

NATIONAL QUALITY FORUM

Measure Evaluation 4.1 December 2009

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The subcriteria and most of the footnotes from the [evaluation criteria](#) are provided in Word comments within the form and will appear if your cursor is over the highlighted area. Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all **yellow highlighted** areas of the form. Evaluate the extent to which each subcriterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the subcriteria (yellow highlighted areas).*

Steering Committee: Complete all **pink** highlighted areas of the form. Review the workgroup/TAP assessment of the subcriteria, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met

C = Completely (unquestionably demonstrated to meet the criterion)

P = Partially (demonstrated to partially meet the criterion)

M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)

N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)

NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) NQF Review #: 0273	NQF Project: Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION	
De.1 Measure Title: Perforated Appendix Admission Rate (PQI 2)	
De.2 Brief description of measure: Percentage of admissions for appendicitis within county with perforated appendix.	
1.1-2 Type of Measure: Outcome	
De.3 If included in a composite or paired with another measure, please identify composite or paired measure Not applicable	
De.4 National Priority Partners Priority Area: Population health, Safety	
De.5 IOM Quality Domain: Effectiveness	
De.6 Consumer Care Need: Staying healthy	

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
<p>A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. <i>Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.</i></p> <p>A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)? Yes</p> <p>A.2 Indicate if Proprietary Measure (as defined in measure steward agreement):</p> <p>A.3 Measure Steward Agreement: Government entity and in the public domain - no agreement necessary</p> <p>A.4 Measure Steward Agreement attached:</p>	<p>A</p> <p>Y <input checked="" type="checkbox"/></p> <p>N <input type="checkbox"/></p>
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least	<p>B</p> <p>Y <input checked="" type="checkbox"/></p>

every 3 years. Yes, information provided in contact section	N <input type="checkbox"/>
C. The intended use of the measure includes <u>both</u> public reporting <u>and</u> quality improvement. ► Purpose: Public reporting, Internal quality improvement	C Y <input checked="" type="checkbox"/> N <input type="checkbox"/>
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1 Testing: Yes, fully developed and tested D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes	D Y <input checked="" type="checkbox"/> N <input type="checkbox"/>
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (if submission returned):	Met Y <input checked="" type="checkbox"/> N <input type="checkbox"/>
Staff Notes to Reviewers (issues or questions regarding any criteria): 1c.8 Citations for Evidence: Updated citations will be presented at the in-person meeting.	
Staff Reviewer Name(s): Alexis Forman	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. 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	Eval Rati ng
(for NQF staff use) Specific NPP goal: Not related to a specific NPP goal.	
1a.1 Demonstrated High Impact Aspect of Healthcare: High resource use, Patient/societal consequences of poor quality 1a.2 1a.3 Summary of Evidence of High Impact: Perforated appendix occurs in one-fourth to one-third of hospitalized acute appendicitis patients [1] Approximately 280,000 appendectomies are performed each year in the United States. [2] Most are performed as emergencies to avoid the complications of perforated appendicitis; an entity believed to result from delay in surgical removal of the appendix after the appendix has become inflamed. The fear of appendicitis complications results in more emergency general surgical operations than any other disease. [3] Negative exploration rates as high as 30% are considered acceptable for women presenting with lower abdominal pain. [4] A retrospective analysis for all patients 18 y of age and over with acute appendicitis between July 1, 2005 and December 31, 2008 at a teaching hospital identified 1003 patients with acute appendicitis of whom 239 (23.8%) had perforated appendix. Patients with public insurance were significantly more likely to have perforated disease (P < 0.001) as were patients in the older age groups (41-64 and =65) (35.8% and 38.24%, respectively, versus 19.2% for those 18-40; P < 0.001). The patients who presented with perforation had a greater length of stay (2.71 ± 2.14 versus 6.04 ± 3.91 d, P < 0.001). [5] 1a.4 Citations for Evidence of High Impact: Updated citations will be presented in the May Steering Committee meeting [1] Braveman P, Schaaf VM, Egerter S, et al. Insurance-related differences in the risk of ruptured appendix	1a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>

[see comments]. N Engl J Med 1994;331(7):444-9. PMID: 7880234.
 [2] National Center for Health Statistics. Ambulatory and Inpatient Procedures in the United States, 1996. National Center for Health Statistics Series 13[No. 139]. 2004.
 [3] Livingston EH, Woodward WA, Sarosi GA, Haley RW. Disconnect Between Incidence of Nonperforated and Perforated Appendicitis: Implications for Pathophysiology and Management. Ann Surg. 2007 June; 245(6): 886-892. doi: 10.1097/01.sla.0000256391.05233.aa.
 [4] Larsson PG, Henriksson G, Olsson M, et al. Laparoscopy reduces unnecessary appendicectomies and improves diagnosis in fertile women. A randomized study. Surg Endosc. 2001;15:200-202. PMID: 11285968
 [5] Boomer L, Freeman J, Landrito E, Feliz A. Perforation in adults with acute appendicitis linked to insurance status, not ethnicity. J Surg Res. 2010 Oct;163(2):221-4. Epub 2010 May 21. PMID: 20599222

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: Timely diagnosis and treatment may reduce the incidence of perforated appendix, and lower rates represent better quality care.

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:

Adjusted per 100,000 rates by patient and hospital characteristics, 2007

Mean	Standard error	Location	P-value: Relative to Northeast
246.2854	0.719	Northeast	1.000
293.2244	0.786	Midwest	0.000
289.0073	0.677	South	0.000
286.8724	0.341	West	0.000

1b.3 Citations for data on performance gap:

See the following report for a complete treatment of the methodology: "Methods: Applying AHRQ Quality Indicators to Healthcare Cost and Utilization Project (HCUP) Data for the National Healthcare Quality Report" [URL: <http://hcupnet.ahrq.gov/QI%20Methods.pdf?JS=Y>]

1b.4 Summary of Data on disparities by population group:

Adjusted per 100,000 rates by patient characteristics, 2007

Estimate	Standard error	Age: for conditions affecting any age
203.578	3.449	18-44
390.937	4.99	45-64
516.977	7.304	65 and over

Estimate	Standard error	Age: for conditions affecting elderly
483.585	11.929	65-69
494.937	14.249	70-74
535.493	15.712	75-79
546.261	18.545	80-84
593.997	22.816	85 and over

Estimate	Standard error	Gender
303.352	3.045	Male
249.47	2.982	Female

Estimate	Standard error	Median income of patient's ZIP code
300.24	4.501	First quartile (lowest income)
283.229	4.24	Second quartile
283.319	4.064	Third quartile
257.117	3.938	Fourth quartile (highest income)

1b
 C
 P
 M
 N

Estimate	Standard error	Location of patient residence (NCHS)	
276.481	3.868	Large central metropolitan	
269.158	4.195	Large fringe metropolitan	
274.846	4.873	Medium metropolitan	
296.272	7.384	Small metropolitan	
299.424	6.613	Micropolitan	
311.401	9.069	Not metropolitan or micropolitan	
Estimate	Standard error	Expected payment source	
Private insurance	263.3063.241	Private insurance	
Medicare	314.21214.617	Medicare	
Medicaid	300.9859.005	Medicaid	
Other insurance	310.40511.033	Other insurance	
Uninsured / self-pay / no charge	296.9538.218	Uninsured / self-pay / no charge	
Estimate	Standard error	Hospital Ownership/control	
274.734	2.479	Private, not-for-profit	
284.808	5.805	Private, for-profit	
305.153	6.159	Public	
Estimate	Standard error	Teaching status	
274.783	4.014	Teaching	
282.269	2.557	Nonteaching	
Estimate	Standard error	Location of hospital	
279.332	3.607	Large central metropolitan	
264.164	4.666	Large fringe metropolitan	
279.024	4.944	Medium metropolitan	
298.37	6.997	Small metropolitan	
292.701	6.42	Micropolitan	
308.891	12.663	Not metropolitan or micropolitan	
Estimate	Standard error	Bed size of hospital	
287.66	5.775	Less than 100	
276.441	3.181	100 - 299	
279.597	4.164	300 - 499	
285.324	5.751	500 or more	
1b.5 Citations for data on Disparities:			
See the following report for a complete treatment of the methodology: "Methods: Applying AHRQ Quality Indicators to Healthcare Cost and Utilization Project (HCUP) Data for the National Healthcare Quality Report" [URL: http://hcupnet.ahrq.gov/QI%20Methods.pdf?JS=Y]			
1c. Outcome or Evidence to Support Measure Focus			
1c.1 Relationship to Outcomes (For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population): With prompt and appropriate care, acute appendicitis should not progress to perforation or rupture. However, rates of perforated appendix are higher in the uninsured or underinsured in both the adult and pediatric population. In addition, perforated appendix rates also vary by race. Potential reasons for differences by insurance status or race include patients failing to seek appropriate care, access to care difficulties, or misdiagnoses and poor quality care that result in delays in receiving surgery.			
Perforated appendix is a potentially avoidable hospitalization / ambulatory care sensitive condition			
			1c <input type="checkbox"/> C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N

indicator. These indicators are not measures of hospital quality, but rather measures of access to high quality outpatient care, and as such are defined with area level denominators.

1c.2-3. Type of Evidence: Expert opinion, Systematic synthesis of research

1c.4 Summary of Evidence (as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome):

Hospital admission for perforated appendix is a PQI that would be of most interest to comprehensive health care delivery systems. With prompt and appropriate care, acute appendicitis should not progress to perforation or rupture. Rates for perforated appendix are higher in the uninsured or underinsured in both adult and pediatric populations, which may be caused by patients failing to seek appropriate care, difficulty in accessing care, or misdiagnoses and poor quality care.

Perforated appendix rates vary systematically by race, although the cause is unknown. Areas with high rates of perforated appendix may want to target points of intervention by using chart reviews and other supplemental data to investigate the reasons for delay in receiving surgery. Hospital contributions to the overall area rate may be particularly useful for this indicator, because misdiagnoses and other delays in receiving surgery in an emergency room may contribute substantially to the rate.

1c.5 Rating of strength/quality of evidence (also provide narrative description of the rating and by whom):

17 Smoothing recommended Testing, rating, and review were conducted by the project team. A full report on the literature review and empirical evaluation can be found in Refinement of the HCUP Quality Indicators by the UCSF-Stanford EPC, Detailed coding information for each QI is provided in the document Prevention Quality Indicators Technical Specifications. Rating of performance on empirical evaluations, ranged from 0 to 26. The scores were intended as a guide for summarizing the performance of each indicator on four empirical tests of precision (signal variance, area-level share, signal ratio, and R-squared) and five tests of minimum bias (rank correlation, top and bottom decile movement, absolute change, and change over two deciles), as described in the previous section.

1c.6 Method for rating evidence: The project team conducted extensive empirical testing of all potential indicators using the 1995-97 HCUP State Inpatient Databases (SID) and Nationwide Inpatient Sample (NIS) to determine precision, bias, and construct validity. The 1997 SID contains uniform data on inpatient stays in community hospitals for 22 States covering approximately 60% of all U.S. hospital discharges. The NIS is designed to approximate a 20% of U.S. community hospitals and includes all stays in the sampled hospitals. Each year of the NIS contains between 6 million and 7 million records from about 1,000 hospitals. The NIS combines a subset of the SID data, hospital-level variables, and hospital and discharge weights for producing national estimates. The project team conducted tests to examine three things: precision, bias, and construct validity.

Precision. The first step in the analysis involved precision tests to determine the reliability of the indicator for distinguishing real differences in provider performance. For indicators that may be used for quality improvement, it is important to know with what precision, or surety, a measure can be attributed to an actual construct rather than random variation.

For each indicator, the variance can be broken down into three components: variation within a provider (actual differences in performance due to differing patient characteristics), variation among providers (actual differences in performance among providers), and random variation. An ideal indicator would have a substantial amount of the variance explained by between-provider variance, possibly resulting from differences in quality of care, and a minimum amount of random variation. The project team performed four tests of precision to estimate the magnitude of between-provider variance on each indicator:

- Signal standard deviation was used to measure the extent to which performance of the QI varies systematically across hospitals or areas.
- Provider/area variation share was used to calculate the percentage of signal (or true) variance relative to the total variance of the QI.
- Signal-to-noise ratio was used to measure the percentage of the apparent variation in QIs across providers that is truly related to systematic differences across providers and not random variations (noise) from year to year.
- In-sample R-squared was used to identify the incremental benefit of applying multivariate signal extraction methods for identifying additional signal on top of the signal-to-noise ratio.

In general, random variation is most problematic when there are relatively few observations per provider, when adverse outcome rates are relatively low, and when providers have little control over patient outcomes

or variation in important processes of care is minimal. If a large number of patient factors that are difficult to observe influence whether or not a patient has an adverse outcome, it may be difficult to separate the “quality signal” from the surrounding noise. Two signal extraction techniques were applied to improve the precision of an indicator:

- Univariate methods were used to estimate the “true” quality signal of an indicator based on information from the specific indicator and 1 year of data.
- Multivariate signal extraction (MSX) methods were used to estimate the “true” quality signal based on information from a set of indicators and multiple years of data. In most cases, MSX methods extracted additional signal, which provided much more precise estimates of true hospital or area quality.

Bias. To determine the sensitivity of potential QIs to bias from differences in patient severity, unadjusted performance measures for specific hospitals were compared with performance measures that had been adjusted for age and gender. All of the PQIs and some of the Inpatient Quality Indicators (IQIs) could only be risk-adjusted for age and sex. The 3M™ APR-DRG System Version 12 with Severity of Illness and Risk of Mortality subclasses was used for risk adjustment of the utilization indicators and the in-hospital mortality indicators, respectively. Five empirical tests were performed to investigate the degree of bias in an indicator:

- Rank correlation coefficient of the area or hospital with (and without) risk adjustment—gives the overall impact of risk adjustment on relative provider or area performance.
- Average absolute value of change relative to mean—highlights the amount of absolute change in performance, without reference to other providers’ performance.
- Percentage of highly ranked hospitals that remain in high decile—reports the percentage of hospitals or areas that are in the highest deciles without risk adjustment that remain there after risk adjustment is performed.
- Percentage of lowly ranked hospitals that remain in low decile—reports the percentage of hospitals or areas that are in the lowest deciles without risk adjustment that remain there after risk adjustment is performed.
- Percentage that change more than two deciles—identifies the percentage of hospitals whose relative rank changes by a substantial percentage (more than 20%) with and without risk adjustment.

Construct validity. Construct validity analyses provided information regarding the relatedness or independence of the indicators. If quality indicators do indeed measure quality, then two measures of the same construct would be expected to yield similar results. The team used factor analysis to reveal underlying patterns among large numbers of variables—in this case, to measure the degree of relatedness between indicators. In addition, they analyzed correlation matrices for indicators.

1c.7 Summary of Controversy/Contradictory Evidence: See the following for a complete treatment of the topic: http://www.qualityindicators.ahrq.gov/downloads/pqi/pqi_guide_v31.pdf

Note: The Literature Review Findings column summarizes evidence specific to each potential concern on the link between the PQIs and quality of care, as described in step 3 above. A question mark (?) indicates that the concern is theoretical or suggested, but no specific evidence was found in the literature. A check mark indicates that the concern has been demonstrated in the literature.

1c.8 Citations for Evidence (other than guidelines): Updated citations will be presented in the May Steering Committee meeting

http://www.qualityindicators.ahrq.gov/downloads/pqi/pqi_guide_v31.pdf

1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number):
Not applicable

1c.10 Clinical Practice Guideline Citation: Not applicable

1c.11 National Guideline Clearinghouse or other URL: Not applicable

1c.12 Rating of strength of recommendation (also provide narrative description of the rating and by whom):
Not applicable

1c.13 Method for rating strength of recommendation (If different from [USPSTF system](#), also describe rating and how it relates to USPSTF):
Not applicable

1c.14 Rationale for using this guideline over others:

Not applicable	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Importance to Measure and Report</i>?	1
Steering Committee: Was the threshold criterion, <i>Importance to Measure and Report</i>, met? Rationale:	1 Y <input type="checkbox"/> N <input type="checkbox"/>
2. 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)	Eval Rati ng
2a. MEASURE SPECIFICATIONS	
<p>S.1 Do you have a web page where current detailed measure specifications can be obtained? S.2 If yes, provide web page URL:</p> <p>2a. Precisely Specified</p>	
<p>2a.1 Numerator Statement (<i>Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome</i>): All discharges with ICD-9-CM diagnosis code for perforations or abscesses of appendix in any field among cases meeting the inclusion rules for the denominator.</p> <p>2a.2 Numerator Time Window (<i>The time period in which cases are eligible for inclusion in the numerator</i>): Time window can be determined by user, but is generally a calendar year.</p> <p>2a.3 Numerator Details (<i>All information required to collect/calculate the numerator, including all codes, logic, and definitions</i>): All discharges with ICD-9-CM diagnosis code for perforations or abscesses of appendix in any field among cases meeting the inclusion rules for the denominator. Include ICD-9-CM diagnosis codes: 5400 AC APPEND W PERITONITIS 5401 ABSCESS OF APPENDIX</p> <p>Exclude cases:</p> <ul style="list-style-type: none"> • transfer from a hospital (different facility) • transfer from a skilled Nursing Facility (SNF) or Intermediate Care Facility (ICF) • transfer from another health care facility • MDC 14 (pregnancy, childbirth, and puerperium) 	
<p>2a.4 Denominator Statement (<i>Brief, text description of the denominator - target population being measured</i>): All non-maternal discharges of age 18 years and older in Metro Area1 or county with diagnosis code for appendicitis in any field.</p> <p>2a.5 Target population gender: Female, Male 2a.6 Target population age range: 18 and older</p>	
<p>2a.7 Denominator Time Window (<i>The time period in which cases are eligible for inclusion in the denominator</i>): Calendar year</p>	2a- spe cs C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
<p>2a.8 Denominator Details (<i>All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions</i>): All non-maternal discharges of age 18 years and older in Metro Area1 or county with diagnosis code for</p>	

appendicitis in any field.

Include ICD-9-CM diagnosis codes (population at risk):

5400

AC APPEND W PERITONITIS

5401

ABSCESS OF APPENDIX

5409

ACUTE APPENDICITIS NOS

541

APPENDICITIS NOS

2a.9 Denominator Exclusions (Brief text description of exclusions from the target population): Not applicable.

2a.10 Denominator Exclusion Details (All information required to collect exclusions to the denominator, including all codes, logic, and definitions):
Not applicable.

2a.11 Stratification Details/Variables (All information required to stratify the measure including the stratification variables, all codes, logic, and definitions):

Observed rates may be stratified by gender, age (5-year age groups), race / ethnicity.

2a.12-13 Risk Adjustment Type: Risk adjustment method widely or commercially available

2a.14 Risk Adjustment Methodology/Variables (List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method):

The predicted value for each case is computed using a logistic regression model and covariates for gender and age in years (in 5-year age groups). The reference population used in the model is the universe of discharges for states that participate in the HCUP State Inpatient Databases (SID) for the year 2007 (updated annually), a database consisting of 43 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., county, state, and region). The risk adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate, multiplied by the reference population rate

2a.15-17 Detailed risk model available Web page URL or attachment: URL

[http://www.qualityindicators.ahrq.gov/downloads/pqi/PQI%20Risk%20Adjustment%20Tables%20\(Versions%204%202\).pdf](http://www.qualityindicators.ahrq.gov/downloads/pqi/PQI%20Risk%20Adjustment%20Tables%20(Versions%204%202).pdf)

2a.18-19 Type of Score: Rate/proportion

2a.20 Interpretation of Score: Better quality = Lower score

2a.21 Calculation Algorithm (Describe the calculation of the measure as a flowchart or series of steps):

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 five 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. 5) Calculate risk-adjusted rate. Use the indirect standardization to account for case-mix. 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. Full information on calculation algorithms and specifications can be found at http://qualityindicators.ahrq.gov/PQI_download.htm

2a.22 Describe the method for discriminating performance (e.g., significance testing):

Significance testing is not prescribed by the software. Users may calculate a confidence interval for the risk-adjusted rates and a posterior probability interval for the smoothed rates at a 95% or 99% level. Users may define the relevant benchmark and the methods of discriminating performance according to their application.

2a.23 Sampling (Survey) Methodology If measure is based on a sample (or survey), provide instructions for

<p>obtaining the sample, conducting the survey and guidance on minimum sample size (response rate): Not applicable</p>	
<p>2a.24 Data Source (Check the source(s) for which the measure is specified and tested) Electronic administrative data/claims</p> <p>2a.25 Data source/data collection instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.): The data source is hospital discharge data such as the HCUP State Inpatient Databases (SID) or equivalent using UB-04 coding standards. The data collection instrument is public-use AHRQ QI software available in SAS or Windows versions.</p> <p>2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL None http://www.qualityindicators.ahrq.gov/software.htm</p> <p>2a.29-31 Data dictionary/code table web page URL or attachment: URL None http://www.qualityindicators.ahrq.gov/downloads/winqi/AHRQ_QI_Windows_Software_Documentation_V41a.pdf</p> <p>2a.32-35 Level of Measurement/Analysis (Check the level(s) for which the measure is specified and tested) Population: states, Population: counties or cities</p> <p>2a.36-37 Care Settings (Check the setting(s) for which the measure is specified and tested) Ambulatory Care: Office</p> <p>2a.38-41 Clinical Services (Healthcare services being measured, check all that apply) Clinicians: Physicians (MD/DO)</p>	
TESTING/ANALYSIS	
<p>2b. Reliability testing</p> <p>2b.1 Data/sample (description of data/sample and size): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>2b.2 Analytic Method (type of reliability & rationale, method for testing): Expert panels and empirical analysis</p> <p>2b.3 Testing Results (reliability statistics, assessment of adequacy in the context of norms for the test conducted): Perforated appendix occurs in one-fourth to one-third of hospitalized acute appendicitis patients.³⁹ Based on empirical evidence, this indicator is precise, with a raw area level rate of 33.3% and a substantial standard deviation of 14.4%. Relative to other indicators, a higher percentage of the variation occurs at the area level rather than the discharge level. However, the signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is low, at 26.5%, indicating that much of the observed differences in age-sex adjusted rates likely do not represent true differences across areas. Applying multivariate signal extraction methods can improve estimation of true differences in area performance.</p>	<p>2b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
<p>2c. Validity testing</p> <p>2c.1 Data/sample (description of data/sample and size): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>2c.2 Analytic Method (type of validity & rationale, method for testing): Expert panels and empirical analysis</p> <p>2c.3 Testing Results (statistical results, assessment of adequacy in the context of norms for the test conducted):</p>	<p>2c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>

<p>Braveman et al. found that the rate of perforated appendix was 50% higher for patients with no insurance or Medicaid than HMO-covered patients, and 20% higher for patients with private fee-for-service insurance. A follow-up study by Blumberg et al. concluded that the high rate of perforated appendix in the black population at an HMO may be explained by delay in seeking care, rather than differences in the quality of health care.⁴² Weissman et al. found that uninsured (but not Medicaid) patients are at increased risk for ruptured appendix after adjusting for age and sex.⁴³ Based on empirical results, areas with high rates of perforated appendix admissions tend to have lower rates of admissions for other ACSCs.</p>	
<p>2d. Exclusions Justified</p> <p>2d.1 Summary of Evidence supporting exclusion(s): Exclusions remove cases where the outcome of interest is less likely to be preventable or with no or very low risk</p> <p>2d.2 Citations for Evidence: Updated citations will be presented in the May Steering Committee meeting</p> <p>Refinement of the HCUP Quality Indicators (Technical Review), May 2001 http://qualityindicators.ahrq.gov/downloads/technical/qi_technical_review.zip</p> <p>2d.3 Data/sample (description of data/sample and size): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>2d.4 Analytic Method (type analysis & rationale): Expert panel and descriptive analyses stratified by exclusion categories</p> <p>2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses): Refinement of the HCUP Quality Indicators (Technical Review), May 2001 http://qualityindicators.ahrq.gov/downloads/technical/qi_technical_review.zip</p>	<p>2d</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>
<p>2e. Risk Adjustment for Outcomes/ Resource Use Measures</p> <p>2e.1 Data/sample (description of data/sample and size): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>2e.2 Analytic Method (type of risk adjustment, analysis, & rationale): Expert panel and descriptive analyses stratified by exclusion categories</p> <p>2e.3 Testing Results (risk model performance metrics): 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</p> <p>2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: Not applicable</p>	<p>2e</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>
<p>2f. Identification of Meaningful Differences in Performance</p> <p>2f.1 Data/sample from Testing or Current Use (description of data/sample and size): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance (type of analysis & rationale): Posterior probability distribution parameterized using the Gamma distribution</p> <p>2f.3 Provide Measure Scores from Testing or Current Use (description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance):</p>	<p>2f</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>

5th	25th	Median	75th	95th	
0.211924	0.250104	0.279212	0.310497	0.359439	
2g. Comparability of Multiple Data Sources/Methods 2g.1 Data/sample (description of data/sample and size): Not applicable 2g.2 Analytic Method (type of analysis & rationale): Not applicable 2g.3 Testing Results (e.g., correlation statistics, comparison of rankings): Not applicable					2g C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
2h. Disparities in Care 2h.1 If measure is stratified, provide stratified results (scores by stratified categories/cohorts): Median income of patient's ZIP code: 1) Estimate 2) Standard error 3) P-value: Relative to marked group-c 4) P-value: 2007 relative to 2006 First quartile (lowest income) 300.240 4.501 0.000 0.097 Second quartile 283.229 4.240 0.000 0.119 Third quartile 283.319 4.064 0.000 0.857 Fourth quartile (highest income)c 257.117 3.938 0.148 2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans: Users may stratify based on gender and race/ethnicity					2h C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Scientific Acceptability of Measure Properties</i>?					2
Steering Committee: Overall, to what extent was the criterion, <i>Scientific Acceptability of Measure Properties</i>, met? Rationale:					2 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
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)					Eval Rati ng
3a. Meaningful, Understandable, and Useful Information 3a.1 Current Use: In use 3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years): 1) State of California: http://www.oshpd.ca.gov/hid/products/preventable_hospitalizations/pdfs/PH_REPORT_WEB.pdf 2) State of New Jersey: Find and Compare Quality Care in New Jersey Hospitals, http://www.nj.gov/health/healthcarequality/ 3) Niagara Health Quality Coalition and Alliance for Quality Health Care: New York State Hospital Report Card, http://www.myhealthfinder.com/ 4) State of Texas: Reports on Hospital Performance, http://www.dshs.state.tx.us/thcic/ 5) Maine: Maine Health Data Organization: http://gateway.maine.gov/mhdo2008Monahrq/home.html 6) Hawaii: Hawaii Health Information Corporation: http://hnic.org/publicreports.asp					3a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>

<p>7) Nevada: Nevada Compare Care: http://www.nevadacomparecare.net/monahrq/home.html</p> <p>In use as a part of the AHRQ Quality Indicators. They are reported in numerous forums including: http://hcupnet.ahrq.gov/Hcupnet.jsp?Id=EB57801381F71C41&Form=MAINSEL&JS=Y&Action=%3E%3ENext%3E%3E&_MAINSEL=AHRQ%20Quality%20Indicators</p> <p>This measure is used in the Monahrq system that is provide for public reporting and quality improvement throughout the United States: http://monahrq.ahrq.gov/</p> <p>3a.3 If used in other programs/initiatives (If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). <u>If not used for QI</u>, state the plans to achieve use for QI within 3 years): The software is publicly available free of charge (www.qualityindicators.ahrq.gov/). Users apply the software to their own administrative data (UB-04 or claims) that is readily available. Hundreds of users have downloaded AHRQ Quality Indicator software.</p> <p>This measure is used in the Monahrq system that is provide for public reporting and quality improvement throughout the United States: http://monahrq.ahrq.gov/</p> <p>Testing of Interpretability (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement)</p> <p>3a.4 Data/sample (description of data/sample and size): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>3a.5 Methods (e.g., focus group, survey, QI project): AHRQ has developed the Quality Indicators Mapping Tool to facilitate use of the Prevention Quality Indicators and incorporated the tool into the MONAHRQ software, which has undergone user beta testing and is now available for download</p> <p>3a.6 Results (qualitative and/or quantitative results and conclusions): Several states including Maine, Hawaii and Nevada have begun public reporting using the MONAHRQ tool. See http://monahrq.ahrq.gov/</p>	
<p>3b/3c. Relation to other NQF-endorsed measures</p> <p>3b.1 NQF # and Title of similar or related measures: No related measures found.</p>	
<p>(for NQF staff use) Notes on similar/related <u>endorsed</u> or submitted measures: No similar/related endorsed or submitted measures.</p>	
<p>3b. Harmonization If this measure is related to measure(s) already endorsed by NQF (e.g., same topic, but different target population/setting/data source <u>or</u> different topic but same target population):</p> <p>3b.2 Are the measure specifications harmonized? If not, why? No related measures found.</p>	<p>3b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/> <input type="checkbox"/></p>
<p>3c. Distinctive or Additive Value</p> <p>3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: No related measures found.</p> <p>5.1 If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), Describe why it is a more valid or efficient way to measure quality: No related measures found.</p>	<p>3c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/> <input type="checkbox"/></p>
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Usability?</p>	<p>3</p>
<p>Steering Committee: Overall, to what extent was the criterion, Usability, met?</p>	<p>3</p>

Rationale:	C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rati ng
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 How are the data elements that are needed to compute measure scores generated? Coding/abstraction performed by someone other than person obtaining original information (E.g., DRG, ICD-9 codes on claims, chart abstraction for quality measure or registry)	4a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4b. Electronic Sources 4b.1 Are all the data elements available electronically? (<i>elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims</i>) Yes 4b.2 If not, specify the near-term path to achieve electronic capture by most providers.	4b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4c. Exclusions 4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications? No 4c.2 If yes, provide justification.	4c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences 4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results. Coding professionals follow detail guidelines, are subject to training and credentialing requirements, peer review and audit. Perforated appendix rates vary systematically by race, although the cause is unknown. Areas with high rates of perforated appendix may want to target points of intervention by using chart reviews and other supplemental data to investigate the reasons for delay in receiving surgery. Hospital contributions to the overall area rate may be particularly useful for this indicator, because misdiagnoses and other delays in receiving surgery in an emergency room may contribute substantially to the rate.	4d C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4e. Data Collection Strategy/Implementation 4e.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/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: As a PQJ, admission for perforated appendix is not a measure of hospital quality, but rather one measure of outpatient and other health care. Relative to other indicators, a higher percentage of the variation occurs at the area level rather than the discharge level. However, the signal ratio (i.e., the proportion of the total variation across areas that is truly related to systematic differences in area performance rather than random variation) is low, at 26.5%, indicating that much of the observed differences in age-sex adjusted rates likely do not represent true differences across areas. Applying multivariate signal extraction methods can improve estimation of true differences in area performance.	4e C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>

<p>4e.2 Costs to implement the measure (<i>costs of data collection, fees associated with proprietary measures</i>): All data necessary to calculate this measure are routinely collected for hospital administrative purposes. The software for calculating the measure is available for free at: http://www.qualityindicators.ahrq.gov/software.htm</p> <p>4e.3 Evidence for costs: All data necessary to calculate this measure are routinely collected for hospital administrative purposes. The software for calculating the measure is available for free at: http://www.qualityindicators.ahrq.gov/software.htm</p> <p>4e.4 Business case documentation: All data necessary to calculate this measure are routinely collected for hospital administrative purposes. The software for calculating the measure is available for free at: http://www.qualityindicators.ahrq.gov/software.htm</p>	
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for Feasibility?</p>	<p>4</p>
<p>Steering Committee: Overall, to what extent was the criterion, <i>Feasibility</i>, met? Rationale:</p>	<p>4 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
<p style="text-align: center;">RECOMMENDATION</p>	
<p>(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.</p>	<p>Time - limit ed <input type="checkbox"/></p>
<p>Steering Committee: Do you recommend for endorsement? Comments:</p>	<p>Y <input type="checkbox"/> N <input type="checkbox"/> A <input type="checkbox"/></p>
<p style="text-align: center;">CONTACT INFORMATION</p>	
<p>Co.1 Measure Steward (Intellectual Property Owner)</p>	
<p>Co.1 Organization Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850</p>	
<p>Co.2 Point of Contact</p>	
<p>John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-</p>	
<p>Measure Developer If different from Measure Steward</p>	
<p>Co.3 Organization Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850</p>	
<p>Co.4 Point of Contact</p>	
<p>John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-</p>	
<p>Co.5 Submitter If different from Measure Steward POC</p>	
<p>John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-, Agency for Healthcare Research and Quality</p>	
<p>Co.6 Additional organizations that sponsored/participated in measure development</p>	
<p>UC Davis, Stanford University, Battelle Memorial Institute</p>	
<p style="text-align: center;">ADDITIONAL INFORMATION</p>	
<p>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.</p>	

None
Ad.2 If adapted, provide name of original measure: None Ad.3-5 If adapted, provide original specifications URL or attachment
Measure Developer/Steward Updates and Ongoing Maintenance Ad.6 Year the measure was first released: 2001 Ad.7 Month and Year of most recent revision: 10, 2010 Ad.8 What is your frequency for review/update of this measure? Annual Ad.9 When is the next scheduled review/update for this measure? 05, 2011
Ad.10 Copyright statement/disclaimers: The AHRQ QI software is publicly available; no copyright disclaimers
Ad.11 -13 Additional Information web page URL or attachment:
Date of Submission (MM/DD/YY): 04/05/2011