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 #:** 0360 **NQF Project:** Surgery Endorsement Maintenance 2010
MEASURE DESCRIPTIVE INFORMATION
De.1 **Measure Title:** Esophageal Resection Mortality Rate (IQI 8)
De.2 **Brief description of measure:** Number of inpatient deaths per 100 discharges with a procedure for esophageal resection
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** Esophageal resection volume (IQI 1)
De.4 **National Priority Partners Priority Area:** Population health, Safety De.5 **IOM Quality Domain:** Effectiveness 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**
A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. *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.* 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 A.2 **Indicate if Proprietary Measure** (*as defined in measure steward agreement*)**:** A.3 **Measure Steward Agreement:** Government entity and in the public domain - no agreement necessary A.4 **Measure Steward Agreement attached:**	**A** Y☐ N☐
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	**B** Y☐

every 3 years. Yes, information provided in contact section	N☐
C. The intended use of the measure includes both public reporting and quality improvement. ►**Purpose:** Public reporting, Internal quality improvement Accountability, Payment incentive	**C** Y☐ N☐
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☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** **Staff Notes to Steward** (*if submission returned*)**:**	**Met** Y☐ N☐
Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**	
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. ***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 Rating**
(for NQF staff use) **Specific NPP goal**:	
1a.1 **Demonstrated High Impact Aspect of Healthcare:** Severity of illness, Patient/societal consequences of poor quality 1a.2 1a.3 **Summary of Evidence of High Impact:** Esophageal resection is a complex cancer surgery, and studies have noted that providers with higher volumes have lower mortality rates. This suggests that providers with higher rates have some characteristics, either structurally or with regard to processes, that influence mortality. 1a.4 **Citations for Evidence of High Impact:** Patti MG, Corvera CU, Glasgow RE, et al. A hospital´s annual rate of esophagectomy influences the operative mortality rate. J Gastrointest Surg 1998;2(2):186-92. Gordon TA, Bowman HM, Bass EB, et al. Complex gastrointestinal surgery: impact of provider experience on clinical and economic outcomes. J Am Coll Surg 1999;189(1):46-56. Dimick JB, Cowan JA, Jr., Ailawadi G, et al. National variation in operative mortality rates for esophageal resection and the need for quality improvement; 2003. Finlayson EV, Goodney PP, Birkmeyer JD. Hospital volume and operative mortality in cancer surgery: a national study. Arch Surg 2003;138(7):721-5; discussion 726.	**1a** C☐ P☐ M☐ N☐
1b. Opportunity for Improvement	**1b**

1b.1 Benefits (improvements in quality) envisioned by use of this measure: Providers can adopt the processes of care of the best performing providers or consumers can select the best performing providers in order to reduce the overall mortality rate

1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers:
5th 25th Median 75th 95th
0.017203 0.037254 0.058397 0.086440 0.140230

1b.3 Citations for data on performance gap:
2007 AHRQ State Inpatient Databases (SID) with 465 hospitals and 1,587 discharges

1b.4 Summary of Data on disparities by population group:
Based on the 2008 national statistics for esophageal resection mortality (http://hcupnet.ahrq.gov) the 2008 rates are as follows:

Overall rate per 100: 5.35 ; Risk adjusted rate: 6.59
Male: 5.75
Female: Too few reported to calculate reliable rates.

Ages 18 to 39: Too few reported to calculate reliable rates.
Ages 40 to 64: 3.15
Ages 65 to 74: 6.38
Ages 75+: 10.17

1b.5 Citations for data on Disparities:
AHRQ 2008 Nationwide Inpatient Sample

C☐ P☐ M☐ N☐

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*)**:** In-hospital death is directly related to the patient experience of care

1c.2-3. Type of Evidence: 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*)**:**
Face validity. Esophageal resection is a complex procedure that requires technical skill. The primary evidence for this indicator arises from the volume-outcome literature. Several studies have found that hospitals that perform more procedures have better mortality rates than lower volume hospitals. The magnitude of this relationship is relatively large as compared to other procedures. A full review of this literature can be found in the discussion of esophageal resection as a volume indicator. This relationship suggests that there may be some differences in processes of care that result in better outcomes. Those processes have not been identified and are subject to controversy, as it is unclear what the causal relationship is, if there truly is one, between hospital volume and mortality.

Precision. Esophageal resection is a relatively uncommon procedure, which may impact the precision of the indicator. Patti et al1 noted that most hospitals perform 10 or fewer procedures during a 5-year period. Utilizing several years of data, which has been done in some of the volume-outcome research, may help improve the precision of this indicator.

Minimum bias. Although we located no studies specifically addressing the need for risk adjustment, most of the volume-outcome studies published have used some sort of risk adjustment, suggesting that risk adjustment may be important for this procedure. Most of those studies used administrative data for risk adjustment.

Practice patterns may influence mortality rate. One such factor is case selection and the practice of "opening and closing" complex cases. Pye at al. identified all patients with oesophagogastric malignancy over

1c
C☐ P☐ M☐ N☐

one year in Wales and showed that 30-day mortality was higher when surgeons operated on more than 70% of their patients. The significant difference in survival when more than 70% of patients were treated surgically compared with less than 70% (18% versus 5%), in conjunction with low overall anastomotic leak rate of 5%, strongly suggests that case selection is a major factor. In this study, the "open and close" rate was 23%, emphasizing the potential importance of preoperative case selection. 2

In addition, patient characteristics have been shown to influence mortality. Some of these patient characteristics can be captured using administrative data. For example, age, urgent or emergent admission, non-white race, and selected comorbidities (e.g., renal disease) have been identified as significant risk factors for in-hospital mortality.4,11 Bias due to these factors can be removed through risk-adjustment using administrative data.

Only a few studies have evaluated potential risk factors that are not available from administrative data. Griffin et al. showed that active smoking, forced vital capacity and forced expiratory volume prior to surgery were associated with severe postoperative pulmonary complications in 228 patients undergoing esophagectomy.3 However, their data base was too small to show whether these factors were also associated with mortality.

One study examined 995 patients undergoing esophagesctomy in 24 hospitals in the United Kingdom. In the analyses, they identified some significant risk factors, including cancer staging, surgeon assessment of disease severity, and score on a standardized physiological assessment (Physiological and Operative Scoring System for enumeration of Morbidity and Mortality (POSSUM). After adjusting for these risk factors, annual hospital volume was still significantly associated with in-hospital mortality, which might be due to some quality effects remaining even after adjusting for other variables.5

As expected, complications following surgery also affect mortality. In a chart review from one tertiary hospital in Texas, all esophagogastrectomy (EG) cases from 1996 to 2002 were examined in relation to in-hospital mortality. Pneumonia was associated with a 20% incidence of death. Patients with pneumonia had significantly worse deglutition and anastomotic integrity on barium esophagogram compared with patients without pneumonia ($p < 0.001$).6

Construct validity. The extensive evidence regarding the association between hospital volume and mortality, summarized elsewhere, supports the construct validity of mortality as an indicator of hospital quality. Patti et al.1 used five volume categories, and found decreasing mortality rates of 17%, 19%, 10%, 16%, and 6% (with volumes of 1-5, 6-10, 11-20, 21-30, and >30 procedures during the 5-year study period). Gordon et al.7 combined all complex gastrointestinal procedures, and found that low volume (11-20 procedures per year) hospitals had an adjusted odds of death of 4.0 as compared to the single high volume hospital. In the most prominent study of the volume-outcome association, Birkmeyer et al used Medicare data from 1994 through 1999 to estimate volume-outcome relationships, imputing total annual hospital volume and adjusting for age, sex, race, year of the procedure, urgency of admission, mean income from Social Security at the ZIP Code level, and coexisting conditions from the index admission and other admissions within the preceding six months (summarized as the Charlson Comorbidity Index). They found that crude mortality rates were 23.1, 18.9, 16.9, 11.7, and 8.1 percent in very low (<2 imputed cases/year), low (2-4), medium (5-7), high (8-19) and very high (>19) volume hospital groups, respectively. Unadjusted and adjusted odds ratios were 0.78 and 0.85, 0.68 and 0.76, 0.44 and 0.51, and 0.29 and 0.36 in low, medium, high and very high volume hospitals, respectively, relative to very low volume hospitals.10 Similar findings (e.g., 2.6 to 2.9-fold variation in adjusted mortality across hospital volume strata) have been reported from studies based on the Nationwide Inpatient Sample, which is designed as a 20% random sample of all hospital discharges in the US.11,12 This association was confirmed in the Netherlands, where hospital mortality was reported as 12.1, 7.5%, and 4.9% at low (1-10 cases/year), medium (11-20), and high (>50) volume centers, respectively.8 A weaker but still significant effect was observed in Ontario, with an adjusted odds ratio of 1.9 at the lowest volume hospitals (mean 2.8 cases/year) relative to the highest volume hospitals (mean 19.0 cases/year).17 The association between hospital volume and mortality also persisted after adjustment for physiologic predictors in one study from the UK.5. Dimick showed that the association between volume and mortality may be mediated by complications such as renal failure, pulmonary failure, septicemia, reintubation and aspiration.9

Dimick also found a significant decline in hospital mortality after esophagectomy in the U.S. from 1988 to 2000 (13.6% to 10.5%, P=0.001). Low volume hospitals had markedly higher mortality rates and showed no improvement over time (15.3% vs 14.5%). In contrast, high volume hospitals experienced a significant reduction in mortality over time (11.0% vs 7.5%, $p = 0.003$). Referral patterns changed over time with the proportion of esophageal resections performed at high volume hospitals increasing from 40% (1988 to 1991) to 57% (1997 to 2000).13

Beyond hospital volume, recent studies have examined other hospital characteristics and their relation to

mortality. Dimick et al. looked at hospital teaching status and found that in analyses adjusted only for patient characteristics, esophageal resection mortality was lower at teaching hospitals than at nonteaching hospitals (OR=1.8, 95% CI 1.1-3.2). However, after adjusting for hospital volume, teaching status was no longer an independent predictor of mortality (OR=1.4, 95% CI 0.7-2.6).14 In a study of 366 patients with esophageal resection, no significant association between the nighttime nurse-to-patient ratio (NNPR) and in-hospital mortality was seen. However, a nurse typically caring for more than two ICU patients at night significantly increased the risk of postoperative pneumonia, reintubation, and septicemia.15
Patients treated at the 51 National Cancer Institute (NCI) cancer centers were compared with patients from 51 control hospitals with the highest volume for esophagectomy. NCI cancer centers had lower adjusted surgical mortality rates than control hospitals for esophagectomy (7.9% vs. 10.9%; P = 0.027).18 Taken together, these findings suggest that risk-adjusted mortality rates may capture other aspects of hospital quality, beyond what volume alone would capture.
Surgeons' training and experience have also been examined as predictors of mortality. Using the national Medicare population in 1998-1999, mortality rates were 37% (odds ratio, 1.37; 95% confidence interval, 1.02 to 1.82) higher for surgeons without specialty training compared with thoracic surgeons (adjusted mortality 16.5% versus 12.4%; p = 0.01). However, differences in mortality between high-volume and low-volume hospitals (24.3% versus 11.4%; p < 0.001) and surgeons (20.7% versus 10.7%; p < 0.001) were larger than those between thoracic and general surgeons.19 Also using Medicare data, Birkmeyer et al. showed that surgeon volume is a strong independent risk factor for esophagectomy mortality (e.g., 18.8% for surgeons with <2 imputed cases/year versus 9.2% for surgeons with >6 imputed cases/year), even after adjusting for hospital volume. For example, even at high-volume hospitals (>13 imputed cases/year), adjusted mortality was 17.2%, 9.8%, and 8.0% for low, medium, and high-volume surgeons.20
Finally, according to a recent meta-analysis of 50 articles comparing surgical techniques for esophageal resection, in-hospital mortality was significantly higher after transthoracic esophageal resection than after transhiatal resection (9.2% versus 5.7%, RR=1.60, 95% CI 1.35-1.89). However, the 3 randomized controlled trials included in that meta-analysis did not support this overall finding (although they collectively included only 106 patients), and the benefits of transhiatal resection disappeared in analyses of 3-year and 5-year survival. Therefore, it is unclear whether hospitals and surgeons can improve their overall outcomes by changing their preferred surgical approach.21

Fosters true quality improvement. Though we found no evidence on whether or not this indicator would stimulate true improvement in quality, it is possible that high risk patients may be denied surgery. This hypothesized effect has not been empirically evaluated or demonstrated. One study found no evidence of discrimination against racial/ethnic minorities or Medicaid or uninsured patients in terms of the odds of receiving esophageal resection at low or high volume (relative to medium volume) hospitals.28

Prior use. This indicator has been utilized in the National Healthcare Quality Report16 and is currently included in the AHRQ Inpatient Quality Indicator set.

1c.5 Rating of strength/quality of evidence (*also provide narrative description of the rating and by whom***):** Not applicable

1c.6 Method for rating evidence: Not applicable

1c.7 Summary of Controversy/Contradictory Evidence: None

1c.8 Citations for Evidence (*other than guidelines***):**
1.Patti MG, Corvera CU, Glasgow RE, et al. A hospital´s annual rate of esophagectomy influences the operative mortality rate. J Gastrointest Surg 1998;2(2):186-92.
2. Pye JK, Crumplin MK, Charles J, et al. One-year survey of carcinoma of the oesophagus and stomach in Wales. In: Br J Surg; 2001. p. 278-85.
3. Griffin SM, Shaw IH, Dresner SM. Early complications after Ivor Lewis subtotal esophagectomy with two-field lymphadenectomy: risk factors and management. In: J Am Coll Surg; 2002. p. 285-97.
4. Dimick JB, Cattaneo SM, Lipsett PA, et al. Hospital volume is related to clinical and economic outcomes of esophageal resection in Maryland. In: Ann Thorac Surg; 2001. p. 334-9; discussion 339-41.
5. McCulloch P, Ward J, Tekkis PP. Mortality and morbidity in gastro-oesophageal cancer surgery: initial results of ASCOT multicentre prospective cohort study. In: Bmj; 2003 Nov 22; 2003. p. 1192-7.
6. Atkins BZ, Shah AS, Hutcheson KA, et al. Reducing hospital morbidity and mortality following esophagectomy. In: Ann Thorac Surg; 2004. p. 1170-6; discussion 1170-6.

7. Gordon TA, Bowman HM, Bass EB, et al. Complex gastrointestinal surgery: impact of provider experience on clinical and economic outcomes. J Am Coll Surg 1999;189(1):46-56.
8. van Lanschot JJ, Hulscher JB, Buskens CJ, et al. Hospital volume and hospital mortality for esophagectomy; 2001.
9. Dimick JB, Pronovost PJ, Cowan JA, et al. Surgical volume and quality of care for esophageal resection: do high-volume hospitals have fewer complications? Ann Thorac Surg 2003 Feb;Sect. 337-41.
10. Birkmeyer JD, Siewers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. In: N Engl J Med; 2002. p. 1128-37.
11. Dimick JB, Cowan JA, Jr., Ailawadi G, et al. National variation in operative mortality rates for esophageal resection and the need for quality improvement; 2003.
12. Finlayson EV, Goodney PP, Birkmeyer JD. Hospital volume and operative mortality in cancer surgery: a national study. Arch Surg 2003;138(7):721-5; discussion 726.
13. Dimick JB, Wainess RM, Upchurch GR, Jr., et al. National trends in outcomes for esophageal resection. In: Ann Thorac Surg; 2005. p. 212-6; discussion 217-8.
14. Dimick JB, Cowan JA, Jr., Colletti LM, et al., inventors; Hospital teaching status and outcomes of complex surgical procedures in the United States. 2004 Feb.
15. Amaravadi RK, Dimick JB, Pronovost PJ, et al. ICU nurse-to-patient ratio is associated with complications and resource use after esophagectomy; 2000.
16. National Healthcare Quality Report. In: Agency for Healthcare Research and Quality; 2003.
17. Urbach DR, Bell CM, Austin PC. Differences in operative mortality between high- and low-volume hospitals in Ontario for 5 major surgical procedures: estimating the number of lives potentially saved through regionalization; 2003.
18. Birkmeyer NJ, Goodney PP, Stukel TA, et al. Do cancer centers designated by the National Cancer Institute have better surgical outcomes? In: Cancer; 2005. p. 435-41.
19. Dimick JB, Goodney PP, Orringer MB, Birkmeyer JD. Specialty training and mortality after esophageal cancer resection. Ann Thorac Surg. 2005;80:282-6.
20. Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. 2003;349:2117-27.
21. Huylscher JBF, Tijssen JGP, Obertop H, van Lanschot JJB. Transthoracic versus transhiatal resection for carcinoma of the esophagus: A meta-analysis. Ann Thorac Surg 2001;72:306-13.
28. Liu JH, Zingmond DS, McGory ML, et al. Disparities in the utilization of high-volume hospitals for complex surgery. In: Jama; 2006. p. 1973-80.

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)*:
During the comprehensive medical literature review, preference was given to high quality systematic reviews, meta-analyses, and clinical trials over the past ten years, plus existing nationally recognized treatment guidelines from the leading specialty societies.

1c.14 Rationale for using this guideline over others:
None

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, met? **Rationale:**	**1** **Y☐** **N☐**

2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)	Eval Rating
2a. MEASURE SPECIFICATIONS	
S.1 Do you have a web page where current detailed measure specifications can be obtained? **S.2 If yes, provide web page URL:** **2a. Precisely Specified**	
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*)**:** Number of deaths among cases meeting the inclusion and exclusion rules for the denominator. **2a.2 Numerator Time Window** (*The time period in which cases are eligible for inclusion in the numerator*)**:** Inpatient admission **2a.3 Numerator Details** (*All information required to collect/calculate the numerator, including all codes, logic, and definitions*)**:** Discharge disposition of death (DISP=20)	
2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*)**:** Discharges, age 18 years and older, with ICD-9-CM esophageal resection procedure code and a diagnosis code of esophageal cancer in any field OR gastrectomy procedure code ONLY if accompanied by selected diagnosis codes. **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*)**:** User defined; usually a calendar year **2a.8 Denominator Details** (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*)**:** ICD-9-CM esophageal resection procedure codes: 424 ESOPHAGECTOMY 4240 ESOPHAGECTOMY NOS 4241 PARTIAL ESOPHAGECTOMY 4242 TOTAL ESOPHAGECTOMY 425 THORAC ESOPHAG ANAST 4251 THORAC ESOPHAGOESOPHAGOS 4252 THORAC ESOPHAGOGASTROST 4253 THORAC SM BOWEL INTERPOS 4254 THORAC ESOPHAGOENTER NEC 4255 THORAC LG BOWEL INTERPOS 4256 THORAC ESOPHAGOCOLOS NEC 4258 THORAC INTERPOSITION NEC 4259 THORAC ESOPHAG ANAST NEC 426 STERN ESOPHAG ANAST 4261 STERN ESOPHAGOESOPHAGOST 4262 STERN ESOPHAGOGASTROSTOM 4263 STERN SM BOWEL INTERPOS 4264 STERN ESOPHAGOENTER NEC 4265 STERN LG BOWEL INTERPOS	**2a-specs** C☐ P☐ M☐ N☐

4266 STERN ESOPHAGOCOLOS NEC
4268 STERN INTERPOSITION NEC
4269 STERN ESOPHAG ANAST NEC
ONLY if selected diagnosis codes:
esophageal cancer (see below)
gastrointestinal-related cancer (see below)

OR:

ICD-9-CM gastrectomy procedure code:
4399 OTHER TOTAL GASTRECTOMY -

ONLY if selected diagnosis codes:
esophageal cancer (see below)

Esophageal cancer:
1500 MALIGNANT NEOPLASM OF ESOPHAGUS, CERVICAL
1501 MALIGNANT NEOPLASM OF ESOPHAGUS, THORACIC
1502 MALIGNANT NEOPLASM OF ESOPHAGUS, ABDOMINAL
1503 MALIGNANT NEOPLASM OF ESOPHAGUS, UPPER THIRD OF
1504 MALIGNANT NEOPLASM OF ESOPHAGUS, MIDDLE THIRD OF
1505 MALIGNANT NEOPLASM OF ESOPHAGUS, LOWER THIRD OF
1508 MALIGNANT NEOPLASM OF ESOPHAGUS, OTHER SPECIFIED PART
1509 MALIGNANT NEOPLASM OF ESOPHAGUS, UNSPECIFIED

Gastrointestinal cancer
1510 MALIGNANT NEOPLASM OF STOMACH, CARDIA
1978 SECONDARY MALIGNANT NEOPLASM OF RESPIRATORY AND DIGESTIVE SYSTEMS, OTHER DIGESTIVE ORGANS AND SPLEEN
2301 CARCINOMA IN SITU OF DIGESTIVE ORGANS, ESOPHAGUS
2355 NEOPLASM OF UNCERTAIN BEHAVIOR OF DIGESTIVE AND RESPIRATORY SYSTEMS, OTHER AND UNSPECIFIED DIGESTIVE ORGANS

2a.9 Denominator Exclusions (*Brief text description of exclusions from the target population***):** Exclude discharges with pregnancy, discharge to a short term hospital or missing information for discharge disposition, age or sex.

2a.10 Denominator Exclusion Details (*All information required to collect exclusions to the denominator, including all codes, logic, and definitions***):**
Exclude cases:
• missing discharge disposition (DISP=missing), gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing) or principal diagnosis (DX1 =missing)
• transferring to another short-term hospital (DISP=2)
• MDC 14 (pregnancy, childbirth, and puerperium)

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 age group, race/ethnicity categories, payer categories and sex.

2a.12-13 Risk Adjustment Type: Case-mix adjustment

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 GEE logistic regression and covariates for age (in 5-year age groups), APR-DRG and MDC. The reference population used in the regression is the universe of discharges for states that participate in the HCUP State Inpatient Databases (SID) for the year 2007, a database consisting of approximately 35 million discharges from 43 states. 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 or state). The risk adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate, multiplied by the reference population rate. The Smoothed Rate is the risk-

adjusted rate shrunken to the volume-specific rate and the prior year smoothed rate.
age 18-24; age 25-29; age 30-34; age 35-39; age 40-44; age 45-49; age 50-54; age 55-59; age 60-64 (omitted); age 65-69; age 70-74; age 75-79; age 80-84; age 85+
each age category*female
APRDRG 2201-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES (MINOR)
APRDRG 2202-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES (MODERATE) ADRG 2203-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES (MAJOR)
APRDRG 2204-MAJOR STOMACH, ESOPHAGEAL & DUODENAL PROCEDURES (EXTREME) ADRG 9999 (OTHER)

2a.15-17 Detailed risk model available Web page URL or attachment: URL
http://www.qualityindicators.ahrq.gov/downloads/iqi/IQI%20Risk%20Adjustment%20Tables%20(Version%204%202)%20wo%20APR-DRG.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 Inpatient Quality Indicator (IQI) expressed as a rate, is defined as outcome of interest/population at risk or numerator/denominator. The Quality Indicators software performs five steps to produce the IQI rates. 1) Discharge-level data is used to mark inpatient records containing outcomes of interest. 2) Identify populations at risk. For provider IQIs populations at risk are derived from hospital discharge records. 3) Calculate observed rates. Using output data from steps 1 and 2, IQI rates are calculated for user-specified combinations of stratifiers. 4) Risk adjust the IQI rates. Regression coefficients from a reference population database are applied to the observed rates in the risk-adjustment process. The risk-adjusted rates will then reflect the age and APR-DRG distribution of data in the reference population. 5) Create multivariate signal extraction (MSX) smoothed rates. Shrinkage factors are applied to the risk-adjusted rates for each IQI in the MSX process. For each IQI, the shrinkage estimate reflects a reliability adjustment unique to each indicator. Full information on IQI algorithms and specification can be found at http://qualityindicators.ahrq.gov/iqi_download.htm.

2a.22 Describe the method for discriminating performance *(e.g., significance testing)***:**
Significance testing is not prescribed by the software. Users may define their methods of discriminating performance according to their application. Although all cases are measured, the rate is considered a sample in time, given the variations in case mix over time. Confidence intervals can be calculated, but again are not prescribed.

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)*
Electronic administrative data/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.***):**
Hospital administrative discharge data. See data requirements in the AHRQ QI Windows Application Documentation: http://www.qualityindicators.ahrq.gov/software.htm

2a.26-28 Data source/data collection instrument reference web page URL or attachment: URL
http://www.qualityindicators.ahrq.gov/software.htm

2a.29-31 Data dictionary/code table web page URL or attachment: URL
http://www.qualityindicators.ahrq.gov/downloads/winqi/AHRQ_QI_Windows_Software_Documentation_V41a.pdf

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

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

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)

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:
Annul review of ICD-9-CM coding updates for denominator inclusion and exclusion criteria

2b.3 Testing Results *(reliability statistics, assessment of adequacy in the context of norms for the test conducted)*:
Not applicable

2b C☐ P☐ M☐ N☐

2c. Validity testing

2c.1 Data/sample *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID)

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:
Annual update of risk-adjustment models and comparative data

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:
Signal variance of 0.001518. Average signal ratio of 0.26.

2c C☐ P☐ M☐ N☐

2d. Exclusions Justified

2d.1 Summary of Evidence supporting exclusion(s):
The only exclusions are for missing data and transfer out to an acute care hospital

2d.2 Citations for Evidence:
Not applicable

2d.3 Data/sample *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID)

2d.4 Analytic Method *(type analysis & rationale)*:
Not applicable

2d.5 Testing Results *(e.g., frequency, variability, sensitivity analyses)*:
Not applicable

2d C☐ P☐ M☐ N☐ NA☐

2e. Risk Adjustment for Outcomes/ Resource Use Measures

2e.1 Data/sample *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID)

2e.2 Analytic Method *(type of risk adjustment, analysis, & rationale)*:
The predicted value for each case is computed using GEE logistic regression and covariates for age (in 5-year age groups), APR-DRG and MDC. The reference population used in the regression is the universe of discharges for states that participate in the HCUP State Inpatient Databases (SID) for the year 2007, a database consisting of approximately 35 million discharges from 43 states. 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 or state). The risk adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate, multiplied by the reference population rate. The Smoothed Rate is the risk-adjusted rate shrunken to the volume-specific rate and the prior year smoothed rate.

2e.3 Testing Results *(risk model performance metrics)*:
c-statistic of 0.766

2e C☐ P☐ M☐ N☐ NA☐

2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: Not applicable	
2f. Identification of Meaningful Differences in Performance **2f.1 Data/sample from Testing or Current Use** *(description of data/sample and size)*: AHRQ 2007 State Inpatient Databases (SID) **2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance** *(type of analysis & rationale)*: Posterior probability distribution (gamma); 95% probability interval **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)*: 5th 25th Median 75th 95th 0.017203 0.037254 0.058397 0.086440 0.140230 Discrimiation above or below the median of 3% of hosptials	**2f** C☐ P☐ M☐ N☐
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☐ P☐ M☐ N☐ NA ☐
2h. Disparities in Care **2h.1 If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)*: Based on the 2008 national statistics for esophageal resection mortality (http://hcupnet.ahrq.gov) the 2008 rates are as follows: Overall rate per 100: 5.35 ; Risk adjusted rate: 6.59 Male: 5.75 Female: Too few reported to calculate reliable rates. Ages 18 to 39: Too few reported to calculate reliable rates. Ages 40 to 64: 3.15 Ages 65 to 74: 6.38 Ages 75+: 10.17 **2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** Rates may be stratified by age, gender and payer categories	**2h** C☐ P☐ M☐ N☐ NA ☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties*, met? **Rationale:**	**2** C☐ P☐ M☐ N☐
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 Rating**
3a. Meaningful, Understandable, and Useful Information	**3a**

C☐ P☐ M☐ N☐

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: Hospital Inpatient Mortality Indicators for California, http://oshpd.ca.gov/HID/Products/PatDischargeData/AHRQ/iqi-imi_overview.html
2) State of Florida: Florida Health Finder, http://www.floridahealthfinder.gov/
3) Norton Healthcare (multi-hospital system): Norton Healthcare Quality Report, http://www.nortonhealthcare.com/body.cfm?id=157
4) State of Massachusetts: My HealthCare Options, http://www.mass.gov/healthcareqc
5) State of New Jersey: Find and Compare Quality Care in New Jersey Hospitals, http://www.nj.gov/health/healthcarequality/
6) Niagara Health Quality Coalition and Alliance for Quality Health Care: New York State Hospital Report Card, http://www.myhealthfinder.com/
7) State of Texas: Reports on Hospital Performance, http://www.dshs.state.tx.us/thcic/
8) Niagara Health Quality Coalition and Alliance for Quality Health Care: Washington State Hospital Report Card, http://www.myhealthfinder.com/wa09/index.php
9) State of Nevada: Nevada Compare Care, http://nevadacomparecare.net/Monahrq/home.html
10) State of Vermont: Department of Banking, Insurance, Securities & Health Care Administration (BISHCA) Comparison Report, http://www.bishca.state.vt.us/health-care/hospitals-health-care-practitioners/2009-vermont-hospital-report-card
11) Wisconsin Hospital Association: CheckPoint, http://www.wicheckpoint.org/index.aspx

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.ord. 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)

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)***:** The AHRQ State Inpatient Databases (SID) consist of approximatley 4,000 hospitals and 38 million 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). The AHRQ hip fracture mortality measure is included in the reports. These reports are designed specifically to report comparative information on hospital performance based on the AHRQ Quality Indicators (QIs). The work was done in close collaboration with AHRQ staff and the AHRQ Quality Indicators team.

The Model Reports (discussed immediately above) are based on: • Extensive search and analysis of the literature on hospital quality measurement and reporting, as well as public reporting on health care quality more broadly; • Interviews with quality measurement and reporting experts, purchasers, staff of purchasing coalitions, and executives of integrated health care delivery systems who are responsible for quality in their facilities; • Two focus groups with chief medical officers of hospitals and/or systems and two focus groups with quality managers from a broad mix of hospitals; • Four focus groups with members of the public who had recently experienced a hospital admission; and • Four rounds of cognitive interviews (a total of 62 interviews) to test draft versions of the two Model Reports with members of the public with recent hospital experience, basic computer literacy but widely varying levels of education. 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.	
3b/3c. Relation to other NQF-endorsed measures 3b.1 **NQF # and Title of similar or related measures:** Leapfrog esophagectomy survival predictor (NQF# Unknown)	
(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**	
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 or different topic but same target population): 3b.2 **Are the measure specifications harmonized? If not, why?** Yes; the Leapfrog specification is based on the AHRQ specification	**3b** C☐ P☐ M☐ N☐ NA ☐
3c. Distinctive or Additive Value 3c.1 **Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:** The AHRQ measure has improved discrimination and predictive properties; the AHRQ measure also has an associated measure of uncertainty. 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 measure has improved discrimination and predictive properties; the AHRQ measure also has an associated measure of uncertainty.	**3c** C☐ P☐ M☐ N☐ NA ☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability?*	**3**
Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**3** C☐ P☐ M☐ N☐
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 Rating
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?**	**4a** C☐ P☐

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)	M☐ N☐
4b. Electronic Sources 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 4b.2 **If not, specify the near-term path to achieve electronic capture by most providers.**	**4b** C☐ P☐ M☐ N☐
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☐ P☐ M☐ N☐ NA ☐
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.** Based on national average mortality rates taken from the 2000 Nationwide Inpatient Sample the minimum hospital caseload necessary to detect a doubling of the mortality rate for esophageal resection is 77 (the rate the authors determined necessary to reliably detect increased mortality in poor performing hospitals). Only 1% of hospitals performed esophageal resections frequently when combining 3 years of data for the authors to advocate use of this indicator as a measure of hospital quality at the hospital-level.[1] AHRQ IQIs, including Esophageal Resection Mortality Rate, were easily applied to Veterans Administration data (2004 – 2007). The relative insensitivity of procuedure-related mortality indicators to detect temporal change or site differences in the VA are hypothesized in this study to be attributable to "the success of longstanding VA programs… or because of inadequate sample sizes (eg. esophageal cancer resection had only 0-12 cases in a given year)." [2] [1] Justin B. Dimick, MD; H. Gilbert Welch, MD, MPH; John D. Birkmeyer, MD. Surgical Mortality as an Indicator of Hospital Quality: The Problem With Small Sample Size. JAMA. 2004;292:847-851. [2] Borzecki Ann M; Christiansen Cindy L; Loveland Susan; Chew Priscilla; Rosen Amy K. Trends in the inpatient quality indicators: the Veterans Health Administration experience. Medical Care. 2010:48:694-702.	**4d** C☐ P☐ M☐ N☐
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:** None 4e.2 **Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:** Administrative data are collected as part of routine operations. Additional staff time required to download and execute the software 4e.3 **Evidence for costs:** Reported user experience 4e.4 **Business case documentation:** None	**4e** C☐ P☐ M☐ N☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility?*	**4**

Steering Committee: Overall, to what extent was the criterion, *Feasibility*, met? **Rationale:**	**4** **C☐** **P☐** **M☐** **N☐**
RECOMMENDATION	
(for NQF staff use) **Check if measure is untested and only eligible for time-limited endorsement.**	**Time-limited ☐**
Steering Committee: Do you recommend for endorsement? **Comments:**	**Y☐** **N☐** **A☐**

CONTACT INFORMATION

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

Co.2 Point of Contact
John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-

Measure Developer If different from Measure Steward
Co.3 Organization
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, Maryland, 20850

Co.4 Point of Contact
John, Bott, MSSW, MBA, john.bott@ahrq.hhs.gov, 301-427-1317-

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

Co.6 Additional organizations that sponsored/participated in measure development
UC Davis
Standford University
Battelle Memorial Institiute

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development
Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
None

Ad.2 If adapted, provide name of original measure: Not applicable
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: 2002
Ad.7 Month and Year of most recent revision: 10, 2010
Ad.8 What is your frequency for review/update of this measure? annually
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. We have no copyright disclaimers.

Ad.11 -13 Additional Information web page URL or attachment: URL
http://www.qualityindicators.ahrq.gov/downloads/pqi/PQI%20Comparative%20Data%202008.pdf

Date of Submission (*MM/DD/YY*)**:** 12/31/2010