

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 #: 0357	NQF Project: Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION	
De.1 Measure Title: Abdominal Aortic Aneurysm (AAA) Repair Volume (IQI 4)	
De.2 Brief description of measure: Count of adult hospital discharges in a one-year time period with a procedure code of AAA repair.	
1.1-2 Type of Measure: Structure	
De.3 If included in a composite or paired with another measure, please identify composite or paired measure Abdominal Aortic Aneurysm (AAA) Repair Mortality Rate (IQI 11) (NQF #0359)	
De.4 National Priority Partners Priority Area: Safety	
De.5 IOM Quality Domain: Effectiveness, Safety	
De.6 Consumer Care Need: Getting better	

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 type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<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>	<p>B</p> <p>Y <input 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, Quality Improvement (Internal to the specific organization)	C Y <input 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 type="checkbox"/> N <input type="checkbox"/>
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (<i>if submission returned</i>):	Met Y <input type="checkbox"/> N <input type="checkbox"/>
Staff Notes to Reviewers (<i>issues or questions regarding any criteria</i>):	
Staff Reviewer Name(s):	

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. <i>Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.</i> (evaluation criteria) 1a. High Impact	Eval Rati ng
(for NQF staff use) Specific NPP goal :	
1a.1 Demonstrated High Impact Aspect of Healthcare: Patient/societal consequences of poor quality 1a.2 1a.3 Summary of Evidence of High Impact: Most studies published since 1985 showed a significant association between either hospital or surgeon volume and inpatient mortality after AAA repair, although these findings may be limited by inadequate risk adjustment of the outcome measure and differ by type of aneurysms (intact vs. ruptured) being considered. Several studies have explored whether experience on related, but not identical, cases may lead to improved outcomes. One study found that hospital volume of surgery for ruptured aneurysms was not associated with postoperative inpatient mortality, but it was associated with fewer inpatient deaths for ruptured aneurysms, suggesting that high-volume hospitals may manage ruptured aneurysms more aggressively. [1] One study that evaluated the impact of total vascular surgery volume found a significant effect for both ruptured and intact aneurysms. [2] Empirical evidence shows that AAA repair volume and mortality—after adjusting for age, sex, and APR-DRG—are independently and negatively correlated with each other (r=-.35, p<.001). [3] 1a.4 Citations for Evidence of High Impact: Updated citations will be presented in the May Steering Committee meeting [1] Kantonen I, Lepantalo M, Brommels M, et al. Mortality in ruptured abdominal aortic aneurysms. The Finnvasc Study Group. . Eur J Vasc Endovasc Surg 1999;17(3):208-12. [2] Amundsen S, Skjaerven R, Trippestad A, et al. Abdominal aortic aneurysms. Is there an association between surgical volume, surgical experience, hospital type and operative mortality? Members of the Norwegian Abdominal Aortic Aneurysm Trial. Acta Chir Scand 1990;156(4):323-7; discussion 327-8.	1a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>

[3] Nationwide Inpatient Sample.

1b. Opportunity for Improvement

1b.1 Benefits (improvements in quality) envisioned by use of this measure: Abdominal Aortic Aneurysm (AAA) repair is a relatively rare procedure that requires proficiency with the use of complex equipment; and technical errors may lead to clinically significant complications, such as arrhythmias, acute myocardial infarction, colonic ischemia, and death. Higher volumes have been associated with better outcomes, which represent better quality.

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

Comparative Data for the IQI based on the 2008 Nationwide Inpatient Sample (NIS):

SEX
7,795 Males
1,996 Females

AGE
12 18 to 39
1,574 40 to 64
3,618 65 to 74
4,587 75+

PAYER
7,377 Medicare
155 Medicaid
2,243 Other

Based on the above, we see AAAs are occurring nearly four times more frequently in males compared to females. We also observe the procedure occurs primarily with the Medicare population; age 65 years and older.

Information about NIS can be found at this AHRQ link: <http://www.hcup-us.ahrq.gov/nisoverview.jsp#Whatis>

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:

Comparative Data for the IQI based on the 2008 Nationwide Inpatient Sample (NIS):

SEX
7,795 Males
1,996 Females

AGE
12 18 to 39
1,574 40 to 64
3,618 65 to 74
4,587 75+

PAYER
7,377 Medicare
155 Medicaid
2,243 Other

Information about NIS can be found at this AHRQ link: <http://www.hcup-us.ahrq.gov/nisoverview.jsp#Whatis>

1b
C
P
M
N

RACE
 29,703 White
 1,350 Black
 949 Hispanic
 457 Asian and NH/PI
 240 Amer Indian/AN
 7,537 Other

Source: 2008 State Inpatient Databases (SID). <http://hcup-us.ahrq.gov/sidoverview.jsp>

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*): Abdominal Aortic Aneurysm (AAA) repair is a relatively rare procedure that requires proficiency with the use of complex equipment; and technical errors may lead to clinically significant complications, such as arrhythmias, acute myocardial infarction, colonic ischemia, and death. Higher volumes have been associated with better outcomes, which represent better quality.

1c.2-3. Type of Evidence: Evidence-based guideline, Expert opinion

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

Most studies published since 1985 showed a significant association between either hospital or surgeon volume and inpatient mortality after AAA repair, although these findings may be limited by inadequate risk adjustment of the outcome measure and differ by type of aneurysms (intact vs. ruptured) being considered. Several studies have explored whether experience on related, but not identical, cases may lead to improved outcomes. One study found that hospital volume of surgery for ruptured aneurysms was not associated with postoperative inpatient mortality, but it was associated with fewer inpatient deaths for ruptured aneurysms, suggesting that high-volume hospitals may manage ruptured aneurysms more aggressively. [1] One study that evaluated the impact of total vascular surgery volume found a significant effect for both ruptured and intact aneurysms. [2] Empirical evidence shows that AAA repair volume and mortality—after adjusting for age, sex, and APR-DRG—are independently and negatively correlated with each other ($r=-.35$, $p<.001$). [3]

[1] Kantonen I, Lepantalo M, Brommels M, et al. Mortality in ruptured abdominal aortic aneurysms. The Finnvasc Study Group. . *Eur J Vasc Endovasc Surg* 1999;17(3):208-12.

[2] Amundsen S, Skjaerven R, Trippestad A, et al. Abdominal aortic aneurysms. Is there an association between surgical volume, surgical experience, hospital type and operative mortality? Members of the Norwegian Abdominal Aortic Aneurysm Trial. *Acta Chir Scand* 1990;156(4):323-7; discussion 327-8.

[3] Nationwide Inpatient Sample.

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

B. 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

1c
 C
 P
 M
 N

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

<p>indicators. In addition, they analyzed correlation matrices for indicators.</p> <p>1c.7 Summary of Controversy/Contradictory Evidence: Some users have questioned the inclusion of both ruptured and unruptured AAA and open and endovascular procedures. However, the experience of repair procedures (open or endovascular) carries over to both types of classes of patients, and total volume was a better predictor of overall mortality than the individual volumes.</p> <p>1c.8 Citations for Evidence (other than guidelines): Updated citations will be presented in the May Steering Committee meeting</p> <p>Hannan EL, Kilburn H, Jr., O'Donnell JF, et al. A longitudinal analysis of the relationship between in-hospital mortality in New York state and the volume of abdominal aortic aneurysm surgeries performed. <i>Health Serv Res</i> 1992;27(4):517-42.</p> <p>Kazmers A, Jacobs L, Perkins A, et al. Abdominal aortic aneurysm repair in Veterans Affairs medical centers. <i>J Vasc Surg</i> 1996;23(2):191-200.</p> <p>Pronovost PJ, Jenckes MW, Dorman T, et al. Organizational characteristics of intensive care units related to outcomes of abdominal aortic surgery. <i>JAMA</i> 1999;281(14):1310-7.</p> <p>1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number): EVAR for AAA represents an advance in patient care, serving as an effective alternative to traditional open surgical AAA repair, and is now the most common treatment method for AAA repair in the United States.</p> <p>1c.10 Clinical Practice Guideline Citation: http://www.sirweb.org/clinical/cpg/QI12.pdf</p> <p>1c.11 National Guideline Clearinghouse or other URL: Not Applicable.</p> <p>1c.12 Rating of strength of recommendation (also provide narrative description of the rating and by whom): Not Applicable.</p> <p>1c.13 Method for rating strength of recommendation (If different from USPSTF system, also describe rating and how it relates to USPSTF): Not Applicable.</p> <p>1c.14 Rationale for using this guideline over others: Not Applicable.</p>	
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Importance to Measure and Report</i>?</p>	<p>1</p>
<p>Steering Committee: Was the threshold criterion, <i>Importance to Measure and Report</i>, met? Rationale:</p>	<p>1 Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p style="text-align: center;">2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>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)</p>	<p>Eval Rati ng</p>
<p style="text-align: center;">2a. MEASURE SPECIFICATIONS</p>	
<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- spe cs</p>
<p>2a.1 Numerator Statement (Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome): Discharges, age 18 years and older, with an abdominal aortic aneurysm (AAA) repair procedure and a principal or secondary diagnosis of AAA</p>	<p>C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>

2a.2 Numerator Time Window (*The time period in which cases are eligible for inclusion in the numerator*):
 Time window can be determined by user, but is generally a calendar year. Note the volume-outcome estimates are based on one year of data.

2a.3 Numerator Details (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*):

ICD-9-CM AAA procedure codes:

3834

AORTA RESECTION & ANAST

3844

RESECT ABDM AORTA W REPL

3864

EXCISION OF AORTA

3971

ENDO IMPLANT OF GRAFT IN AORTA

ICD-9-CM AAA diagnosis codes:

4413

RUPT ABD AORTIC ANEURYSM

4414

ABDOM AORTIC ANEURYSM

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*):

Not applicable

2a.5 Target population gender: Female, Male

2a.6 Target population age range: 18 and older

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*):

Not applicable

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):

Not applicable

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*):

The stratification of the denominator for open vs. endovascular and ruptured vs. unruptured involve the following codes in the denominator specification:

/* AAA Repair */

/* ICD-9-CM Procedure Codes: */

/* OPEN */;

3834 = 1 /* AORTA RESECTION & ANAST */

3844 = 1 /* RESECT ABDM AORTA W REPL */

3864 = 1 /* EXCISION OF AORTA */

/* ENDOVASCULAR */;

3971 = 1 /* ENDO IMPL GRFT ABD AORTA */

/* Include Only: AAA */

/* ICD-9-CM Diagnosis Codes: */

/* RUPTURED */;
 ^4413 ^ = ^1^ /* RUPT ABD AORTIC ANEURYSM */
 /* UNRUPTURED */;
 ^4414 ^ = ^1^ /* ABDOM AORTIC ANEURYSM */

2a.12-13 Risk Adjustment Type: No risk adjustment necessary

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method*):
 Not applicable

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

2a.18-19 Type of Score: Count
2a.20 Interpretation of Score: Better quality = Higher score
2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps*):
 The volume is the number of discharges with a diagnosis of, and a procedure for AAA. There are four volume strata: open vs. endovascular, and ruptured vs. un-ruptured.

2a.22 Describe the method for discriminating performance (*e.g., significance testing*):
 Performance discrimination is based on pre-defined thresholds derived from the literature. Threshold 1: 10 or more procedures per year Threshold 2: 32 or more procedures per year.

2a.23 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

2a.24 Data Source (*Check the source(s) for which the measure is specified and tested*)
 Administrative claims

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.

2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL None
<http://qualityindicators.ahrq.gov/software/default.aspx>

2a.29-31 Data dictionary/code table web page URL or attachment: URL None
<http://qualityindicators.ahrq.gov/Downloads/Software/WinQI/V43/AHRQ%20QI%20Software%20Instructions,%20WinQI.pdf>

2a.32-35 Level of Measurement/Analysis (*Check the level(s) for which the measure is specified and tested*)
 Facility

2a.36-37 Care Settings (*Check the setting(s) for which the measure is specified and tested*)
 Hospital/Acute Care Facility

2a.38-41 Clinical Services (*Healthcare services being measured, check all that apply*)
 Clinicians: Physicians (MD/DO)

TESTING/ANALYSIS

2b. Reliability testing

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

2b.2 Analytic Method (*type of reliability & rationale, method for testing*):
 Literature summary, expert panels and empirical analysis

2b
 C
 P
 M
 N

<p>2b.3 Testing Results (<i>reliability statistics, assessment of adequacy in the context of norms for the test conducted</i>): AAA repair is an uncommon cardiovascular procedure—only 50,000 were performed in the United States in 2007. Although AAA repair is measured accurately with discharge data, the relatively small number of procedures performed annually at most hospitals suggests that volume may be subject to much random variation.</p>																																																		
<p>2c. Validity testing</p> <p>2c.1 Data/sample (<i>description of data/sample and size</i>): AHRQ 2007 State Inpatient Databases (SID) with 4,000 hospitals and 30 million adult discharges</p> <p>2c.2 Analytic Method (<i>type of validity & rationale, method for testing</i>): Literature summary, expert panels and empirical analysis</p> <p>2c.3 Testing Results (<i>statistical results, assessment of adequacy in the context of norms for the test conducted</i>): Most studies published since 1985 showed a significant association between either hospital or surgeon volume and inpatient mortality after AAA repair, although these findings may be limited by inadequate risk adjustment of the outcome measure and differ by type of aneurysms (intact vs. ruptured) being considered.</p> <p>Several studies have explored whether experience on related, but not identical, cases may lead to improved outcomes. One study found that hospital volume of surgery for ruptured aneurysms was not associated with postoperative inpatient mortality, but it was associated with fewer inpatient deaths for ruptured aneurysms, suggesting that high-volume hospitals may manage ruptured aneurysms more aggressively. [3] One study that evaluated the impact of total vascular surgery volume found a significant effect for both ruptured and intact aneurysms. [2] Empirical evidence shows that AAA repair volume and mortality—after adjusting for age, sex, and APR-DRG—are independently and negatively correlated with each other ($r = -.35$, $p < .001$). [3]</p> <p>References: [1] Kantonen I, Lepantalo M, Brommels M, et al. Mortality in ruptured abdominal aortic aneurysms. The Finnvasc Study Group. Eur J Vasc Endovasc Surg 1999;17(3):208-12. [2] Amundsen S, Skjaerven R, Trippestad A, et al. Abdominal aortic aneurysms. Is there an association between surgical volume, surgical experience, hospital type and operative mortality? Members of the Norwegian Abdominal Aortic Aneurysm Trial. Acta Chir Scand 1990;156(4):323-7; discussion 327-8. [3] Nationwide Inpatient Sample.</p> <p>Table 1. Reference Population Volume</p> <p>Column Headings: Column 1: Year Column 2: Open procedure, ruptured diagnosis Column 3: Open procedure, unruptured diagnosis Column 4: Endovascular procedure, ruptured diagnosis Column 5: Endovascular procedure, unruptured diagnosis Column 6: Original(Composite)</p> <table border="1"> <thead> <tr> <th>Volume</th> <th>2004</th> <th>2005</th> <th>2006</th> <th>2007</th> <th>2008</th> <th>%Change</th> </tr> </thead> <tbody> <tr> <td></td> <td>3,241</td> <td>2,876</td> <td>2,652</td> <td>2,445</td> <td>2,352</td> <td></td> </tr> <tr> <td></td> <td>15,723</td> <td>12,941</td> <td>11,152</td> <td>9,693</td> <td>8,851</td> <td>-32.1%</td> </tr> <tr> <td></td> <td>456</td> <td>568</td> <td>647</td> <td>799</td> <td>1,068</td> <td>-57.5%</td> </tr> <tr> <td></td> <td>17,438</td> <td>19,981</td> <td>22,778</td> <td>25,101</td> <td>28,103</td> <td>85.1%</td> </tr> <tr> <td></td> <td>36,768</td> <td>36,292</td> <td>37,156</td> <td>37,970</td> <td>40,293</td> <td>47.7%</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>9.2%</td> </tr> </tbody> </table> <p>Source: State Inpatient Databases (SID), 2008, Healthcare Cost and Utilization Project (HCUP)</p>	Volume	2004	2005	2006	2007	2008	%Change		3,241	2,876	2,652	2,445	2,352			15,723	12,941	11,152	9,693	8,851	-32.1%		456	568	647	799	1,068	-57.5%		17,438	19,981	22,778	25,101	28,103	85.1%		36,768	36,292	37,156	37,970	40,293	47.7%							9.2%	<p>2c</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>
Volume	2004	2005	2006	2007	2008	%Change																																												
	3,241	2,876	2,652	2,445	2,352																																													
	15,723	12,941	11,152	9,693	8,851	-32.1%																																												
	456	568	647	799	1,068	-57.5%																																												
	17,438	19,981	22,778	25,101	28,103	85.1%																																												
	36,768	36,292	37,156	37,970	40,293	47.7%																																												
						9.2%																																												
<p>2d. Exclusions Justified</p> <p>2d.1 Summary of Evidence supporting exclusion(s):</p>	<p>2d</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p>																																																	

<p>Not applicable</p> <p>2d.2 Citations for Evidence: Not applicable</p> <p>2d.3 Data/sample (description of data/sample and size): Not applicable</p> <p>2d.4 Analytic Method (type analysis & rationale): Not applicable</p> <p>2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses): Not applicable</p>	<p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p> <p><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): Not applicable</p> <p>2e.2 Analytic Method (type of risk adjustment, analysis, & rationale): Not applicable</p> <p>2e.3 Testing Results (risk model performance metrics): Not applicable</p> <p>2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: Volume</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><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): Predefined thresholds based on the literature</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> <table border="0" data-bbox="107 1255 454 1312"> <tr> <td>Q1</td> <td>Q2</td> <td>Q3</td> <td>Q4</td> </tr> <tr> <td>1.9</td> <td>5.6</td> <td>13.8</td> <td>47.3</td> </tr> </table> <p>N = 1,963</p>	Q1	Q2	Q3	Q4	1.9	5.6	13.8	47.3	<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>
Q1	Q2	Q3	Q4						
1.9	5.6	13.8	47.3						
<p>2g. Comparability of Multiple Data Sources/Methods</p> <p>2g.1 Data/sample (description of data/sample and size): Not applicable</p> <p>2g.2 Analytic Method (type of analysis & rationale): Not applicable</p> <p>2g.3 Testing Results (e.g., correlation statistics, comparison of rankings): Not applicable</p>	<p>2g</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><input type="checkbox"/></p>								
<p>2h. Disparities in Care</p> <p>2h.1 If measure is stratified, provide stratified results (scores by stratified categories/cohorts): Not applicable</p> <p>2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans: Not applicable</p>	<p>2h</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><input type="checkbox"/></p>								

<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Scientific Acceptability of Measure Properties</i>?</p>	<p>2</p>
<p>Steering Committee: Overall, to what extent was the criterion, <i>Scientific Acceptability of Measure Properties</i>, met? Rationale:</p>	<p>2 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
<p>3. USABILITY</p>	
<p>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)</p>	<p>Eval Rati ng</p>
<p>3a. Meaningful, Understandable, and Useful Information</p>	
<p>3a.1 Current Use: In use</p>	
<p>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):</p>	
<p>California (state) Hospital Volume and Utilization Indicators for California http://www.oshpd.ca.gov/HID/Products/PatDischargeData/ResearchReports/HospIPQualInd/Vol-Util_IndicatorsRpt/index.html</p>	
<p>Colorado (state hospital association) Colorado Hospital Report Card http://www.cohospitalquality.org/index.php?option=com_frontpage&Itemid=1</p>	
<p>Illinois (state hospital association) Illinois Hospitals Caring for You www.illinoishospitals.org</p>	
<p>Kentucky (Norton Healthcare, a hospital system) Norton Healthcare Quality Report http://www.nortonhealthcare.com/body.cfm?id=157</p>	
<p>New Jersey (state) Find and Compare Quality Care in NJ Hospitals http://www.nj.gov/health/healthcarequality/</p>	
<p>New York (health care coalition) New York State Hospital Report Card http://www.myhealthfinder.com/</p>	
<p>Oregon (state) Oregon Hospital Quality Indicators http://www.oregon.gov/OHPPR/HQ/</p>	
<p>Texas (state) Reports on Hospital Performance http://www.dshs.state.tx.us/thcic/</p> <p>Vermont (state) Dept of Banking, Insurance, Securities & Health Care Administration Comparison Report http://www.bishca.state.vt.us/health-care/hospitals-health-care-practitioners/2009-vermont-hospital-</p>	
<p>3a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>	

report-card

Washington (health care coalition)
Washington State Hospital Report Card
<http://www.myhealthfinder.com/wa09/index.php>

The measure is also reported on HCUPnet:
http://hcupnet.ahrq.gov/HCUPnet.jsp?Id=EB57801381F71C41&Form=MAINSEL&JS=Y&Action=%3E%3ENext%3E%3E&_MAINSEL=AHRO%20Quality%20Indicators

This measure is used in the MONAHRQ system that is provided for public reporting and quality improvement throughout the United States: <http://monahrq.ahrq.gov/>

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). If not used for QI, state the plans to achieve use for QI within 3 years:*)

University Healthcare Consortium - An alliance of 103 academic medical centers and 219 of their affiliated hospitals. Reporting the AHRQ QIs to their member hospitals. (see www.uhc.edu. Note: measure results reported to hospitals; not reported on site).

Dallas Fort Worth Hospital Council - Reporting on measure results to over 70 hospitals in Texas (see www.dfwhc.org. Note: measure results reported to hospitals; not reported on site).

Norton Healthcare - a multi-hospital system in Kentucky (see http://www.nortonhealthcare.com/about/Our_Performance/index.aspx)
Ministry Health Care - a multi-hospital system in Wisconsin (see <http://ministryhealth.org/display/router.aspx>. Note: measure results reported to hospitals; not reported on site).

Minnesota Hospital Association
<http://www.mnhospitals.org/> Note: measure used in quality improvement. Not reported publicly by the association).

This measure is used in the MONAHRQ system that is provided for public reporting and quality improvement throughout the United States: <http://monahrq.ahrq.gov/>

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*)

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

3a.5 Methods (*e.g., focus group, survey, QI project*):

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

<p>of education</p> <p>3a.6 Results (qualitative and/or quantitative results and conclusions): Given the above review of the literature and original research that was conducted, a Model report was the result that could help sponsors use the best evidence on public reports so they are most likely to have the desired effects on quality</p>	
<p>3b/3c. Relation to other NQF-endorsed measures</p> <p>3b.1 NQF # and Title of similar or related measures:</p>	
<p>(for NQF staff use) Notes on similar/related <u>endorsed</u> or submitted measures:</p>	
<p>3b. Harmonization If this measure is related to measure(s) already <u>endorsed by NQF</u> (e.g., same topic, but different target population/setting/data source <u>or</u> different topic but same target population): 3b.2 Are the measure specifications harmonized? If not, why? Leapfrog measure specification is based on the AHRQ QI, but is not reported separately</p>	<p>3b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/></p>
<p>3c. Distinctive or Additive Value 3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: The AHRQ QI measure is paried with a risk-adjusted mortality measure</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: The AHRQ QI measure is paried with a risk-adjusted mortality measure</p>	<p>3c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/></p>
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Usability</i>?</p>	<p>3</p>
<p>Steering Committee: Overall, to what extent was the criterion, <i>Usability</i>, met? Rationale:</p>	<p>3 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
4. FEASIBILITY	
<p>Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)</p>	<p>Eval Rati ng</p>
<p>4a. Data Generated as a Byproduct of Care Processes</p> <p>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)</p>	<p>4a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
<p>4b. Electronic Sources</p> <p>4b.1 Are all the data elements available electronically? (elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims) Yes</p> <p>4b.2 If not, specify the near-term path to achieve electronic capture by most providers.</p>	<p>4b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
<p>4c. Exclusions</p> <p>4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?</p>	<p>4c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/></p>

<p>No</p> <p>4c.2 If yes, provide justification.</p>	<p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>
<p>4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences</p> <p>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.</p> <p>AAA repair volume is measured with great precision, although volume indicators overall are not direct measures of quality and are relatively insensitive. For this reason, this indicator should be used in conjunction with other measures of mortality to ensure that increasing volumes truly improve patient outcomes. The volume-outcome relationship on which this indicator is based may not hold over time, as providers become more experienced or as technology changes.</p>	<p>4d</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>4e. Data Collection Strategy/Implementation</p> <p>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: Low-volume providers may attempt to increase their volume without improving quality of care by performing the procedure on patients who may not qualify or benefit. Additionally, shifting procedures to high-volume providers may impair access to care for certain types of patients.</p> <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://qualityindicators.ahrq.gov/software/default.aspx</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://qualityindicators.ahrq.gov/software/default.aspx</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://qualityindicators.ahrq.gov/software/default.aspx</p>	<p>4e</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>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Feasibility</i>?</p>	<p>4</p>
<p>Steering Committee: Overall, to what extent was the criterion, <i>Feasibility</i>, met? Rationale:</p>	<p>4</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>RECOMMENDATION</p>	
<p>(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.</p>	<p>Time - limit ed</p> <p><input type="checkbox"/></p>
<p>Steering Committee: Do you recommend for endorsement? Comments:</p>	<p>Y <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>A <input type="checkbox"/></p>
<p>CONTACT INFORMATION</p>	

<p>Co.1 Measure Steward (Intellectual Property Owner) Co.1 Organization Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850</p> <p>Co.2 Point of Contact John, Bott, Contractor, AHRQ Quality Indicators Measure Expert Center for Delivery, Organization and Markets, John.Bott@ahrq.hhs.gov, 301-427-1317-</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 John, Bott, MSSW, MBA, John.Bott@AHRQ.hhs.gov, 301-427-1317-</p>
<p>Co.5 Submitter If different from Measure Steward POC 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 UC Davis, Stanford University, Battelle Memorial Institute</p>
<p>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. None</p>
<p>Ad.2 If adapted, provide name of original measure: None Ad.3-5 If adapted, provide original specifications URL or attachment</p>
<p>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: 08, 2011 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? 12, 2011</p>
<p>Ad.10 Copyright statement: The AHRQ QI software is publicly available; no copyright disclaimers</p>
<p>Ad.11 Disclaimers: None</p>
<p>Ad.12 -14 Additional Information web page URL or attachment:</p>
<p>Date of Submission (MM/DD/YY): 02/01/2011</p>