

# 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> 0349 <b>NQF Project:</b> <a href="#">Patient Safety Measures-Complications Project</a>
(for Endorsement Maintenance Review) <b>Original Endorsement Date:</b> <a href="#">May 15, 2008</a> <b>Most Recent Endorsement Date:</b> <a href="#">May 15, 2008</a>
<b>BRIEF MEASURE INFORMATION</b>
<b>De.1 Measure Title:</b> <a href="#">Transfusion Reaction (PSI 16)</a>
<b>Co.1.1 Measure Steward:</b> <a href="#">Agency for Healthcare Research and Quality</a>
<b>De.2 Brief Description of Measure:</b> <a href="#">The count of medical and surgical discharges for patients age greater than or equal to 18 or in MDC 14 with ICD-9-CM code for transfusion reaction in any secondary diagnosis field.</a>
<b>2a1.1 Numerator Statement:</b> <a href="#">Discharges 18 years and older or in MDC 14 with ICD-9-CM codes for transfusion reaction in any secondary diagnosis field of all medical and surgical discharges defined by specific DRGs or MS-DRGs</a>
See Patient Safety Indicators Appendices: <ul style="list-style-type: none"> <li>- Appendix B – Medical Discharge DRGs</li> <li>- Appendix C – Medical Discharge MS-DRGs</li> <li>- Appendix D – Surgical Discharge DRGs</li> <li>- Appendix E – Surgical Discharge MS-DRGs</li> </ul> Link to PSI appendices: <a href="http://qualityindicators.ahrq.gov/Downloads/Software/SAS/V43/TechnicalSpecifications/PSI%20Appendices.pdf">http://qualityindicators.ahrq.gov/Downloads/Software/SAS/V43/TechnicalSpecifications/PSI%20Appendices.pdf</a>  Exclude cases: <ul style="list-style-type: none"> <li>-with principal diagnosis of transfusion reaction or secondary diagnosis present on admission</li> <li>-with missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing) or principal diagnosis (DX1=missing)</li> </ul>
<b>2a1.4 Denominator Statement:</b> <a href="#">Not applicable</a>
<b>2a1.8 Denominator Exclusions:</b> <a href="#">Not applicable</a>
<b>1.1 Measure Type:</b> <a href="#">Outcome</a> <b>2a1. 25-26 Data Source:</b> <a href="#">Administrative claims</a> <b>2a1.33 Level of Analysis:</b> <a href="#">Facility</a>  <b>1.2-1.4 Is this measure paired with another measure?</b> <a href="#">No</a>  <b>De.3 If included in a composite, please identify the composite measure (title and NQF number if endorsed):</b> <a href="#">Not applicable</a>

<b>STAFF NOTES</b> <i>(issues or questions regarding any criteria)</i>
<b>Comments on Conditions for Consideration:</b>
<b>Is the measure untested?</b> Yes <input type="checkbox"/> No <input type="checkbox"/> <b>If untested, explain how it meets criteria for consideration for time-limited</b>

<b>endorsement:</b>
<b>1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure</b> ( <i>check De.5</i> ): <b>5. Similar/related <u>endorsed</u> or submitted measures</b> ( <i>check 5.1</i> ): <b>Other Criteria:</b>
<b>Staff Reviewer Name(s):</b>

**1. IMPACT, OPPORTUNITY, EVIDENCE - IMPORTANCE TO MEASURE AND REPORT**

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a 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): [Surgery : General Surgery](#)

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

**1a.1 Demonstrated High Impact Aspect of Healthcare:** [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):

There were 153 cases of transfusion reaction in 2008 in SID participating states; however, 89 of these cases were reported as present on admission, leaving only 64 that occurred as a result of care during the same hospital stay [1]. Excess resource use among such cases is 3.4 inpatient days and \$18,900 in hospital charges, relative to carefully matched controls without this complication (results not significant due to small numbers) [2]. A similar analysis using regression methods on the Thomson Reuters Projected Inpatient Database for FY 2007 generated estimates of 4.5 inpatient days and \$18,405 in hospital charges, respectively.

**1a.4 Citations for Evidence of High Impact cited in 1a.3:** [HCUP State Inpatient Databases \(SID\). Healthcare Cost and Utilization Project \(HCUP\). 2008. Agency for Healthcare Research and Quality, Rockville, MD.](#)

[Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. JAMA 2003;290\(14\):1868-1874.](#)

[Foster D, Young J, Heller S. US national estimates of mortality, length of stay, and costs attributable to inpatient complications of care. Abstract presented at 2009 Annual Research Meeting of AcademyHealth.](#)

**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:**

The Transfusion Reaction indicator is intended to flag cases of major reactions due to transfusions (ABO and Rh).

**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.]**

This is a “sentinel event” indicator, which is described as a count rather than a rate. Accordingly, it is impossible to provide a distribution of scores across measured entities.

**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]**

[\[1\] HCUP State Inpatient Databases \(SID\). Healthcare Cost and Utilization Project \(HCUP\). 2008. Agency for Healthcare Research and Quality, Rockville, MD. Includes approximately 30 million adult discharges for 4,000 hospitals.](#)

[2]Source: Agency for Healthcare Research and Quality (AHRQ), Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2007, and AHRQ Quality Indicators, version 3.1.

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

This is a “sentinel event” indicator, which is described as a count rather than a rate. Accordingly, it is impossible to provide descriptive statistics for performance results across population groups.

**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]**

Source: Agency for Healthcare Research and Quality (AHRQ), Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2007, and AHRQ Quality Indicators, version 3.1.

**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="checkbox"/>
L	M-H	M	Yes <input type="checkbox"/> IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No <input type="checkbox"/>
M-H	L	M-H	Yes <input type="checkbox"/> IF potential benefits to patients clearly outweigh potential harms: otherwise No <input type="checkbox"/>
L-M-H	L-M-H	L	No <input type="checkbox"/>

**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):**

Transfusion reaction is a health outcome measure. This measure captures illness or injury resulting from administration of mismatched blood or blood products, based on ABO or Rh antigens. These events are considered to be almost entirely preventable. For example, the 2011 Update of the NQF Serious Reportable Events in Healthcare includes this specification of “Patient death or serious injury associated with unsafe administration of blood products”: “Unsafe administration includes, but is not limited to hemolytic reactions and administering a) blood or blood products to the wrong patient; b) the wrong type; or c) blood or blood products that have been improperly stored or handled.” Similarly, “Hemolytic transfusion reaction involving administration of blood or blood products having major blood group incompatibilities” is classified as a sentinel event by The Joint Commission. Preoperative evaluation of a patient for blood transfusion includes (1) reviewing previous medical records, (2) conducting a patient or family interview, and (3) reviewing laboratory test results.

American Society of Anesthesiologists Task Force on Perioperative Blood Transfusion and Adjuvant Therapies. Practice guidelines for perioperative blood transfusion and adjuvant therapies: an updated report by the American Society of Anesthesiologists Task Force on Perioperative Blood Transfusion and Adjuvant Therapies. *Anesthesiology*. 2006 Jul;105(1):198-208.

According to one recent review (Janatpour KA, Kalmin ND, Jensen HM, Holland PV. Clinical outcomes of ABO-incompatible RBC transfusions. *Am J Clin Pathol* 2008; 129(2):276-81), “the most frequent error leading to transfusion of ABO-incompatible blood is failure of the final patient identification check at the bedside, leading to transfusion of properly labeled blood to a recipient other than the one intended. In a recent report from Ireland’s hemovigilance system, more than half of all adverse reactions to blood transfusion were caused by the patient being given the wrong blood component. The relative distribution of errors in our cases and survey results are similar to those in other reports, with failures in pretransfusion verification of patient identification comprising a majority of all errors, followed by laboratory errors, and errors in sample collection and labeling... With an increased awareness of the root causes of transfusion errors, hospitals have taken steps to address them, such as requiring 2 pre-transfusion samples to confirm a patient’s initial ABO blood type result (independent of the American Association of Blood Banks standard requiring 2

determinations of the recipient's ABO type if using computer crossmatching). In theory, requiring a second sample to confirm the ABO blood type could significantly reduce ABO-incompatible transfusion because the vast majority of errors are due to sample collection and labeling and bedside errors. A reduction in the use of stationary refrigerators in the operating room is reported to have reduced some transfusion errors... Various devices have also been introduced to minimize errors in sample collection and transfusion to the intended recipient and have prevented some errors. These are summarized in a recent review. However, it is difficult to know whether actual use of these devices is widespread and their effectiveness in preventing ABO-incompatible transfusions.... Quality improvement dictates that analysis of adverse sentinel events such as ABO-incompatible transfusions be performed. When such an event has been identified, corrective measures should be instituted to prevent recurrences."

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

**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):  
Not applicable

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

**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): Not applicable

**1c.7 Consistency of Results across Studies** (Summarize the consistency of the magnitude and direction of the effect): Not applicable

**1c.8 Net Benefit** (Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms):  
Not applicable

**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:** Not applicable

**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: Not applicable

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

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? Yes

S.2 If yes, provide web page URL: [http://qualityindicators.ahrq.gov/modules/psi\\_resources.aspx](http://qualityindicators.ahrq.gov/modules/psi_resources.aspx)

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

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):

Discharges 18 years and older or in MDC 14 with ICD-9-CM codes for transfusion reaction in any secondary diagnosis field of all medical and surgical discharges defined by specific DRGs or MS-DRGs

See Patient Safety Indicators Appendices:

- Appendix B – Medical Discharge DRGs
- Appendix C – Medical Discharge MS-DRGs
- Appendix D – Surgical Discharge DRGs
- Appendix E – Surgical Discharge MS-DRGs

Link to PSI appendices:

<http://qualityindicators.ahrq.gov/Downloads/Software/SAS/V43/TechnicalSpecifications/PSI%20Appendices.pdf>

Exclude cases:

- with principal diagnosis of transfusion reaction or secondary diagnosis present on admission
- with missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing) or principal diagnosis (DX1=missing)

**2a1.2 Numerator Time Window** *(The time period in which the target process, condition, event, or outcome is eligible for inclusion):*  
 User may specify the time window; generally one calendar 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):*

ICD-9-CM Transfusion reaction diagnosis codes:

9996

ABO INCOMPATIBILITY REACTION

99960

ABO INCOMPATIBILITY REACTION NOS

99961

ABO INCOMP/HTR NEC

99962

ABO INCOMPAT/ACUTE HTR

99963

ABO INCOMPAT/DELAY HTR

99969

ABO INCOMPAT REACTN NEC

9997

RH INCOMPATIBILITY REACTION

99970

RH INCOMPAT REACTION NOS

99971

RH INCOMP/HTR NEC

99972

RH INCOMPAT/ACUTE HTR

99973

RH INCOMPAT/DELAY HTR

99974

RH INCOMPAT REACTION NEC

E8760

MISMATCHED BLOOD IN TRANSFUSION

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

Not applicable

**2a1.5 Target Population Category** *(Check all the populations for which the measure is specified and tested if any):* Adult/Elderly Care

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

Not applicable

**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):*

Not applicable

**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:** Count

**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.):

Identify cases meeting the target outcome. Exclude cases meeting the exclusion criteria. Count the number of case at the hospital level.

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

URL

<http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/QI%20Empirical%20Methods%2005-03-11.pdf>

Not applicable

**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.): HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality, Rockville, MD.

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

<http://www.hcup-us.ahrq.gov/sidoverview.jsp>

Not applicable

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

URL

<http://qualityindicators.ahrq.gov/Downloads/Software/WinQI/V43/AHRQ%20QI%20Software%20Instructions,%20WinQI.pdf>

Not applicable

**2a1.33 Level of Analysis** (Check the levels of analysis for which the measure is specified and tested): Facility

**2a1.34-35 Care Setting** (Check all the settings for which the measure is specified and tested): Hospital/Acute Care Facility

**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):

HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2008. Agency for Healthcare Research and Quality, Rockville, MD. Consists of approximately 30 million adult discharges and 4,000 hospitals.

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

Not applicable

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

Because this indicator is expressed as a count, no reliability testing was conducted.

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

**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:**

In FY2011 there was a coding change that improved the specificity of this indicator. Diagnoses now assigned to 999.75-999.79 (non-ABO incompatibility, including incompatibility related to minor antigens such as Duffy, Kell, Kidd, Lewis, M, N, P, or S) were previously assigned to 999.7. These diagnoses are no longer part of the numerator inclusion, because these cases are generally non-preventable reactions to "minor" blood group antigen reactions.

**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):

Due to the extreme rarity of this event, formal validation studies based on randomly sampled data cannot be undertaken. However, user feedback regarding internal review of flagged cases is continuously solicited. This user feedback indicated that most false positive cases were attributable to minor or unspecified antigens rather than the ABO and Rh antigens. To eliminate this problem, we petitioned the ICD-9-CM Coordination and Maintenance Committee to create new codes (effective October 2010) for non-ABO, non-Rh incompatibility reactions and to clarify that the ICD-9-CM codes specified above should only be used for ABO and Rh incompatibility reactions.

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

A structured review of each indicator was undertaken to evaluate the face validity (from a clinical perspective) of the indicators. Specifically, the panels approach sought to establish consensual validity, which "extends face validity from one expert to a panel of experts who examine and rate the appropriateness of each item..." The methodology for the structured review was adapted from the RAND/UCLA Appropriateness Method and consisted of an initial independent assessment of each indicator by clinician panelists using an initial questionnaire, a conference call among all panelists, followed by a final independent assessment by clinician panelists using the same questionnaire. The panel process served to refine definitions of some indicators, add new measures, and dismiss indicators with major concerns from further consideration.

Twenty-one professional clinical organizations were invited to submit nominations. These organizations were selected based on the applicability of the specialty or subspecialty to our quality indicators. Organizations that represented general practitioners (e.g., general surgeons, internists, critical care physicians, perioperative nurses, and critical care nurses) were asked to nominate more panelists than those representing sub-specialties. Fifteen organizations submitted nominations: American Association of Critical-Care Nurses; American Academy of Family Physicians; American College of Cardiology; American College of Nurse-Midwives; American College of Obstetricians and Gynecologists; American College of Physicians/American Society of Internal Medicine;



American College of Radiology; American College of Surgeons; American Geriatric Society; Association of Perioperative Nurses; American Society of Anesthesiologists; American Society of Health-system Pharmacists; American Thoracic Society; Association of Women's Health Obstetric and Neonatal Nurses; and National Association of Inpatient Physicians.

These professional organizations nominated a total of 162 clinicians. Each nominee was invited to participate in the evaluation. In order to be eligible to participate, nominees were required to spend at least 30% of their work time on patient care, including hospitalized patients. Ninety-two nominees accepted this invitation. Five nominees were ineligible to participate. Nominees were asked to provide information regarding their practice characteristics, including specialty and subspecialty and setting (i.e., urban vs. rural location, region of country, and service to underserved populations), information regarding primary hospital of practice (i.e., funding source) and personal information (i.e., clinical education history, academic affiliation).

For assignments to each panel, a list of applicable specialties was identified for the indicators to be evaluated by a given panel. Panelists were selected so that each panel had diverse membership in terms of practice characteristics and setting. Thus, when a specific area was over-represented by the pool of eligible nominees, randomly drawn members from that specific sub-group were contacted first to fill the panels. In addition, conference call scheduling logistics influenced assignments. Fifty-seven of the eligible panelists accepted the invitation to participate on specific panels. Four did not participate in the conference call, and thus were removed from the panels. All other panelists (53) completed the evaluation in full.

Eight panels were formed. Complications of medical care indicators were examined by two panels. Surgical complications indicators were reviewed by three panels. Another panel assessed indicators related to procedural complications. Finally, two panels examined obstetric complications indicators. All panels had diversity in the geographic location of panelists, and the type of practice.

**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 was systematically assessed using an expert panel process, as described in our original submission documents (McDonald KM, Romano PS, Geppert J, Davies SM, Duncan BW, Shojania KG. Measures of Patient Safety Based on Hospital Administrative Data: The Patient Safety Indicators. Technical Review Number 5. Rockville, MD: Agency for Healthcare Research and Quality, 2002). A Multi-specialty Panel and a Surgical Panel both rated the indicator as acceptable on overall usefulness as an indicator of potentially preventable complications of care. Specifically, a multi-specialty expert panel gave this indicator an overall usefulness rating of 8 (on a scale of 1-9) with agreement, and a preventability rating of 7 with disagreement. A surgical expert panel gave this indicator a median usefulness rating of 7.8 (on a scale of 1-9) with agreement, and a median preventability rating of 8 with indeterminate agreement. The major reason for disagreement was that some panelists expressed concern about the fact that the original numerator specification for PSI 16 may have included reactions caused by minor antigens in patients with complex hematologic histories who had been sensitized by multiple prior transfusions. These complications were seen as less preventable than Rh or ABO incompatibility reactions, and clinically different. This concern was resolved by petitioning the ICD-9-CM Coordination and Maintenance Committee for appropriate coding changes, as described above.

**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):

HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2008. Agency for Healthcare Research and Quality, Rockville, MD. Consists of approximately 30 million adult discharges and 4,000 hospitals.

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

Exclude cases with principal diagnosis of transfusion reaction or secondary diagnosis present on admission.

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

Of 153 cases with a secondary diagnosis of transfusion reaction, 89 cases were present on admission.

**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):

HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2008. Agency for Healthcare Research and Quality, Rockville, MD. Consists of approximately 30 million adult discharges and 4,000 hospitals.

**2b5.2 Analytic Method** (Describe methods and rationale to identify statistically significant and practically/meaningfully differences in performance):

Sum the number of cases identified with the outcome of interest by hospital

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

The measure is a serious reportable event so differences in performance are not applicable

**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:**

URL

[http://qualityindicators.ahrq.gov/Downloads/Modules\\_Non\\_Software/Modules%20Development%20Bullet/psi\\_development.zip](http://qualityindicators.ahrq.gov/Downloads/Modules_Non_Software/Modules%20Development%20Bullet/psi_development.zip)  
 Not applicable

**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 Purpose/ Use** (Check all the purposes and/or uses for which the measure is intended): [Public Reporting](#), [Quality Improvement](#) (Internal to the specific organization)

**3.1 Current Use** (Check all that apply; for any that are checked, provide the specific program information in the following questions): [Public Reporting](#), [Quality Improvement](#) (Internal to the specific organization)

**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 is used for public reporting in 9 realms.

Arizona (NY QIO)

Why Not the Best?

<http://www.whynotthebest.org/>

Illinois (state hospital association)

Illinois Hospitals Caring for You

[www.illinoishospitals.org](http://www.illinoishospitals.org)

Iowa (Iowa Healthcare Collaborative)

Iowa Healthcare Collaborative

<http://www.ihconline.org/asp/publicreporting/iowareport.aspx>

Kentucky (Norton Healthcare, a hospital system)

Norton Healthcare Quality Report

<http://www.nortonhealthcare.com/body.cfm?id=157>

Kentucky (state hospital association)

Kentucky Hospital Association Quality Data

<http://info.kyha.com/QualityData/IQISite/>

Louisiana (state)

Louisiana Health Finder

<http://www.healthfinderla.gov/default.aspx>

New Jersey (state)

Find and Compare Quality Care in NJ Hospitals  
<http://www.nj.gov/health/healthcarequality/>

New York (health care coalition)  
 New York State Hospital Report Card  
<http://www.myhealthfinder.com/>

**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: A research team from the School of Public Affairs, Baruch College, under contracts with the Department of Public Health, Weill Medical College and Battelle, Inc., has developed a pair of Hospital Quality Model Reports at the request of the Agency for Healthcare Research & Quality (AHRQ). These reports are designed specifically to report comparative information on hospital performance based on the AHRQ Quality Indicators (QIs). The work was done in close collaboration with AHRQ staff and the AHRQ Quality Indicators team. The Model Reports (discussed immediately above) are based on:

- Extensive search and analysis of the literature on hospital quality measurement and reporting, as well as public reporting on health care quality more broadly;
- Interviews with quality measurement and reporting experts, purchasers, staff of purchasing coalitions, and executives of integrated health care delivery systems who are responsible for quality in their facilities;
- Two focus groups with chief medical officers of hospitals and/or systems and two focus groups with quality managers from a broad mix of hospitals;
- Four focus groups with members of the public who had recently experienced a hospital admission; and
- Four rounds of cognitive interviews (a total of 62 interviews) to test draft versions of the two Model Reports with members of the public with recent hospital experience, basic computer literacy but widely varying levels of education

**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].**

The Patient Safety Indicators (PSIs) are a set of indicators providing information on potential in hospital complications and adverse events following surgeries, procedures, and childbirth. The PSIs were developed after a comprehensive literature review, analysis of ICD-9-CM codes, review by a clinician panel, implementation of risk adjustment, and empirical analyses.

The PSIs can be used to help hospitals identify potential adverse events that might need further study; provide the opportunity to assess the incidence of adverse events and in hospital complications using administrative data found in the typical discharge record; include indicators for complications occurring in hospital that may represent patient safety events; and, indicators also have area level analogs designed to detect patient safety events on a regional level.

[http://qualityindicators.ahrq.gov/modules/psi\\_overview.aspx](http://qualityindicators.ahrq.gov/modules/psi_overview.aspx)

The following are several entities that use the measure in quality improvement:

1) University Healthcare Consortium (UHC)

UHC is an alliance of 103 academic medical centers and 219 of their affiliated hospitals. UHC reports this and other AHRQ QIs to their member hospitals for their internal quality improvement purposes.

2) Ministry

Ministry is a 14 hospital system in WI, which includes the Marshfield Clinic in its system.

3) Premier

Premier uses the measure in their "QUEST" tool, which is used by hundreds of hospitals in their quality assurance and improvement work.

**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 AHRQ QI support line receives approximately 150 user queries per month and almost 50 user per month download the AHRQ QI PSI software. Users have used the PSI since the release in 2003.

Users can identify cases with the serious reportable event and conduct root cause analysis

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:**

Coding professionals follow detail guidelines, are subject to training and credentialing requirements, peer review and audit

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

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

**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 AHRQ QI software has been publicly available at no cost since 2001; Users have over ten years of experience using the AHRQ QI software in SAS and Windows.

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

NQF #0349 Transfusion Reaction (PSI 16)

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):**

**CONTACT INFORMATION**

**Co.1 Measure Steward (Intellectual Property Owner):** [Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850](#)

**Co.2 Point of Contact:** [John, Bott, Contractor, AHRQ Quality Indicators Measure Expert Center for Delivery, Organization and Markets, John.Bott@ahrq.hhs.gov, 301-427-1317-](#)

**Co.3 Measure Developer if different from Measure Steward:** [Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850](#)

**Co.4 Point of Contact:** [John, Bott, Contractor, AHRQ Quality Indicators Measure Expert Center for Delivery, Organization and Markets, John.Bott@ahrq.hhs.gov, 301-427-1317-](#)

**Co.5 Submitter:** [John, Bott, Contractor, AHRQ Quality Indicators Measure Expert Center for Delivery, Organization and Markets, John.Bott@ahrq.hhs.gov, 301-427-1317-, Agency for Healthcare Research and Quality](#)

**Co.6 Additional organizations that sponsored/participated in measure development:**  
[Battelle Memorial Institute,](#)  
[University of California-Davis,](#)  
[Stanford University](#)

**Co.7 Public Contact:** [John, Bott, Contractor, AHRQ Quality Indicators Measure Expert Center for Delivery, Organization and Markets, John.Bott@ahrq.hhs.gov, 301-427-1317-, Agency for Healthcare Research and Quality](#)

**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.**

[Multi-specialty Panel and Surgical Panel members are listed in the technical report:](#)

[http://qualityindicators.ahrq.gov/Downloads/Modules\\_Non\\_Software/Modules%20Development%20Bullet/psi\\_development.zip](http://qualityindicators.ahrq.gov/Downloads/Modules_Non_Software/Modules%20Development%20Bullet/psi_development.zip)

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

<p><b>adapting the original measure and any work with the original measure steward:</b> This indicator was originally proposed by lezzoni et al. as part of the Complications Screening Program (CSP “sentinel events”)</p>
<p><b>Measure Developer/Steward Updates and Ongoing Maintenance</b>  <b>Ad.3 Year the measure was first released:</b> 2003  <b>Ad.4 Month and Year of most recent revision:</b> 08, 2011  <b>Ad.5 What is your frequency for review/update of this measure?</b> Annual  <b>Ad.6 When is the next scheduled review/update for this measure?</b> 12, 2011</p>
<p><b>Ad.7 Copyright statement:</b> Not applicable</p>
<p><b>Ad.8 Disclaimers:</b> Not applicable</p>
<p><b>Ad.9 Additional Information/Comments:</b> Not applicable</p>
<p><b>Date of Submission (MM/DD/YY):</b> 09/14/2011</p>