NQF #0363 Foreign Body Left During Procedure (PSI 5)

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 <u>submitting standards web page</u>.

NQF #: 0363 NQF Project: Patient Safety Measures-Complications Project

(for Endorsement Maintenance Review)

Original Endorsement Date: Jun 24, 2010 Most Recent Endorsement Date: Jun 24, 2010

BRIEF MEASURE INFORMATION

De.1 Measure Title: Foreign Body Left During Procedure (PSI 5)

Co.1.1 Measure Steward: Agency for Healthcare Research and Quality

De.2 Brief Description of Measure: Count of discharges with foreign body left in during procedure in medical and surgical discharges among patients 18 years and older or MDC 14 (pregnancy, childbirth, and puerperium)

2a1.1 Numerator Statement: Discharges, 18 years and older or MDC 14 (pregnancy, childbirth, and puerperium), with ICD-9-CM codes for foreign body left in during procedure in any secondary diagnosis field of medical and surgical discharges defined by specific DRGs or MS-DRGs. (Details of medical and surgical discharges defined by specific DRGs or MS-DRGs and exclusions noted in 2a1.3).

2a1.4 Denominator Statement: Not applicable

2a1.8 Denominator Exclusions: Not applicable

1.1 Measure Type: Outcome 2a1. 25-26 Data Source: Administrative claims 2a1.33 Level of Analysis: Facility

1.2-1.4 Is this measure paired with another measure? No

De.3 If included in a composite, please identify the composite mea	asure (title and NQF number if endorsed)
Not applicable	

STAFF NOTES (issues or questions regarding any criteria)

Comments on Conditions for Consideration:

Is the measure untested?	Yes No	If untested, explain how it meets	criteria for consideration for time-limited
endorsement:			

1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (*check De.5*):
5. Similar/related <u>endorsed</u> or submitted measures (*check 5.1*):

Other Criteria:

Staff Reviewer Name(s):

1. IMPACT, OPPORTUITY, 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 <u>guidance on evidence</u>.

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

NQF #0363 Foreign Body Left During Procedure (PSI 5) 1a. High Impact: (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): HealthGrades noted significant occurrence, cost and death with the occurrence of a patient safety event (http://www.healthgrades.com/business/img/HealthGradesPatientSafetyInAmericanHospitalsStudy2011.pdf). An earlier study that used multivariable matching on the 2000 Nationwide Inpatient Sample database (Healthcare Cost and Utilization Project) attributed significant excess length of stay, charges and mortality to this QI. {Zhan, 2003} Cases that were flagged by this PSI had 2.1% excess mortality, 2.1 days of excess hospitalization, and \$13,315 in excess hospital charges, relative to carefully matched controls that were not flagged. A similar analysis using regression methods on the Thomson Reuters Projected Inpatient Database for FY 2007 generated estimates of 4.5 inpatient days and \$13,202 in hospital charges, respectively. Although the four year trends for this QI were consistent compared to other PSI, PSI 05 was one of the most rare QI in a sample of VA data. {Rosen, 2006} About 184 of these events are estimated to have occurred in US community hospitals in 2008. This QI was also included as part of an international consortium on the conversion of the PSI to ICD-10.{Quan, 2008} International data from the Organization for Economic Cooperation and Development show substantial variation across countries, with 6-fold variation between the countries with the lowest and highest rates (Denmark and Switzerland, respectively). 1a.4 Citations for Evidence of High Impact cited in 1a.3: Bottle A, Aylin P. Application of AHRQ patient safety indicators to English hospital data. Qual Saf Health Care 2009;18(4):303-308. Drösler SE, Romano PS, Tancredi DJ, Klazinga NS. International comparability of Patient Safety Indicators in 15 OECD member countries: A methodological approach of adjustment by secondary diagnoses. Health Serv Res 2011; Jul 15. [Epub ahead of print]. Rosen AK, Rivard P, Zhao S et al. Evaluating the patient safety indicators: how well do they perform on Veterans Health Administration data? Med Care 2005;43(9):873-884. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. JAMA 2003;290(14):1868-1874. 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: This indicator is intended to flag cases of a foreign body accidentally left in a patient during a procedure. **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 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 <u>Maintenance</u> – 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 out	tcome OR meets tl	he 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	
L	M-H	М	Yes IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No	
M-H	L	M-H	Yes IF potential benefits to patients clearly outweigh potential harms: otherwise No	
L-M-H	L-M-H	L	No 🗌	
Health ou	tcome – ra	tionale support	s relationship to at least	Does the measure pass subcriterion1c?

one healthcare structure, process, intervention, or service

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):

Foreign body left during procedure is a health outcome measure. This measure captures illness or injury resulting from a foreign body that was left during a procedure, when the retention of that foreign body was not part of the originally planned procedure. 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 "Unintended retention of a foreign object in a patient after surgery or other invasive procedure": "includes medical or surgical items intentionally placed by provider(s) that are unintentionally left in place. Excludes (a) objects present prior to surgery or other invasive procedure that are intentionally left in place; (b) objects intentionally implanted as part of a planned intervention; and (c) objects not present prior to surgery/procedure that are intentionally left in when the risk of removal exceeds the risk of retention (such as microneedles, broken screws)." Similarly, "Unintended retention of a foreign object in an individual after surgery or other procedure" is classified as a reviewable sentinel event by The Joint Commission. It is also classified as a "hospital-acquired condition" by the Center for Medicare & Medicaid Services, based on a similar ICD-9-CM specification as PSI 05.

Retained surgical items have become a major focus of safety improvement effort within health care organizations. Current thinking is summarized as follows by The Joint Commission:

"Retained foreign objects have the potential to cause a multitude of problems including post-procedure infections, bowel perforation, abscess, undue pain, return to surgery, and even death.

"The consequences extend beyond clinical complications and often include additional financial burdens such as extended lost time from work, additional expenses related to frequent follow up visits, and additional medications. Such complications also present a significant financial risk to health care providers. Effective October 1, 2008, CMS identified a list of hospital-acquired conditions deemed to be "preventable" and may deny payment for complications resulting from such events or conditions: retained foreign objects after surgery is one of those conditions. Other third party payors have implemented similar no-pay policies.

- "Organizations should take steps to mitigate the occurrence of retained objects. Such steps may include the following:
- · Auditing of operative/procedural records to ensure required counts are completed and documented

• Reviewing policies and procedures to ensure they are consistent with current practice and ensuring current practice is consistent with policies and procedures

- · Conducting random "real time" observations of staff to monitor compliance with safe practice
- · Ensure that initial and ongoing competencies are assessed regularly
- Consider incorporating the use of technology into the process
- · Use "near miss" incidents as learning experiences for staff"

Another widely cited clinical practice guideline relevant to PSI 05 was published by the Institute for Clinical Systems Improvement: "Accurately accounting for all items that could potentially become unintentionally retained is a priority of the entire surgical team, though the primary responsibility for performing the count process belongs to the circulator and scrub. There must be no distractions (e.g., extraneous conversation, music, unnecessary interruptions). The circulator must be a registered nurse [R]. Radiographic imaging is not a substitute for performing accurate count procedures. Count procedures may be omitted or modified in an extreme patient emergency. This exception will be documented in the patient's medical record and when the patient's condition allows, radiographic imaging should be obtained to rule out the possibility of an unintentionally retained foreign object." Perioperative protocol. Health care protocol. 2009 Sep (revised 2010 Oct). NGC:008119 Institute for Clinical Systems Improvement See also recent publications and guidelines from the "Nothing Left Behind" project (http://nothingleftbehind.org/) and the Association of Perioperative Registered Nurses, which offers "Recommended Practices for Prevention of Retained Surgical Items" (http://www.aorn.org/News/August2010News/RSI).

Our literature review identified four empirical studies supporting the relationship of this health outcome to processes of care. In the medical field, one research team used a computer-assisted material testing machine to study the mechanical properties of three different types of epidural catheters: polyurethane, radiopaque and clear nylon. In this blinded study, polyurethane was the least prone to breakage, whereas radiopaque catheters had the highest stiffness (Ates, Yucesoy et al. 2000).

At least two case-control studies have evaluated risk factors associated with retained objects after major surgery. The first study involved retrospective review of medical records flagged using the PSI 5 numerator code. In the multivariate analysis, factors associated with a significantly higher risk of retained foreign bodies were the total number of major procedures performed [OR 1.6; 95% CI 1.1-2.3, P=0.008] and an incorrect count [OR 16.2; 95% CI 1.3-197.8; P=0.02], highlighting the importance of ending surgery with correct instrument and sponge counts (Lincourt, Harrell et al. 2007). In the second study, cases were drawn from a large malpractice insurer representing one third of the physicians in Massachusetts from 1985 to 2001. Emergency surgery [RR 8.8; 95% CI 2.4-31.9], unplanned change in the operation [RR 4.1; 95% CI 1.4-12.4], and body-mass index [RR for each one unit interval 1.1; 95% CI 1.0-1.2] were among the significant risk factors for retained objects (Gawande, Studdert et al. 2003). Finally, in a review of lawsuits claiming retained objects during surgery, 10% involved procedures during which no sponge count was performed. In the case of non-vaginal surgery, a falsely correct sponge count was documented in 76% of the claims (Kaiser, Friedman et al. 1996). Falsely correct sponge counts were attributed to team fatigue, difficult operations, sponges "sticking together," or a poor counting system. Incorrect sponge counts that were accepted prior to closure resulted from either surgeons' dismissing the incorrect count without re-exploring the wound, or nursing staff allowing an incorrect count to be accepted. Ates, Y., C. A. Yucesoy, et al. (2000). "The mechanical properties of intact and traumatized epidural catheters." Anesth Analg 90(2): 393-9.

Gawande, A. A., D. M. Studdert, et al. (2003). "Risk factors for retained instruments and sponges after surgery." N Engl J Med 348(3): 229-35.

Kaiser, C. W., S. Friedman, et al. (1996). "The retained surgical sponge." Ann Surg 224(1): 79-84. Lincourt, A. E., A. Harrell, et al. (2007). "Retained foreign bodies after surgery." J Surg Res 138(2): 170-4.

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

Clinical Practice Guideline, 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): 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

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=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 <u>developer's assessment</u> of the quantity, quality, and consistency of the body of evidence? 1c.25 Quantity: Moderate 1c.26 Quality: Moderate1c.27 Consistency: Moderate
Was the threshold criterion, <i>Importance to Measure and Report</i> , 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, <u>as specified</u>, 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 <u>guidance on measure testing</u>.

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.ahrg.gov/modules/psi resources.aspx 2a. RELIABILITY. Precise Specifications and Reliability Testing: H M L 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 MDC 14 (pregnancy, childbirth, and puerperium), with ICD-9-CM codes for foreign body left in during procedure in any secondary diagnosis field of medical and surgical discharges defined by specific DRGs or MS-DRGs. (Details of medical and surgical discharges defined by specific DRGs or MS-DRGs and exclusions noted in 2a1.3). 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 Foreign body left in during procedure diagnosis codes: 9984 FOREIGN BODY ACCIDENTALLY LEFT DURING A PROCEDURE 9987 ACUTE REACTIONS TO FOREIGN SUBSTANCE ACCIDENTALLY LEFT DURING A PROCEDURE Foreign body left in during: E8710 SURGICAL OPERATION E8711 INFUSION OR TRANSFUSION E8712 KIDNEY DIALYSIS OR OTHER PERFUSION E8713 INJECTION OR VACCINATION E8714 ENDOSCOPIC EXAMINATION E8715 ASPIRATION OF FLUID OR TISSUE, PUNCTURE, AND CATHETERIZATION E8716 HEART CATHETERIZATION E8717 **REMOVAL OF CATHETER OR PACKING** E8718 OTHER SPECIFIED PROCEDURES E8719 UNSPECIFIED PROCEDURE 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

Numerator exclusions include:

- cases with the outcome of interest noted as present on admission

- cases with the following missing variables: gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing) or principal diagnosis (DX1=missing)

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

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:

Similar to the NQF Serious Reportable Event for "Unintended retention of a foreign object in a patient after surgery or other invasive procedure", PSI 05 is intended to capture "occurrences of unintended retention of objects at any point after the surgery/procedure ends regardless of setting (post anesthesia recovery unit, surgical suite, emergency department, patient bedside) and regardless of whether the object is to be removed after discovery" and "unintentionally retained objects (including such things as wound packing material, sponges, catheter tips, trocars, guide wires) in all applicable settings." The Editorial Advisory Board for Coding Clinics for ICD-9-CM reinforced this alignment in 2009 (26(1):18) by instructing the coding community to follow current NQF guidance related to when surgery ends in assigning the ICD-9-CM code 998.4 (Foreign body accidentally left during a procedure). Specifically,

according to the 2011 update, "Surgery ends after all incisions or procedural access routes have been closed in their entirety, device(s) such as probes or instruments have been removed, and, if relevant, final surgical counts confirming accuracy of counts and resolving any discrepancies have concluded and the patient has been taken from the operating/procedure room." The coding community may not yet be aware of NQF's 2011 change to the definition of when surgery ends, because the previous definition permitted a wider range of interpretation ("Surgery ends after counts have concluded, the surgical incision has been closed, and/or operative device(s) such as probes have been removed, regardless of setting"), but educational efforts are now underway. However, there is one substantive difference between the PSI 05 measure specification and the corresponding NQF SRE specification. The NQF SRE specification excludes "objects not present prior to surgery/procedure that are intentionally left in when the risk of removal exceeds the risk of retention (such as microneedles, broken screws)." ICD-9-CM does not recognize this exclusion, based on the principle that retention of the foreign body was not part of the originally planned procedure (and was not the original intent of the surgeon), and therefore it represents a complication of that procedure, even if the surgeon later decides that the risk of removal exceeds the risk of retention. See this guestion and answer in Coding Clinics for ICD-9-CM (1st guarter 2011): "During surgery a needle was placed along the right lateral aspect of the ring of the aortic valve and passed through tissue deep in the recess between the aortic root and the left atrium to try to cinch the valve down. The suture broke from the needle and the needle was lost within this tissue. Multiple attempts to find it were unsuccessful. An x-ray was obtained which did not reveal the needle. The chest was closed and a second x-ray showed that the needle was where the surgeon thought, to the right of the aortic valve. The chest was reopened but the needle still could not be located or palpated. The surgeon decided that continuing to look for the needle was likely to cause the patient more harm than good. Therefore, the chest was reapproximated...

Assign code 998.4, Foreign body accidentally left during a procedure. Although the surgeon made the decision to leave the needle because continuing to search for it might cause harm to the patient, it was not the intent of the original procedure to leave a foreign body behind."

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):

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

Only one published study (Chen Q, Rosen AK, Cevasco M, Shin M, Itani KM, Borzecki AM. Detecting patient safety indicators: How valid is "foreign body left during procedure" in the Veterans Health Administration? J Am Coll Surg 2011; 212(6):977-83) has reported the criterion validity of PSI 05 in the US, based on clinical data abstracted from the Veterans Health Administration (VA) electronic medical record (EMR) as the gold standard. A smaller study from England reported the criterion validity of an ICD-10 version of PSI 05 using similar methods with administrative and clinical data from 18 regional trusts in 2007 (Bottle A, Aylin P. How NHS trusts could use patient safety indicators to help improve care. Health Care Risk Rep 2008; 12-14).

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 a median usefulness rating of 8 (on a scale of 1-9) with agreement, and a preventability rating of 8 with agreement. A surgical expert panel gave this indicator a median usefulness rating of 7 (on a scale of 1-9) with indeterminate agreement, and a median preventability rating of 7.5 with agreement. Panelists noted that sometimes removal of a foreign body may cause more damage than leaving it in, in which case the foreign body may represent a device malfunction but not a medical error. Panelists felt that this indicator should be stratified or risk adjusted for the type of procedure, although AHRQ determined that it is not technically possible to do so. Panelists were also concerned about the fact that coding of this indicator requires the physician to note that the foreign body was accidentally left in, which may be inconsistently documented. Finally, panelists noted that some foreign bodies left in do not cause substantial morbidity, and thus may not represent clinically significant complications.

In the criterion validation study by Chen et al., 93 cases were flagged by PSI 05 and 42 were true positives, yielding a PPV of 45% (95% CI 35% to 56%). False positives were due to foreign bodies present on admission (57%) or coding errors (43%). The former cases would automatically be excluded by users with "present on admission" (POA) data. Adjusting for the availability of POA data, the estimated PPV would be 66% in the VA. Similarly, estimated PPV in the English study based on ICD-10 codes was 52% (15/29), or 58% after adjusting for the availability of POA data.

True foreign bodies in the VA study were associated with both surgical (n=23) and medical (n=19) procedures. The most common type of surgical foreign body was a sponge (52%). Overall, approximately 40% of foreign bodies were related to a device failure or malfunction (30% of surgical vs 53% of medical foreign bodies). Postoperative complications included pain (24%), infection (12%), adhesions (5%), and bowel obstruction (5%).

Due to the extreme rarity of this event, no studies have been able to assess the sensitivity of PSI 05. However, user feedback regarding internal review of flagged cases is continuously solicited, and has not revealed major concerns about false negative cases.

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): Exclude:

- patients with ICD-9-CM codes for foreign body left in during procedure in the principal diagnosis field or secondary diagnosis present on admission.

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

NQF #0363 Foreign Body Left During Procedure (PSI 5)

Count of cases with principal diagnosis field or secondary diagnosis present on admission.

2b3.3 Results (Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses): Of 1378 cases identified with the outcome of interest, 534 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 guantitative 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):

Not applicable

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):

The measure is a serious reportable event so differences in peformance 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 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)*). <u>If not publicly reported in a national or community program</u>, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [For <u>Maintenance</u> – 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 was used for public reporting in 8 Initatives.

Illinois (state hospital association) Illinois Hospitals Caring for You www.illinoishospitals.org

Iowa (Iowa Healthcare Collaborative) Iowa Healthcare Collaborative http://www.ihconline.org/aspx/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 Hampshire (NY QIO) New York State Health Accountability Foundation http://nyshaf.org/juice/IPROSpikeChart.html

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. <u>If usefulness was demonstrated</u> (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.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):

[For <u>Maintenance</u> – 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://gualityindicators.ahrg.gov/modules/psi_overview.aspx

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

1)Norton Healthcare

A multi-hospital system in Kentucky that uses a number of measures in both quality improvement and publicly reporting.

2) Minnesota Hospital Association

3) Ministry

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

4) 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., *Ql 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:

	SHITARII ITY	ORSEMEN
UVLNALL	JULIADILL	

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

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 <u>NQF-endorsed measure(s)</u>: 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: University of California-Davis Stanford University Battelle Memorial Institute

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

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: This indicator was originally proposed by lezzoni et al. as part of the

Complications Screening Program (CSP "sentinel events")

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.3 Year the measure was first released: 2003

Ad.4 Month and Year of most recent revision: 08, 2011

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

Ad.6 When is the next scheduled review/update for this measure? 12, 2011

Ad.7 Copyright statement: Not applicable

Ad.8 Disclaimers: Not applicable

Ad.9 Additional Information/Comments: Not applicable

Date of Submission (MM/DD/YY): 09/14/2011