Exclusions.** *(Exclusions were supported by the clinical evidence in 1c or appropriately tested with results demonstrating the need to specify them.)*

**2b3.1 Data/Sample for analysis of exclusions** *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*  
Not applicable.

**2b3.2 Analytic Method** *(Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference):*

Not applicable.

**2b3.3 Results** *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*

Not applicable.

**2b4. Risk Adjustment Strategy.** *(For outcome measures, adjustment for differences in case mix (severity) across measured entities was appropriately tested with adequate results.)*

**2b4.1 Data/Sample** *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*

Not applicable.

**2b4.2 Analytic Method** *(Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables):*

Not applicable.

**2b4.3 Testing Results** *(Statistical risk model: Provide quantitative assessment of relative contribution of model risk factors; risk model performance metrics including cross-validation discrimination and calibration statistics, calibration curve and risk decile plot, and assessment of adequacy in the context of norms for risk models. Risk stratification: Provide quantitative assessment of relationship of risk factors to the outcome and differences in outcomes among the strata):*

Not applicable.

**2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment:** Not applicable.

**2b5. Identification of Meaningful Differences in Performance.** *(The performance measure scores were appropriately analyzed and discriminated meaningful differences in quality.)*

**2b5.1 Data/Sample** *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*

Using Medicaid Analytic Extract (MAX) claims data from 2007 we included beneficiaries from 22 states who met the following criteria (1) enrolled in fee-for-service plans\* (2) disability as the basis of eligibility; and (3) continuously enrolled in Medicaid for 10 months. From these beneficiaries we drew two analytic samples. Beneficiaries who had a primary diagnosis of schizophrenia on either one inpatient or two outpatient claims on different days were included in our schizophrenia sample. Overall, there were 98,412 beneficiaries in the schizophrenia sample.

Data from the following states were included in both analytic samples: Alabama, Alaska, California, Connecticut, DC, Georgia, Idaho, Illinois, Indiana, Iowa, Louisiana, Maryland, Missouri, Mississippi, Nevada, New Hampshire, North Carolina, North Dakota, Oklahoma, South Dakota, West Virginia and Wyoming.

Beneficiaries ranged in age from 25 – 64 years. Just under half of the schizophrenia population was female (49.2%). About 7% and 34% of the sample was Hispanic and African-American, respectively.

(\*Beneficiaries enrolled in managed care plans (e.g. BHO or HMO plans) that provided usable claims records were included. About 1% of the schizophrenia sample was enrolled in a BHO (1.4%) and 11.5% were enrolled in an HMO).

**2b5.2 Analytic Method** *(Describe methods and rationale to identify statistically significant and*

*practically/meaningfully differences in performance):*

Pearson correlations, means and percentiles are reported.

**2b5.3 Results** *(Provide measure performance results/scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance):*

Among 22 states, the measure had a minimum value of 22.3%, mean=31.0%, 25th percentile=26.8%, median=32.5%, 75th percentile=34.4% and a maximum value of 36.8%.

**2b6. Comparability of Multiple Data Sources/Methods.** *(If specified for more than one data source, the various approaches result in comparable scores.)*

**2b6.1 Data/Sample** *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included):*

Not applicable.

**2b6.2 Analytic Method** *(Describe methods and rationale for testing comparability of scores produced by the different data sources specified in the measure):*

Not applicable.

**2b6.3 Testing Results** *(Provide statistical results, e.g., correlation statistics, comparison of rankings; assessment of adequacy in the context of norms for the test conducted):*

Not applicable.

**2c. Disparities in Care: H ☐ M ☐ L ☐ I ☐ NA ☐** *(If applicable, the measure specifications allow identification of disparities.)*

**2c.1 If measure is stratified for disparities, provide stratified results** *(Scores by stratified categories/cohorts):* Not applicable.

**2c.2 If disparities have been reported/identified (e.g., in 1b), but measure is not specified to detect disparities, please explain:**

Not applicable.

**2.1-2.3 Supplemental Testing Methodology Information:**

**Steering Committee: Overall, was the criterion, Scientific Acceptability of Measure Properties, met? (Reliability and Validity must be rated moderate or high) Yes ☐ No ☐**

Provide rationale based on specific subcriteria:

**If the Committee votes No, STOP**

### 3. USABILITY

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. (evaluation criteria)

**C.1 Intended Actual/Planned Use** *(Check all the planned uses for which the measure is intended):* Public Reporting, Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

**3.1 Current Use** *(Check all that apply; for any that are checked, provide the specific program information in the following questions):* Not in use

### 3a. Usefulness for Public Reporting: H● M● L● I●

(The measure is meaningful, understandable and useful for public reporting.)

**3a.1. Use in Public Reporting - disclosure of performance results to the public at large** (If used in a public reporting program, provide name of program(s), locations, Web page URL(s)). If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: **[For Maintenance – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.]**

This measure may be used to evaluate State Medicaid programs. The public comment and Technical Advisory Group generally supported this measure for public reporting.

**3a.2. Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting.** If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: The measure captures important variation in performance between states. High and low performers (first v. last quartile) on the measure were had better performance on related metrics (e.g., hospital and ED utilization), and high performers tended to be high performers over time, except in smaller states. The measure was rated as meaningful, understandable, and useful for public reporting by participants in our focus groups with representatives from the Medicaid Medical Directors Learning Network, state mental health program directors, and MBHOs. These representatives were drawn from a diverse array of states and have expertise in the construction of claims/encounter-based measurement of quality and use for public reporting programs.

**3.2 Use for other Accountability Functions (payment, certification, accreditation).** If used in a public accountability program, provide name of program(s), locations, Web page URL(s): Not applicable.

### 3b. Usefulness for Quality Improvement: H● M● L● I●

(The measure is meaningful, understandable and useful for quality improvement.)

**3b.1. Use in QI.** If used in quality improvement program, provide name of program(s), locations, Web page URL(s):

**[For Maintenance – If not used for QI, indicate the reasons and describe progress toward using performance results for improvement].**

This proposed measure will be used to evaluate State Medicaid programs.

**3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement.** If usefulness was demonstrated (e.g., QI initiative), describe the data, method and results:

The measure was deemed usable and feasible by the focus groups, public comment, and Technical Advisory Group.

**Overall, to what extent was the criterion, Usability, met? H● M● L● I●**

**Provide rationale based on specific subcriteria:**

## 4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)

### 4a. Data Generated as a Byproduct of Care Processes: H● M● L● I●

**4a.1-2 How are the data elements needed to compute measure scores generated?** (Check all that apply).



Data used in the measure are:

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

#### 4b. Electronic Sources: H ☐ M ☐ L ☐ I ☐

**4b.1 Are the data elements needed for the measure as specified available electronically** (*Elements that are needed to compute measure scores are in defined, computer-readable fields*): ALL data elements in electronic claims

**4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources:**

#### 4c. Susceptibility to Inaccuracies, Errors, or Unintended Consequences: H ☐ M ☐ L ☐ I ☐

**4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:**

Validity and reliability testing of the measures has been performed. To our knowledge, there are no known inaccuracies, errors, or unintended consequences of measurement identified during testing, however, there may be potential for underreporting of services that are not billed by Medicaid.

#### 4d. Data Collection Strategy/Implementation: H ☐ M ☐ L ☐ I ☐

**A.2 Please check if either of the following apply** (*regarding proprietary measures*): Proprietary measure

**4d.1 Describe what you have learned/modified 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** (e.g., fees for use of proprietary measures):

The proposed measure is claims-based. We identified 22 states in the Medicaid Analytic Extract (MAX) data files with valid and reliable claims data, and we were able to calculate the measure for all states. We observed substantial variability in performance between states, but we believe in nearly all cases that those are related to performance differences rather than data availability differences. Based upon our focus group testing with representatives from the Medicaid Medical Directors Learning Network, state mental health program directors, and MBHOs, we have confidence that states are able to capture these performance data in claims/encounter systems and are capable of programming, reporting, and using the metric.

**Overall, to what extent was the criterion, *Feasibility*, met?** H ☐ M ☐ L ☐ I ☐

**Provide rationale based on specific subcriteria:**

### OVERALL SUITABILITY FOR ENDORSEMENT

**Does the measure meet all the NQF criteria for endorsement?** Yes ☐ No ☐

**Rationale:**

**If the Committee votes No, STOP.**

**If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.**

### 5. COMPARISON TO RELATED AND 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 before a final recommendation is made.



**5.1 If there are related measures (*either same measure focus or target population*) or competing measures (*both the same measure focus and same target population*), list the NQF # and title of all related and/or competing measures:**

#### 5a. Harmonization

**5a.1 If this measure has EITHER the same measure focus OR the same target population as [NQF-endorsed measure\(s\)](#): Are the measure specifications completely harmonized?**

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

#### 5b. Competing Measure(s)

**5b.1 If this measure has 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.](#)

### CONTACT INFORMATION

**Co.1 Measure Steward (Intellectual Property Owner):** [National Committee for Quality Assurance, 1100 13th Street, NW, District Of Columbia, 20005](#)

**Co.2 Point of Contact:** [Bob, Rehm, \[nqf@ncqa.org\]\(mailto:nqf@ncqa.org\), 202-955-1728-](#)

**Co.3 Measure Developer if different from Measure Steward:** [National Committee for Quality Assurance, 1100 13th St, NW, Suite 1000, Washington, District Of Columbia, 20005](#)

**Co.4 Point of Contact:** [Jill Marie, Farrell, \[farrell@ncqa.org\]\(mailto:farrell@ncqa.org\), 202-955-1785-](#)

**Co.5 Submitter:** [Rita, Lewis, MPH, \[lewis@ncqa.org\]\(mailto:lewis@ncqa.org\), 202-955-5102-, National Committee for Quality Assurance](#)

**Co.6 Additional organizations that sponsored/participated in measure development:**  
[Mathematica Policy Research, Inc.](#)

**Co.7 Public Contact:** [Bob, Rehm, Assistant Vice President, Performance Measurement, \[rehm@ncqa.org\]\(mailto:rehm@ncqa.org\), 202-955-1728-, National Committee for Quality Assurance](#)

### ADDITIONAL INFORMATION

**Workgroup/Expert Panel involved in measure development**

**Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.**

[Technical Advisory Group Roster](#)

[Alisa Busch, MD, MS](#)

[Enola Proctor, PhD, MSW](#)

[David Shern, PhD](#)

[Wilma Townsend, MSW](#)

Dan Ford, MD, MPH  
Lorrie Rickman-Jones, PhD  
Eric Hamilton  
Alexander Young, MD, MHS  
Peter Delany, PhD  
Ben Druss, MD, MPH  
Maureen Corcoran, MSN, MBA  
Mike Fitzpatrick, MSW  
Anita Yuskas  
Bob Heinssen, PhD

**Consultants:**

Lisa Dixon, MD, MPH  
Julie Kreyenbul, PharmD, PhD

The Technical Advisory Group advised Mathematica Policy Research, Inc. and the National Committee for Quality Assurance during measure development. The TAG was responsible for providing feedback on measure concepts, specifications, results from field and data testing. The TAG consisted of a multistakeholder group of experts with knowledge in behavioral health and quality measurement.

**Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward:**

**Measure Developer/Steward Updates and Ongoing Maintenance**

**Ad.3 Year the measure was first released:**

**Ad.4 Month and Year of most recent revision:**

**Ad.5 What is your frequency for review/update of this measure?**

**Ad.6 When is the next scheduled review/update for this measure?**

**Ad.7 Copyright statement:** © 2012 by the National Committee for Quality Assurance  
1100 13th Street, NW, Suite 1000  
Washington, DC 20005

**Ad.8 Disclaimers:**

**Ad.9 Additional Information/Comments:**

**Date of Submission (MM/DD/YY):** 02/14/2012