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.

NQF #: 0347 NQF Project: Patient Safety Measures-Complications Project

(for Endorsement Maintenance Review)
Original Endorsement Date: May 15, 2008 Most Recent Endorsement Date: May 15, 2008 Last Updated Date: Jul 12, 2012

BRIEF MEASURE INFORMATION

De.1 Measure Title: Death Rate in Low-Mortality Diagnosis Related Groups (PSI 2)

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

De.2 Brief Description of Measure: Percent of discharges with disposition of "deceased" (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator

2a1.1 Numerator Statement: Discharges with disposition of "deceased" (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator

2a1.4 Denominator Statement: Discharges, 18 years and older or MDC 14 (pregnancy, childbirth, and puerperium), in DRGs or MS-DRGs with less than 0.5% mortality rate. If a DRG is divided into two groups with or without "comorbidities or complications" or an MS-DRG is divided into three groups - with major, other, or no comorbidities or complications - then both DRGs or all MS-DRGs must have mortality rates below 0.5% to qualify for inclusion.

2a1.8 Denominator Exclusions: Exclude cases:
- with any code for trauma, cancer, or immunocompromised state
- transfer to an acute care facility (DISP = 2)
- with missing discharge disposition (DISP=missing), gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing) or principal diagnosis (DX1=missing)

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 measure (*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 endorsed or submitted measures (*check 5.1*):
Other Criteria:

Staff Reviewer Name(s):

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable
Created on: 07/17/2012 at 02:47 PM

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 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)*:
The utility of examining deaths in low-mortality DRGs was recognized in early efforts to develop healthcare quality outcomes. Based on two-stage implicit review of 8,109 randomly selected records from 104 New York hospitals in 1985-86, Hannan et al. found that patients in low-mortality DRGs (<0.5%) were 5.2 times more likely than non-targeted patients (9.8% versus 1.7%) to have received "care that departed from professionally recognized standards," after adjusting for patient demographic, geographic, and hospital characteristics. A total of 301 inpatient deaths were captured by this indicator, or an average of 3 per hospital during the study period.

Based on the Nationwide Inpatient Sample from AHRQ's Healthcare Cost and Utilization Project, there were 3,524 (±62) cases of PSI 2 in the United States in 2009, representing 0.046% of eligible hospitalizations.

Rosen and colleagues examined the incidence of PSI 2 using Veterans Health Administration (VA) data, and the association between this indicator and hospital length of stay and costs. Across the 127 acute care hospitals in the VA system, there were 178, 139, 208, and 191 deaths flagged by this indicator in FY 2001, 2002, 2003, and 2004, respectively. Hospitalizations flagged by this indicator had significantly longer median length of stay (7 versus 3 days, difference=4 days) and significantly higher median estimated cost ($12,005 versus $4,409, difference=$7,595) than unflagged hospitalizations in the same low-mortality DRGs.

1a.4 Citations for Evidence of High Impact cited in 1a.3: (1) Hannan EL, Bernard HR, O´Donnell JF, Kilburn H, Jr. A methodology for targeting hospital cases for quality of care record reviews. Am J Public Health 1989;79(4):430-436.
(2) Romano PS, Geppert JJ, Davies S, et al. A national profile of patient safety in US hospitals. Health Aff (Millwood). 2003;22:154-66.
(3) Rosen AK, Rivard P, Zhao S, et al. Evaluating the patient safety indicators (PSIs): how well do they perform on VA data? Med Care. 2005;43:873-84.
(4) Rosen AK, Zhao S, Rivard P, et al. Tracking rates of Patient Safety Indicators over time: Lessons from the VA. Med Care. 2006;44(9):850 61.

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 identify in-hospital deaths among patients unlikely to die during their hospitalizations. The underlying assumption is that when patients admitted for an extremely low-mortality condition or procedure die, a health care error is more likely to be responsible. AHRQ views this measure as complementary to "death among surgical inpatients with serious treatable complications" (NQF-endorsed PSI 4, NQF #0351), which also focuses on a subset of deaths that are more likely to reflect challenges and opportunities in the area of patient safety. In comparison with all-patient Hospital Standardized Mortality Ratios (HSMRs), PSI 2 excludes patients with elevated risk of non-preventable mortality, such as patients who experienced trauma or who

have an immunodeficiency condition or cancer. (As a result, the few remaining patients in low-mortality MS-DRGs such as "total mastectomy" actually have relatively benign conditions such as ductal carcinoma-in-situ.)

Based on two-stage implicit review of 8,109 randomly selected records from 104 New York hospitals in 1985-86, Hannan et al. found that patients in low-mortality DRGs (<0.5%) were 5.2 (95% CI, 3.2-8.4) times more likely than non-targeted cases (9.8% versus 1.7%) to have received "care that departed from professionally recognized standards," after adjusting for patient demographic, geographic, and hospital characteristics. In 15 of these 26 cases (58%) of substandard care, the patient's death was attributed to that care. The association with substandard care was stronger for the DRG-based definition of this indicator than for an alternative definition based on primary surgical procedures with similarly low mortality (9.8% versus 5.7%). Other outcome metrics, such as death within 1 day or 2 days of a surgical procedure, fluid or electrolyte imbalance complicating a surgical case, and cardiac and urinary complications of surgery, also had weaker associations with substandard care than death in low-mortality DRGs (e.g., 2.2% to 7.1%, odds ratios 1.1-3.2).

Hannan's study validated the concept that deaths in low-mortality DRGs represent an opportunity to identify safety-related problems and to intervene to improve patient outcomes. Mihrshani et al. (Intern Med J 2010;40:250-7) reviewed all of the published literature on this indicator, including Hannan's paper and two other studies "that provide direct evidence that the quality of care gap is higher" in deaths flagged by PSI 2 than in other patients. They conclude that "the indicator has utility as a screening tool to enable institutions to quickly and easily identify a manageable number of medical records to investigate more fully, for example, by using chart reviews or a mortality review" but they emphasize the need for "robust analytic research" to understand better the indicator's current performance.

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.*]
In regard to figures below:
1st figure: estimate per 1,000, risk adjusted rates
2nd figure: standard error
3rd figure: p value relative to marked group (marked group = "c")
4th figure: p value: current year relative to prior year

Key:
"c": Reference for p-value test statistics
"*": Data do not meet criteria for statistical reliability, data quality, or confidentiality

Hospital characteristic:
Location of inpatient treatment:
Northeast c 0.409 0.031 0.034
Midwest 0.351 0.028 0.169 0.411
South 0.451 0.024 0.289 0.316
West 0.453 0.034 0.336 0.204

Ownership/control:
Private, not-for-profit c 0.401 0.016 0.495
Private, for-profit 0.446 0.039 0.288 0.855
Public 0.471 0.040 0.109 0.909

Teaching status:
Teaching 0.402 0.025 0.436 0.052
Nonteaching c 0.425 0.018 0.288

Location of hospital (NCHS):
Large central metropolitan 0.387 0.023 0.097 0.237

Large fringe metropolitan c 0.454 0.033 0.040
Medium metropolitan 0.454 0.034 0.991 0.781
Small metropolitan 0.421 0.042 0.539 0.582
Micropolitan 0.402 0.044 0.346 0.992
Not metropolitan or micropolitan 0.382 0.079 0.403 0.631

Bed size of hospital:
Less than 100 0.410 0.041 0.945 0.739
100 - 299 c 0.413 0.022 0.677
300 - 499 0.426 0.027 0.718 0.850
500 or more 0.416 0.034 0.938 0.142

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*]
In regard to figures below:
1st figure: estimate per 1,000, risk adjusted rates
2nd figure: standard error
3rd figure: p value relative to marked group (marked group = “c”)
4th figure: p value: current year relative to prior year

Key:
"c": Reference for p-value test statistics
"*": Data do not meet criteria for statistical reliability, data quality, or confidentiality

Patient characteristic:
Age groups for conditions affecting any age
18-44 c 0.142 0.012 0.917
45-64 0.448 0.037 0.000 0.336
65 and over 1.928 0.093 0.000 0.397

Age groups for conditions affecting primarily elderly
65-69 c 0.645 0.103 0.072
70-74 1.030 0.147 0.031 0.263
75-79 1.670 0.180 0.000 0.844
80-84 2.479 0.244 0.000 0.821
85 and over 4.506 0.343 0.000 0.892

Gender:
Male c 0.593 0.037 0.241
Female 0.357 0.016 0.000 0.506

Median income of patient´s ZIP code:
First quartile (lowest income) 0.474 0.028 0.018 0.744
Second quartile 0.406 0.029 0.492 0.124
Third quartile 0.400 0.029 0.587 0.844
Fourth quartile (highest income) c 0.378 0.029 0.393

Location of patient residence (NCHS):
Large central metropolitan 0.386 0.026 0.090 0.377
Large fringe metropolitan c 0.454 0.030 0.118
Medium metropolitan 0.442 0.035 0.795 0.716
Small metropolitan 0.363 0.043 0.084 0.465
Micropolitan 0.443 0.046 0.841 0.570
Not metropolitan or micropolitan 0.413 0.052 0.501 0.167

Expected payment source:
Private insurance c 0.388 0.045 0.049
Medicare 0.537 0.055 0.036 0.509
Medicaid 0.464 0.083 0.421 0.995
Other insurance 0.603 0.173 0.228 DNC
Uninsured / self-pay / no charge 0.452 0.133 0.648 0.342

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☐
L	M-H	M	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 outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service	Does the measure pass subcriterion1c? Yes☐ IF rationale supports relationship

1c.1 Structure-Process-Outcome Relationship *(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome)*:
This is a health outcome measure that has been linked to the process of care.

1c.2-3 Type of Evidence *(Check all that apply)*:
Selected individual studies (rather than entire body of evidence)

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

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: Moderate1c.27 Consistency: Moderate
1c.28 Attach evidence submission form:
1c.29 Attach appendix for supplemental materials:

Was the threshold criterion, *Importance to Measure and Report*, met?
(*1a & 1b must be rated moderate or high and 1c yes*) Yes☐ No☐

See Guidance for Definitions of Rating Scale: H=High; M=Moderate; L=Low; I=Insufficient; NA=Not Applicable

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

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 with disposition of "deceased" (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator

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*:
Discharges with disposition of "deceased" (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.
User may specify the time window; generally one calendar year.

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
Discharges, 18 years and older or MDC 14 (pregnancy, childbirth, and puerperium), in DRGs or MS-DRGs with less than 0.5% mortality rate. If a DRG is divided into two groups with or without "comorbidities or complications" or an MS-DRG is divided into three groups - with major, other, or no comorbidities or complications - then both DRGs or all MS-DRGs must have mortality rates below 0.5% to qualify for inclusion.

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)*:
User may specify the time window; generally one calendar year

2a1.7 Denominator Details (*All information required to identify and calculate the target population/denominator such as definitions, codes with descriptors, and/or specific data collection items/responses)*:
Presently low-mortality MS DRGs are used in the denominator definition.

Please note that the low-mortality DRGs are no longer in use, but are presented for historical compatibility only.

Low-mortality MS-DRG codes:

069 TRANSIENT ISCHEMIA
113 ORBITAL PROCEDURES W CC/MCC
114 ORBITAL PROCEDURES W/O CC/MCC
123 NEUROLOGICAL EYE DISORDERS
139 SALIVARY GLAND PROCEDURES
149 DYSEQUILIBRIUM
202 BRONCHITIS & ASTHMA W CC/MCC
203 BRONCHITIS & ASTHMA W/O CC/MCC
311 ANGINA PECTORIS
312 SYNCOPE & COLLAPSE
313 CHEST PAIN
483 MAJOR JOINT & LIMB REATTACHMENT PROC OF UPPER EXTREMITY W CC/MCC
484 MAJOR JOINT & LIMB REATTACHMENT PROC OF UPPER EXTREMITY W/O CC/MCC
488 KNEE PROCEDURES W/O PDX OF INFECTION W CC/MCC
489 KNEE PROCEDURES W/O PDX OF INFECTION W/O CC/MCC
490 BACK & NECK PROC EXC SPINAL FUSION W CC/MCC OR DISC DEVICE/NEUROSTIM
491 BACK & NECK PROC EXC SPINAL FUSION W/O CC/MCC
506 MAJOR THUMB OR JOINT PROCEDURES
513 HAND OR WRIST PROC, EXCEPT MAJOR THUMB OR JOINT PROC W CC/MCC
514 HAND OR WRIST PROC, EXCEPT MAJOR THUMB OR JOINT PROC W/O CC/MCC
537 SPRAINS, STRAINS, & DISLOCATIONS OF HIP, PELVIS & THIGH W CC/MCC
538 SPRAINS, STRAINS, & DISLOCATIONS OF HIP, PELVIS & THIGH W/O CC/MCC
582 MASTECTOMY FOR MALIGNANCY W CC/MCC
583 MASTECTOMY FOR MALIGNANCY W/O CC/MCC
691 URINARY STONES W ESW LITHOTRIPSY W CC/MCC
692 URINARY STONES W ESW LITHOTRIPSY W/O CC/MCC
697 URETHRAL STRICTURE
707 MAJOR MALE PELVIC PROCEDURES W CC/MCC
708 MAJOR MALE PELVIC PROCEDURES W/O CC/MCC
742 UTERINE & ADNEXA PROC FOR NON-MALIGNANCY W CC/MCC
743 UTERINE & ADNEXA PROC FOR NON-MALIGNANCY W/O CC/MCC
748 FEMALE REPRODUCTIVE SYSTEM RECONSTRUCTIVE PROCEDURES
760 MENSTRUAL & OTHER FEMALE REPRODUCTIVE SYSTEM DISORDERS W CC/MCC
761 MENSTRUAL & OTHER FEMALE REPRODUCTIVE SYSTEM DISORDERS W/O CC/MCC
765 CESAREAN SECTION W CC/MCC
766 CESAREAN SECTION W/O CC/MCC
767 VAGINAL DELIVERY W STERILIZATION &/OR D&C
768 VAGINAL DELIVERY W O.R. PROC EXCEPT STERIL &/OR D&C
769 POSTPARTUM & POST ABORTION DIAGNOSES W O.R. PROCEDURE
770 ABORTION W D&C, ASPIRATION CURETTAGE OR HYSTEROTOMY
774 VAGINAL DELIVERY W COMPLICATING DIAGNOSES
775 VAGINAL DELIVERY W/O COMPLICATING DIAGNOSES
776 POSTPARTUM & POST ABORTION DIAGNOSES W/O O.R. PROCEDURE
777 ECTOPIC PREGNANCY
778 THREATENED ABORTION
779 ABORTION W/O D&C
780 FALSE LABOR
781 OTHER ANTEPARTUM DIAGNOSES W MEDICAL COMPLICATIONS
782 OTHER ANTEPARTUM DIAGNOSES W/O MEDICAL COMPLICATIONS
793 FULL TERM NEONATE W MAJOR PROBLEMS
794 NEONATE W OTHER SIGNIFICANT PROBLEMS

880 ACUTE ADJUSTMENT REACTION & PSYCHOSOCIAL DYSFUNCTION
881 DEPRESSIVE NEUROSES
882 NEUROSES EXCEPT DEPRESSIVE
883 DISORDERS OF PERSONALITY & IMPULSE CONTROL
885 PSYCHOSES
886 BEHAVIORAL & DEVELOPMENTAL DISORDERS
887 OTHER MENTAL DISORDER DIAGNOSES
894 ALCOHOL/DRUG ABUSE OR DEPENDENCE, LEFT AMA
895 ALCOHOL/DRUG ABUSE OR DEPENDENCE W REHABILITATION THERAPY
906 HAND PROCEDURES FOR INJURIES

Low-mortality DRG codes:

037 ORBITAL PROCEDURES
045 NEUROLOGICAL EYE DISORDERS
050 SIALOADENECTOMY
051 SALIVARY GLAND PROCEDURES EXCEPT SIALOADENECTOMY
065 DYSEQUILIBRIUM
096 BRONCHITIS & ASTHMA AGE >17 W CC
097 BRONCHITIS & ASTHMA AGE >17 W/O CC
140 ANGINA PECTORIS
141 SYNCOPE & COLLAPSE W CC
142 SYNCOPE & COLLAPSE W/O CC
143 CHEST PAIN
228 MAJOR THUMB OR JOINT PROC,OR OTH HAND OR WRIST PROC W CC
229 HAND OR WRIST PROC, EXCEPT MAJOR JOINT PROC, W/O CC
237 SPRAINS, STRAINS, & DISLOCATIONS OF HIP, PELVIS & THIGH
257 TOTAL MASTECTOMY FOR MALIGNANCY W CC
258 TOTAL MASTECTOMY FOR MALIGNANCY W/O CC
323 URINARY STONES W CC, &/OR ESW LITHOTRIPSY
328 URETHRAL STRICTURE AGE >17 W CC
329 URETHRAL STRICTURE AGE >17 W/O CC
334 MAJOR MALE PELVIC PROCEDURES W CC
335 MAJOR MALE PELVIC PROCEDURES W/O CC
356 FEMALE REPRODUCTIVE SYSTEM RECONSTRUCTIVE PROCEDURES
358 UTERINE & ADNEXA PROC FOR NON-MALIGNANCY W CC
359 UTERINE & ADNEXA PROC FOR NON-MALIGNANCY W/O CC
369 MENSTRUAL & OTHER FEMALE REPRODUCTIVE SYSTEM DISORDERS
370 CESAREAN SECTION W CC
371 CESAREAN SECTION W/O CC
372 VAGINAL DELIVERY W COMPLICATING DIAGNOSES
373 VAGINAL DELIVERY W/O COMPLICATING DIAGNOSES
374 VAGINAL DELIVERY W STERILIZATION &/OR D&C
375 VAGINAL DELIVERY W O.R. PROC EXCEPT STERIL &/OR D&C
376 POSTPARTUM & POST ABORTION DIAGNOSES W/O O.R. PROCEDURE
377 POSTPARTUM & POST ABORTION DIAGNOSES W O.R. PROCEDURE
378 ECTOPIC PREGNANCY
379 THREATENED ABORTION
380 ABORTION W/O D&C
381 ABORTION W D&C, ASPIRATION CURETTAGE OR HYSTEROTOMY
382 FALSE LABOR
383 OTHER ANTEPARTUM DIAGNOSES W MEDICAL COMPLICATIONS

384 OTHER ANTEPARTUM DIAGNOSES W/O MEDICAL COMPLICATIONS
389 FULL TERM NEONATE W MAJOR PROBLEMS
390 NEONATE W OTHER SIGNIFICANT PROBLEMS
425 ACUTE ADJUSTMENT REACTION & PSYCHOSOCIAL DYSFUNCTION
426 DEPRESSIVE NEUROSES
427 NEUROSES EXCEPT DEPRESSIVE
428 DISORDERS OF PERSONALITY & IMPULSE CONTROL 430 PSYCHOSES
431 CHILDHOOD MENTAL DISORDERS
432 OTHER MENTAL DISORDER DIAGNOSES
433 ALCOHOL/DRUG ABUSE OR DEPENDENCE, LEFT AMA
441 HAND PROCEDURES FOR INJURIES
491 MAJOR JOINT & LIMB REATTACHMENT PROCEDURES OF UPPER EXTREMITY
499 BACK & NECK PROCEDURES EXCEPT SPINAL FUSION W CC
500 BACK & NECK PROCEDURES EXCEPT SPINAL FUSION W/O CC
503 KNEE PROCEDURES W/O PDX OF INFECTION 521 ALCOHOL/DRUG ABUSE OR DEPENDENCE W CC 522 ALC/DRUG ABUSE OR DEPEND W REHABILITATION THERAPY W/O CC
524 TRANSIENT ISCHEMIA

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
Exclude cases:
- with any code for trauma, cancer, or immunocompromised state
- transfer to an acute care facility (DISP = 2)
- with missing discharge disposition (DISP=missing), gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing) or principal diagnosis (DX1=missing)

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)*:
See Patient Safety Indicators Appendices:
- Appendix G – Trauma Diagnosis Codes
- Appendix H – Cancer Diagnosis Codes
- Appendix I – Immunocompromised State Diagnosis and Procedure Codes

Link to PSI appendices:
http://qualityindicators.ahrq.gov/Downloads/Software/SAS/V43/TechnicalSpecifications/PSI%20Appendices.pdf

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)*: Statistical risk model 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.)*:
The predicted value for each case is computed using a hierarchical model (logistic regression with hospital random effect) and covariates for gender, age (in 5-year age groups), modified CMS DRG, and the AHRQ Comorbidity category. The reference population used in the regression is the universe of discharges for states that participate in the HCUP State Inpatient Data (SID) for the years 2008, a database consisting of 42 states and approximately 30 million adult discharges. The expected rate is computed as the sum of the predicted value for each case divided by the number of cases for the unit of analysis of interest (i.e., hospital). The risk adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate, multiplied by the reference population rate.
Sex Female
Age 18 to 24

Age	25 to 29
Age	30 to 59
Age	65 to 69
Age	70 to 74
Age	75 to 79
Age	80 to 84
Age	85+
MDRG	413
MDRG	533
MDRG	1915
MDRG	2019
MDC	19
TRNSFER	Transfer-in
NOPRDAY	Procedure Days Data Not Available
COMORB	CHF
COMORB	NEURO
COMORB	CHRNLUNG
COMORB	HYPOTHY
COMORB	RENLFAIL
COMORB	OBESE
COMORB	ANEMDEF

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:
URL
http://qualityindicators.ahrq.gov/Downloads/Software/SAS/V43/Risk%20Adjustment%20Tables%20PSI%204.3.pdf
Not applicable

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

2a1.19 Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*: Better quality = Lower score

2a1.20 Calculation Algorithm/Measure Logic*(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*:
Each indicator is expressed as a rate, is defined as outcome of interest / population at risk or numerator / denominator. The AHRQ Quality Indicators (AHRQ QI) software performs six steps to produce the rates. 1) Discharge-level data is used to mark inpatient records containing the outcome of interest and 2) the population at risk. For provider indicators, the population at risk is also derived from hospital discharge records; for area indicators, the population at risk is derived from U.S. Census data. 3) Calculate observed rates. Using output from steps 1 and 2, rates are calculated for user-specified combinations of stratifiers. 4) Calculate expected rates. Regression coefficients from a reference population database are applied to the discharge records and aggregated to the provider or area level. For indicators that are not risk-adjusted, this is the reference population rate. 5) Calculate risk-adjusted rate. Use the indirect standardization to account for case-mix. For indicators that are not risk-adjusted, this is the same as the observed rate. 6) Calculate smoothed rate. A Univariate shrinkage factor is applied to the risk-adjusted rates. The shrinkage estimate reflects a reliability adjustment unique to each indicator

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)*:
The signal to noise ratio is the ratio of the between hospital variance (signal) to the within hospital variance (noise). The formula is signal / (signal + noise). The ratio itself is only a diagnostic for the degree of variance in the risk-adjusted rate systematically associated with the provider. Therefore, what matters is the magnitude of the variance in the "smoothed" rate (that is, the variance in the risk-adjusted rate after the application of the univariate shrinkage estimator based on the signal ratio).

2a2.3 Testing Results *(Reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
What the data demonstrate is systematic variation in the provider level rate of 0.025 to 0.779 to per 1,000 from the 5th to 95th percentile after a signal ratio of 0.502 is applied as the shrinkage estimator (that is, after accounting for variation due to random factors).

These findings were confirmed by Bernal-Delgado et al. (BMC Med Res Methodol 2012;12:19), who analyzed data from 171-175 Spanish hospitals in 2005-2006. They estimated PSI 2 virtually unchanged (as Spain uses ICD-9-CM for inpatient coding of both diagnoses and procedures, and MS DRGs for resource allocation). The Empirical Bayes estimator of systematic hospital-level variation in a two-stage hierarchical random effects model was 0.32, which was quite similar to the values for other NQF-endorsed AHRQ patient safety measures, such as Pressure Ulcer (PDI 2) and Postoperative DVT/PE (PSI 12). The rho statistic for overall cluster effect in the hierarchical model was 6%.

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:
No identified differences

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)*:
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.

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 panel approach was designed 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 tested in previous AHRQ projects. The process consisted of an initial independent assessment of each indicator by clinician panelists using a structured questionnaire, anonymized written feedback of the results of this initial assessment, a conference call among all panelists with facilitated discussion of each candidate indicator (focusing on areas of disagreement and suggestions for improving the indicator), followed by a final independent reassessment using the same questionnaire.

Twenty-one professional clinical organizations were invited to submit nominations. These organizations were selected based on the applicability of the specialty or subspecialty to patient safety indicators. Fifteen organizations submitted a total of 162 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.

To be eligible to participate, nominees were required to spend at least 30% of their work time on patient care, including hospitalized patients. 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), primary hospital of practice (i.e., funding source) and professional history (i.e., clinical education history, academic affiliation). Panelists were randomly selected from the pool of eligible clinicians so that each panel had diverse membership in terms of practice characteristics and setting. Fifty-seven of the 87 eligible panelists accepted the invitation to participate on specific panels.

Eight panels were formed, and PSI 2 was reviewed by a panel on general indicators of complications of medical care. In the final evaluation step, panelists were asked to rate the indicator on its overall usefulness, based on the panel discussion and all of the information that they had received about the indicator's rationale and characteristics. Indicators were considered as "endorsed" if the median "usefulness" score was 7 or greater (on a 1-9 scale) without significant disagreement (defined as one or fewer panelists who rated the indicator in the 1-3 range on the same scale).

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)*:
The overall panel rating of the usefulness of this indicator was 7.5 (on a 1-9 scale) with agreement.

Rosen et al. (Med Care. 2005;43:873-84) evaluated the construct validity of the AHRQ PSIs using VA inpatient data. PSI 2 was significantly correlated at the hospital level with 'failure to rescue' (now known as NQF-endorsed PSI 4, or Death among surgical inpatients with serious treatable complications), with a Spearman's rank correlation coefficient of 0.245. In an exploratory factor analysis, PSI 2 loaded together with PSI 3 (Pressure ulcer) and PSI 4 on a single factor that explained 31% of total variation (with loadings of 0.772, 0.768, and 0.835, respectively). See "Importance 1a.3" for other published evidence regarding the construct validity of this indicator.

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)*:
The impact of each category of exclusion (e.g., trauma, immunodeficiency, cancer, transfer out) on the number of denominator-eligible records was examined, both overall and by MS-DRG.

If the user´s data lacks present on admission (POA) information, then the likelihood that the outcome of interest and the covariates are present on admission is estimated using a Markov Chain Monte Carlo (MCMC) estimation procedure. That likelihood is then used to adjust the denominator to account for excluded cases. In this manner, we avoid specifically excluding records with missing POA information.

2b3.3 Results *(Provide statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses):*
The percentages of patients excluded because of trauma diagnoses, immunodeficiency diagnoses, cancer diagnoses, and transfer-out status vary widely across MS-DRGs, as would be expected based on the principal diagnoses assigned to that MS-DRG. Denominator exclusions are shown in the table below; numerator exclusions follow the same pattern but data confidentiality restrictions preclude enumerating them due to small cell sizes. The highlighted MS-DRGs are largely excluded from PSI 2, in practice, because they overlap so greatly with the exclusions for cancer or trauma.

[FOLLOWING TABLE EMAILED TO JESSICA WEBER, ANDREW LYZENGA, JESSE PINES, HEIDI BOSSLEY]

MS-DRG	Label	Trauma	Immune	Cancer	Transfer	Cases	Total	%excluded
69	Transient ischemia	1,247	3,751	18,628	1,922	141,207	166,755	15%
113	Orbital procedures w CC/MCC	1,380	75	206	41	325	2,027	84%
114	Orbital procedures w/o CC/MCC	1,544	*	216	28	567	2,356	76%
123	Neurological eye disorders	37	199	761	128	7,049	8,174	14%
139	Salivary gland procedures	21	43	800	*	3,465	4,336	20%
149	Dysequilibrium	803	1,189	7,396	492	59,568	69,448	14%
202	Bronchitis & asthma w CC/MCC	679	6,220	9,123	898	66,353	83,273	20%
203	Bronchitis & asthma w/o CC/MCC	159	433	5,584	673	117,082	123,931	6%
311	Angina pectoris	93	887	2,502	12,202	28,874	44,558	35%
312	Syncope & collapse	18,866	10,251	35,457	5,171	220,166	289,911	24%
313	Chest pain	1,583	15,910	38,784	14,296	468,507	539,080	13%
483	Major joint & limb reattachment proc of upper extremity w CC/MCC						3,379	339	1,852	117	8,107	13,794	41%
484	Major joint & limb reattachment proc of upper extremity w/o CC/MCC						3,906	23	2,810	86	23,078	29,903	23%
488	Knee procedures w/o pdx of infection w CC/MCC					2,413	192	566	103	4,163	7,437	44%
489	Knee procedures w/o pdx of infection w/o CC/MCC					5,070	12	841	88	12,922	18,933	32%
490	Back & neck proc exc spinal fusion w CC/MCC or disc device/neurostim						1,645	960	5,429	508	39,861	48,403	18%
491	Back & neck proc exc spinal fusion w/o CC/MCC					1,172	72	9,053	367	135,603	146,267	7%
506	Major thumb or joint procedures	575	53	138	19	1,955	2,740	29%
513	Hand or wrist proc, except major thumb or joint proc w CC/MCC						2,446	225	243	58	2,750	5,722	52%
514	Hand or wrist proc, except major thumb or joint proc w/o CC/MCC						2,293	*	182	40	5,053		

	7,577 33%											
537	Sprains, strains, & dislocations of hip, pelvis & thigh w CC/MCC					341	75	215	31	805		
	1,467 45%											
538	Sprains, strains, & dislocations of hip, pelvis & thigh w/o CC/MCC					587	*	122	34	1,060		
	1,804 41%											
582	Mastectomy for malignancy w CC/MCC	*	24	13,258	81	483	13,848	97%				
583	Mastectomy for malignancy w/o CC/MCC	*	*	19,694	61	3,682	23,441	84%				
691	Urinary stones w esw lithotripsy w CC/MCC	*	46	237	13	3,079	3,382	9%				
692	Urinary stones w esw lithotripsy w/o CC/MCC		*	*	89	*	2,000	2,093	4%			
697	Urethral stricture 17	22	227	14	774	1,054	27%					
707	Major male pelvic procedures w CC/MCC	16	24	13,473	46	1,091	14,650	93%				
708	Major male pelvic procedures w/o CC/MCC	*	*	60,856	56	1,646	62,567	97%				
742	Uterine & adnexa proc for non-malignancy w CC/MCC			946	1,191	4,294	278	78,479	85,188	8%		
743	Uterine & adnexa proc for non-malignancy w/o CC/MCC			158	135	13,192	327	373,125	386,937	4%		
748	Female reproductive system reconstructive procedures			68	117	3,038	54	42,514	45,791	7%		
760	Menstrual & other female reproductive system disorders w CC/MCC				80	568	755	297	7,764			
	9,464 18%											
761	Menstrual & other female reproductive system disorders w/o CC/MCC				84	52	516	265	19,993			
	20,910 4%											
765	Cesarean section w CC/MCC	845	1,094	1,515	1,218	416,529	421,201	1%				
766	Cesarean section w/o CC/MCC	165	52	1,535	902	822,823	825,477	0%				
767	Vaginal delivery w sterilization &/or D&C	45	63	163	66	85,114	85,451	0%				
768	Vaginal delivery w O.R. proc except steril &/or D&C	*	*	*	*	1,775	1,804	2%				
769	Postpartum & post abortion diagnoses w O.R. procedure			60	69	96	98	11,273	11,596	3%		
770	Abortion w D&C, aspiration curettage or hysterotomy	26	75	76	37	17,322	17,536	1%				
774	Vaginal delivery w complicating diagnoses	118	464	636	446	301,160	302,824	1%				
775	Vaginal delivery w/o complicating diagnoses	362	295	3,205	1,804	2,152,513	2,158,179	0%				
776	Postpartum & post abortion diagnoses w/o O.R. procedure			135	234	194	649	50,871	52,083	2%		
777	Ectopic pregnancy	31	28	75	56	23,256	23,446	1%				
778	Threatened abortion	47	59	147	4,244	64,772	69,269	6%				
779	Abortion w/o D&C19	58	41	81	17,481	17,680	1%					
780	False labor *	*	*	136	6,546	6,692	2%					
781	Other antepartum diagnoses w medical complications			1,619	1,003	983	6,066	189,047	198,718	5%		
782	Other antepartum diagnoses w/o medical complications				26	*	75	3,044	38,072	41,225	8%	
793	Full term neonate w major problems0	0	0	0	0	0	0%					
794	Neonate w other significant problems	0	0	0	0	65	65	0%				
880	Acute adjustment reaction & psychosocial dysfunction			440	626	2,241	941	28,813	33,061	13%		
881	Depressive neuroses	2,671	1,097	2,182	2,370	80,417	88,737	9%				
882	Neuroses except depressive	968	226	658	723	27,294	29,869	9%				
883	Disorders of personality & impulse control	243	438	144	328	7,389	8,542	13%				
885	Psychoses 18,326	11,092	24,502	30,018	822,301	906,239	9%					
886	Behavioral & developmental disorders	170	47	128	226	4,832	5,403	11%				
887	Other mental disorder diagnoses	34	58	132	54	1,995	2,273	12%				
894	Alcohol/drug abuse or dependence, left ama	763	903	578	0	39,884	42,128	5%				
895	Alcohol/drug abuse or dependence w rehabilitation therapy			342	643	734	782	32,277	34,778	7%		
906	Hand procedures for injuries	2,528	45	126	41	3,340	6,080	45%				

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

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

2b4.2 Analytic Method (*Describe methods and rationale for development and testing of risk model or risk stratification including selection of factors/variables)*:
Risk-adjustment models use a standard set of categories based on readily available classification systems for demographics, severity of illness and comorbidities. Within each category, covariates are initially selected based on a minimum of 30 cases in the outcome of interest. Then a stepwise regression process on a development sample is used to select a parsimonious set of covariates where p<.05. Model is then tested on a validation sample.

If the user´s data lacks present on admission information, then the likelihood that the outcome of interest and the covariates are present on admission is estimated using a Markov Chain Monte Carlo (MCMC) estimation procedure. That likelihood is then used to adjust the oberved and expected rates.

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)*:
AHRQ's current risk-adjustment model has a c-statistic = 0.839.

Barker et al. (Med J Australia 2011;195:89-94) confirmed the validity of the AHRQ risk-adjustment model in an entirely different setting, using data from 122 public hospitals in Victoria (Australia) in 2006-2008. The Australian translation was inexact because Australia uses ICD-10-AM and Australian DRGs. Their basic finding was that the indicator was sensitive to patient characteristics that are included in AHRQ's risk-adjustment model: age, male sex, comorbidities, inter-hospital transfer, SNF transfer.

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)*:
Posterior probability distribution parameterized using the Gamma distribution

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)*:
Raw Rates (numerator / denominator):

5th	25th	Median	75th	95th
0.000025	0.000102	0.000214	0.000391	0.000779

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)*: Note that this indicator demonstrates significant disparities by socioeconomic status, with average rates of 0.474 in the lowest income quartile, 0.406 in the lower-middle income quartile, 0.400 in the upper-middle income quartile, and 0.378 in the highest income quartile. Similarly, the average national rates in 2007 were 0.388 for privately insured patients versus 0.464 for Medicaid patients and 0.452 for uninsured and self-pay patients.

In regard to figures below:
1st figure: estimate per 1,000, risk adjusted rates
2nd figure: standard error
3rd figure: p value relative to marked group (marked group = “c”)
4th figure: p value: current year relative to prior year

Key:
"c": Reference for p-value test statistics
"*": Data do not meet criteria for statistical reliability, data quality, or confidentiality

Patient characteristic:
Age groups for conditions affecting any age
18-44 c 0.142 0.012 0.917
45-64 0.448 0.037 0.000 0.336
65 and over 1.928 0.093 0.000 0.397

Age groups for conditions affecting primarily elderly:
65-69 c 0.645 0.103 0.072
70-74 1.030 0.147 0.031 0.263
75-79 1.670 0.180 0.000 0.844
80-84 2.479 0.244 0.000 0.821
85 and over 4.506 0.343 0.000 0.892

Gender:
Male c 0.593 0.037 0.241
Female 0.357 0.016 0.000 0.506

Median income of patient´s ZIP code:
First quartile (lowest income) 0.474 0.028 0.018 0.744
Second quartile 0.406 0.029 0.492 0.124
Third quartile 0.400 0.029 0.587 0.844
Fourth quartile (highest income) c 0.sss378 0.029 0.393

Location of patient residence (NCHS):
Large central metropolitan 0.386 0.026 0.090 0.377
Large fringe metropolitan c 0.454 0.030 0.118

Medium metropolitan 0.442 0.035 0.795 0.716
Small metropolitan 0.363 0.043 0.084 0.465
Micropolitan 0.443 0.046 0.841 0.570
Not metropolitan or micropolitan 0.413 0.052 0.501 0.167

Expected payment source:
Private insurance c 0.388 0.045 0.049
Medicare 0.537 0.055 0.036 0.509
Medicaid 0.464 0.083 0.421 0.995
Other insurance 0.603 0.173 0.228 DNC
Uninsured / self-pay / no charge 0.452 0.133 0.648 0.342

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 Actual/Planned Use *(Check all the planned 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 10 realms.

Colorado (state hospital association)
Colorado Hospital Report Card
http://www.cohospitalquality.org/index.php?option=com_frontpage&Itemid=1

Florida (state)
Florida Health Finder

http://www.floridahealthfinder.gov/

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

Maine (state)
Maine Health Data Organization
http://gateway.maine.gov/mhdo2008Monahrq/home.html

Nevada (state hospital association)
Nevada Hospital Association Hospital Performance
http://www.nvhospitalquality.net/

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

Oklahoma (state)
Oklahoma Hospital Report
http://www.ok.gov/health/documents/08%20Hospital%20AR.pdf

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

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) 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., 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 readily use the risk-adjusted rate and the observed to expected results to identify opportunities for improvment for specific patient populations based on default stratifiers or risk adjustment model covariates. In addition, comparative data from the AHRQ SID and NIS databases provides relative performance information.

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

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure before a final recommendation is made.

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

5a. Harmonization

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

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

5b. Competing Measure(s)

5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures *(e.g., a more valid or efficient way to measure quality);* OR provide a rationale for the additive value of endorsing an additional measure. *(Provide analyses when possible)*:

Not applicable

CONTACT INFORMATION

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

Co.2 Point of Contact: John, Bott, AHRQ Quality Indicators Senior Analyst, 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, AHRQ Quality Indicators Senior Analyst, John.Bott@ahrq.hhs.gov, 301-427-1317-

Co.5 Submitter: John, Bott, AHRQ Quality Indicators Senior Analyst, 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, AHRQ Quality Indicators Senior Analyst, 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 Iezzoni 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