5.3 Abdominal Aortic Aneurysm Repair Volume (IQI 4)

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.

Relationship to Quality	Higher volumes have been associated with better outcomes, which represent better quality.
Benchmark	Threshold 1: 10 or more procedures per year[34] Threshold 2: 32 or more procedures per year[35 36 37]
Definition	Raw volume of provider-level AAA repair.
Numerator	Discharges, age 18 years and older, with ICD-9-CM codes of 3834, 3844, 3864 and 3971 in any procedure field with a diagnosis code of AAA in any field. Exclude cases: • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Denominator	Not applicable.
Type of Indicator	Provider Level, Procedure Volume Indicator

Summary of Evidence

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.

As noted in the literature, higher volume hospitals have lower mortality than lower volume hospitals, and the differences in patient case-mix do not account fully for these relationships.

Empirical evidence shows that a moderate to low percentage of procedures were performed at high-volume hospitals, depending on which threshold is used. At threshold 1, 83.9% of AAA repair procedures were performed at high-volume providers (and 44.3% of providers are high volume). At threshold 2, 43.0% were performed at high-volume providers (and 12.2% of providers are high volume).[34 35 36 37]

Limitations on Use

As a volume indicator, AAA repair is a proxy measure for quality and should be used with other indicators.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

The face validity of AAA repair depends on whether a strong association with outcomes of care is widely accepted in the professional community. No consensus recommendations about minimum procedure volume currently exist.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

[34] 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. Health Serv Res 1992;27(4):517-42.

[35] Kazmers A, Jacobs L, Perkins A, et al. Abdominal aortic aneurysm repair in Veterans Affairs medical centers. J Vasc Surg 1996;23(2):191-200.

[36] Pronovost PJ, Jenckes MW, Dorman T, et al. Organizational characteristics of intensive care units related to outcomes of abdominal aortic surgery. JAMA 1999;281(14):1310-7.

[37] Nationwide Inpatient Sample and State Inpatient Databases. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup

AAA repair is an uncommon cardiovascular procedure—only 48,600 were performed in the United States in 1997.[38] 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.

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Risk adjustment is not appropriate, because volume measures are not subject to bias due to disease severity and comorbidities.

Construct validity: Does the indicator perform well in identifying true (or actual) quality of care problems?

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.[39] One study that evaluated the impact of total vascular surgery volume found a significant effect for both ruptured and intact aneurysms.[40] 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).[41]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

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.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

The Center for Medical Consumers posts volumes of "resection of aorta with replacement" for New York hospitals.[42] The Pacific Business Group on Health states that "one marker of how well a hospital is likely to perform is...the number of (AAA) surgeries a hospital performs."[43]

[38] HCUPnet. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup/.
[39] 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.
[40] 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.
[41] Nationwide Inpatient Sample.
[42] The Center for Medical Consumers. (http://www.medicalconsumers.org/)
[43] http://www.pbgh.org/

5.4 Coronary Artery Bypass Graft Volume (IQI 5)

Coronary artery bypass graft (CABG) requires proficiency with the use of complex equipment; and technical errors may lead to clinically significant complications, such as myocardial infarction, stroke, and death.

Relationship to Quality	Higher volumes have been associated with better outcomes, which represent better quality.
Benchmark	Threshold 1: 100 or more procedures per year[44] Threshold 2: 200 or more procedures per year[45 46]
Definition	Raw volume of provider-level CABG.
Numerator	Discharges, age 18 years and older, with ICD-9-CM codes of 3610 through 3619 in any procedure field. Exclude cases: • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Denominator	Not applicable.
Type of Indicator	Provider Level, Procedure Volume Indicator

Summary of Evidence

CABG is measured with great precision, although volume indicators overall are not direct measures of quality and are relatively insensitive. For this reason, CABG should be used in conjunction with other measures of mortality to ensure that increasing volumes truly improve patient outcomes.

As noted in the literature, higher volumes of CABG have been associated with fewer deaths. However, the American Heart Association (AHA) and the American College of Cardiology (ACC) recommend that since some low-volume hospitals have very good outcomes, other measures besides volume should be used to evaluate individual surgeon's performance.

Empirical evidence shows that a high percentage of procedures were performed at high-volume hospitals. At threshold 1, 98.3% of CABG procedures were performed at high-volume providers (and 88% of providers are high volume).[44] At threshold 2, 90.7% were performed at high-volume providers (and 68% of providers are high volume).[45 46]

Limitations on Use

As a volume indicator, CABG is a proxy measure for quality and should be used with other indicators.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

The face validity of CABG depends on whether a strong association with outcomes of care is both plausible and widely accepted in the professional community. The AHA and ACC have argued for "careful outcome tracking" and supported "monitoring institutions and individuals who annually perform fewer than 100 cases," although the panel noted that "some institutions and practitioners maintain excellent outcomes despite relatively low volumes."[47]

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

[44]Eagle KA, Guyton RA, Davidoff R, et al. ACC/AHA Guidelines for Coronary Artery Bypass Graft Surgery: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1991 Guidelines for Coronary Artery Bypass Graft Surgery). American College of Cardiology/American Heart Association. J Am Coll Cardiol 1999;34(4):1262-347.

[45]Hannan EL, Kilburn H, Jr., Bernard H, et al. Coronary artery bypass surgery: the relationship between inhospital mortality rate and surgical volume after controlling for clinical risk factors. Med Care 1991;29(11):1094-107.

[46]Nationwide Inpatient Sample and State Inpatient Databases. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup

[47]Eagle et al. 1999.

CABG is measured accurately with discharge data. The large number of procedures performed annually at most hospitals suggests that annual volume is not subject to considerable random variation. Hannan et al. reported year-to-year hospital volume correlations of 0.96-0.97 in New York.[48]

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Risk adjustment is not appropriate, because volume measures are not subject to bias due to disease severity and comorbidities.

Construct validity: Does the indicator perform well in identifying true (or actual) quality of care problems?

Higher volumes have been repeatedly associated with better outcomes of care, although these findings may be limited by inadequate risk adjustment of the outcome measure.

Hannan found that the adjusted relative risk of inpatient death at high-volume hospitals (more than 200 cases per year) in 1989-92 was 0.84, compared with low-volume hospitals.[49] However, only 3.3% of patients in that study underwent CABG at a low-volume hospital. Analyses using instrumental variables suggested that much of the volume effect may be due to "selective referral" of patients to high-quality centers.[50 51]
Empirical evidence shows that CABG volume and mortality—after adjusting for age, sex, and APR-DRG—is independently and negatively correlated with mortality for CABG (r=-.29, p<.001).[52]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

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 from the procedure. Additionally, shifting procedures to high-volume providers may impair access to care for certain types of patients.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

Specific CABG volume thresholds have been suggested as "standards" for the profession. The Pacific Business Group on Health states that "one marker of how well a hospital is likely to perform is...the number of (CABG) surgeries a hospital performs."[53]

[48]Hannan EL, Kilburn H Jr., Racz M, et al. Improving the outcomes of coronary artery bypass surgery in New York state. JAMA 1994;271(10):761-6.
[49]Hannan et al. 1994.
[50]Farley, DE, Ozminkowski RJ. Volume-outcome relationships and in-hospital mortality: the effect of changes in volume over time. Med Care 1992;30(1):77-94.
[51]Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? Health Serv Res 1987;22(2):157-82.
[52]Nationwide Inpatient Sample.

[53]http://www.pbgh.org/

5.5 Percutaneous Transluminal Coronary Angioplasty Volume (IQI 6)

Percutaneous transluminal coronary angioplasty (PTCA) is a relatively common procedure that requires proficiency with the use of complex equipment, and technical errors may lead to clinically significant complications. The definition for PTCA mortality rate (IQI 30) is also noted below. The QI software calculates mortality for PTCA, so that the volumes for this procedure can be examined in conjunction with mortality. However, the mortality measure should not be examined independently, because it did not meet the literature review and empirical evaluation criteria to stand alone as its own measure.

Relationship to Quality	Higher volumes have been associated with better outcomes, which represent better quality.
Benchmark	Threshold 1: 200 or more procedures per year[54] Threshold 2: 400 or more procedures per year[55 56]
Definition	Raw volume of PTCA.
Numerator	Discharges, age 18 years and older, with ICD-9-CM codes 0066, 3601, 3602, 3605 in any procedure field. Exclude cases: • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Denominator	Not applicable.
Type of Indicator	Provider Level, Procedure Volume Indicator

5.6 PTCA Mortality Rate (IQI 30)

Relationship to Quality	Better processes of care may reduce short-term mortality, which represents better quality.
Definition	Number of deaths per 100 PTCAs.
Numerator	Number of deaths among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	Discharges, age 40 years and older, with ICD-9-CM codes 0066, 3601, 3602, 3605 in any procedure field. Exclude cases: • missing discharge disposition (DISP=missing) • transferring to another short-term hospital (DISP=2) • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Type of Indicator	Provider Level, Mortality Indicator – Recommended for use only with the corresponding volume indicator above.

Summary of Evidence

PTCA is measured with great precision, although volume indicators overall are not direct measures of quality and are relatively insensitive. For this reason, PTCA should be used in conjunction with measures of mortality and quality of care within cardiac care to ensure that increasing volumes truly improve patient outcomes. As noted in the literature, higher volumes of PTCA have been associated with fewer deaths and post-procedural coronary artery bypass grafts (CABG).

Empirical evidence shows that a moderate to high percentage of procedures were performed at high-volume hospitals. At threshold 1, 95.7% of PTCA procedures were performed at high-volume providers (and 69% of the providers are high volume).[54] At threshold 2, 77.0% were

[54]Ryan TJ, Bauman WB, Kennedy JW, et al. Guidelines for percutaneous transluminal coronary angioplasty. .A report of the American Heart Association/American College of Cardiology Task Force on Assessment of Diagnostic and Therapeutic Cardiovascular Procedures (Committee on

performed at high-volume providers (and 42% of providers are high volume).[55 56]

Limitations on Use

As a volume indicator, PTCA is a proxy measure for quality and should be used with other indicators.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

The face validity of PTCA depends on whether a strong association with outcomes of care is both plausible and widely accepted in the professional community. The American Heart Association (AHA) and the American College of Cardiology (ACC) have stated that "a significant number of cases per institution—at least 200 PTCA procedures annually—is essential for the maintenance of quality and safe care."[57] Providers may wish to examine rates by surgeon with this indicator.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

PTCA is an increasingly common procedure (16.7 per 10,000 persons in 1997[58]) and is measured accurately with discharge data. The large number of procedures performed annually at most hospitals suggests that annual volume is not subject to considerable random variation.

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Risk adjustment is not appropriate, because volume measures are not subject to bias due to disease severity and comorbidities.

Construct validity: Does the indicator perform well in identifying true (or actual) quality of care problems?

Higher volumes have been repeatedly associated with better outcomes of care, although these findings may be limited by inadequate risk adjustment of the outcome measure.

Using hospital discharge data to adjust for age, gender, multilevel angioplasty, unstable angina, and six comorbidities, one study found that high-volume hospitals had significantly lower rates of same-stay coronary artery bypass surgery (CABG) and inpatient mortality than low-volume hospitals.[59] Better studies based on clinical data systems (adjusting for left ventricular function) have confirmed higher risk-adjusted mortality and CABG rates at low-volume hospitals relative to high-volume hospitals.[60]

Empirical evidence shows that PTCA volume is negatively related to several other post-procedural mortality rates: CABG (r=-.21, p<.001), craniotomy (r=-.200, p<.0001), and abdominal aortic aneurysm (AAA) repair (r=-.45, p<.0001).[61]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

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 from the procedure. Additionally, shifting procedures to high-volume providers may impair access to care for certain types of patients.

Percutaneous Transluminal Coronary Angioplasty). Circulation 1993;88(6):2987-3007.

[55]Hannan EL, Racz M, Ryan TJ, et al. .Coronary angioplasty volume-outcome relationships for hospitals and cardiologists. JAMA 1997;277(11):892-8.

[56]Nationwide Inpatient Sample and State Inpatient Databases. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup

[57]Ryan et al., 1993.

[58]Kozak LJ, Lawrence L. National Hospital Discharge Survey: annual summary, 1997. Vital Health Stat 13 1999(144):i-iv, 1-46.

[59]Ritchie JL, Maynard C, Chapko MK, et al. Association between percutaneous transluminal coronary angioplasty volumes and outcomes in the Healthcare Cost and Utilization Project 1993-1994. Am J Cardiol 1999;83(4):493-7.

[60]Hannan et al. 1997.

[61]Nationwide Inpatient Sample.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

PTCA volume has not been widely used as an indicator of quality, although specific volume thresholds have been suggested as "standards" for the profession.[62]

[62]Hirshfeld JW, Jr., Ellis SG, Faxon DP. Recommendations for the assessment and maintenance of proficiency in coronary interventional procedures: Statement of the American College of Cardiology. J Am Coll Cardiol 1998;31(3):722-43.

5.11 Abdominal Aortic Aneurysm Repair Mortality Rate (IQI 11)

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.

Relationship to Quality	Better processes of care may reduce mortality for AAA repair, which represents better quality care.
Benchmark	State, regional, or peer group average.
Definition	Number of deaths per 100 discharges with procedure code of AAA repair.
Numerator	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator
Denominator	Discharges, age 18 years and older, with ICD-9-CM codes of 3834, 3844, 3864, 3971 in any procedure field and a diagnosis code of AAA in any field. Exclude cases: • missing discharge disposition (DISP=missing) • transferring to another short-term hospital (DISP=2) • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Type of Indicator	Provider Level, Mortality Indicator for Inpatient Procedures

Summary of Evidence

AAA repair is a technically difficult procedure with a relatively high mortality rate. Higher volume hospitals have been noted to have lower mortality rates, which suggests that some differences in the processes of care between lower and higher volume hospitals result in better outcomes.

Empirical analyses of demographic risk adjustment noted some potential bias for this indicator. Additional medical chart review or analyses of laboratory data may be helpful in determining whether more detailed risk adjustment is necessary. This indicator should also be considered with length of stay and transfer rates to account for differing discharge practices among hospitals.

Limitations on Use

Risk adjustment for clinical factors is recommended because of the confounding bias for AAA repair mortality rate. In addition, little evidence exists supporting the construct validity of this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Studies have reported 40-55% in-hospital mortality after emergent repair of ruptured aneurysms.[84] [85] [86] These data suggest that improved quality of care could have a substantial impact on public health.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

The relatively small number of AAA resections performed by each hospital suggests that mortality rates at the hospital level are likely to be unreliable. Empirical evidence shows that his indicator is precise, with a raw provider level

[84] Dardik A, Burleyson GP, Bowman H, et al. Surgical repair of ruptured abdominal aortic aneurysms in the state of Maryland: factors influencing outcome among 527 recent cases. J Vasc Surg 1998;28(3):413-20.
[85] Kazmers A, Jacobs L, Perkins A, et al. Abdominal aortic aneurysm repair in Veterans Affairs medical centers. J Vasc Surg 1996;23(2):191-200.
[86] Rutledge R, Oller DW, Meyer AA, et al. A statewide, population-based time-series analysis of the outcome of ruptured abdominal aortic aneurysm. Ann Surg 1996;223(5):492-502.

mean of 21.5% and a substantial standard deviation of 26.8%.[87]

Relative to other indicators, a higher percentage of the variation occurs at the provider level, rather than the discharge level. The signal ratio (i.e., the proportion of the total variation across providers that is truly related to systematic differences in provider performance rather than random variation) is low, at 30.7%, indicating that some of the observed differences in provider performance likely do not represent true differences.

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

The known predictors of in-hospital mortality include whether the aneurysm is intact or ruptured, age, female gender, admission through an emergency room, various comorbidities such as renal failure and dysrhythmias, and Charlson's comorbidity index.[88 89 90] In the absence of studies explicitly comparing models with and without additional clinical elements, it is difficult to assess whether administrative data contain sufficient information to remove bias.

Construct validity: Does the indicator perform well in identifying true (or actual) quality of care problems?

The correlation between hospital or physician characteristics and in-hospital mortality in most studies supports the validity of in-hospital mortality as a measure of quality.[91 92] Finally, excessive blood loss, which is a potentially preventable complication of surgery, has been identified as the most important predictor of mortality after elective AAA repair.[93]

Empirical evidence shows that AAA repair mortality is positively related to other post-procedural mortality measures, such as craniotomy (r=.28, p<.0001) and coronary artery bypass graft (CABG) (r=.17, p<.01).[94]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

All in-hospital mortality measures may encourage earlier post-operative discharge, and thereby shift deaths to skilled nursing facilities or outpatient settings. Another potential response would be to avoid operating on high-risk patients.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

The Pennsylvania Health Care Cost Containment Council includes AAA repair in the "Other major vessel operations except heart (DRG 100)" indicator. It is also used by HealthGrades.com.

[87] Nationwide Inpatient Sample and State Databases. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup/

[88] Manheim LM, Sohn MW, Feinglass J, et al. Hospital vascular surgery volume and procedure mortality rates in California, 1982-1994. J Vasc Surg 1998;28(1):45-56.

[89] 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. Health Serv Res 1992;27(4):517-42.

[90] Wen SW, Simunovic M, Williams JI, et al. Hospital volume, calendar age, and short term outcomes in patients undergoing repair of abdominal aortic aneurysm: the Ontario experience, 1988-92. J Epidemiol Community Health 1996;50(2):207-13.

[91] Pearce WH, Parker MA, Feinglass J, et al. The importance of surgeon volume and training in outcomes for vascular surgical procedures. J Vasc Surg 1999;29(5):768-76.

[92] Rutledge et al., 1996.

[93] Pilcher DB, Davis JH, Ashikaga T, et al. Treatment of abdominal aortic aneurysm in an entire state over 7½ years. Am J Surg 1980;139(4):487-94.

[94] Nationwide Inpatient Sample.

5.12 Coronary Artery Bypass Graft Mortality Rate (IQI 12)

Coronary artery bypass graft (CABG) is a relatively common procedure that requires proficiency with the use of complex equipment; and technical errors may lead to clinically significant complications such as myocardial infarction, stroke, and death.

Relationship to Quality	Better processes of care may reduce mortality for CABG, which represents better quality care.
Benchmark	State, regional, or peer group average.
Definition	Number of deaths per 100 discharges with procedure code of CABG.
Numerator	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator
Denominator	Discharges, age 40 years and older, with ICD-9-CM codes of 3610 through 3619 in any procedure field. Exclude cases: • missing discharge disposition (DISP=missing) • transferring to another short-term hospital (DISP=2) • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Type of Indicator	Provider Level, Mortality Indicator for Inpatient Procedures

Summary of Evidence

CABG mortality is one of the most widely used and publicized post-procedural mortality indicators. Demographics, comorbidities, and clinical characteristics of severity of disease are important predictors of outcome that may vary systematically by provider. Chart review may help distinguish comorbidities from complications.

This indicator should be considered with length of stay and transfer rates to account for differing discharge practices among hospitals. The use of smoothed estimates to help avoid the erroneous labeling of outlier hospitals is recommended.

Limitations on Use

Some selection of the patient population may lead to bias; providers may perform more CABG procedures on less clinically complex patients with questionable indications. Risk adjustment for clinical factors, or at a minimum APR-DRGs, is recommended because of the confounding bias of this indicator. Finally, the evidence for the construct validity of this indicator is limited.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Post-CABG mortality rates have recently become the focus of State public reporting initiatives.[95] Studies suggest that these reports serve as the basis for discussions between physicians and patients about the risks of cardiac surgery.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

Without applying hierarchical statistical models to remove random noise, it is likely that hospitals will be identified as outliers as a result of patient variation and other factors beyond the hospital's control. Empirical evidence shows that this indicator is precise, with a raw provider level mean of 5.1% and a standard deviation of 6.2%.[96]

Relative to other indicators, a lower percentage of the variation occurs at the provider level, rather than the discharge level. The signal ratio (i.e., the proportion of the total variation across providers that is truly related to systematic differences in provider performance rather than

[95] Localio AR, Hamory BH, Fisher AC, et al. The public release of hospital and physician mortality data in Pennsylvania. A case study. Med Care 199;35(3):272-286.
[96] Nationwide Inpatient Sample and State Databases. Healthcare Cost and Utilization Project. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup/

random variation) is moderate, at 54.5%, indicating that some of the observed differences in provider performance likely do not represent true differences.

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Based on studies using large databases, cardiac function, coronary disease severity, and the urgency of surgery appear to be powerful predictors of mortality.[97] Some of these risk factors are not available from administrative data.

Construct validity: Does the indicator perform well in identifying true (or actual) quality of care problems?

Numerous studies have reported an association between hospital volume and mortality after CABG surgery. However, experienced surgeons and surgical teams should be able to improve post-operative mortality by reducing aortic cross-clamp time, which has been repeatedly associated with post-operative mortality after adjusting for a variety of patient characteristics.[98] It is unknown how performance of these processes of care would affect hospital-level mortality rates.

Empirical evidence shows that CABG mortality is positively related to bilateral catheterization and negatively related to laparoscopic cholecystectomy.[99]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

Public reporting of CABG mortality rates may cause providers to avoid high-risk patients. Sixty-three percent of cardiothoracic surgeons surveyed in Pennsylvania reported that they were "less willing" to operate on the most severely ill patients since mortality data were released.[100] However, one study using Medicare data shows no evidence that cardiac surgeons in New York, which also reports CABG mortality rates, avoided high-risk patients.[101] All in-hospital mortality measures may encourage earlier post-operative discharge, shifting deaths to skilled nursing facilities or outpatient settings and causing biased comparisons across hospitals with different mean lengths of stay.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

CABG mortality is publicly reported by California, New Jersey, New York, and Pennsylvania. Recent users of CABG mortality as a quality indicator include the University Hospital Consortium, the Joint Commission on Accreditation of Healthcare Organizations' (JCAHO's) IMSystem, Greater New York Hospital Association, the Maryland Hospital Association (as part of the Maryland QI Project) and HealthGrades.com.

[97] Higgins TL, Estafanous FG, Loop FD, et al. Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients. A clinical severity score. JAMA 1992;267(17):2344-8.
[98] Ottino G, Bergerone S, Di Leo M, et al. Aortocoronary bypass results: a discriminant multivariate analysis of risk factors of operative mortality. J Cardiovasc Surg (Torino) 1990;31(1):20-5.
[99] Nationwide Inpatient Sample.
[100] Hannan EL, Siu AL, Kumar D, et al. Assessment of coronary artery bypass graft surgery performance in New York. Is there a bias against taking high-risk patients? Med Care 1997;35(1):49-56.
[101] Peterson ED, DeLong ER, Jollis JG, et al. Public reporting of surgical mortality: a survey of new York State cardiothoracic surgeons. Ann Thorac surg 1999;68(4):1195-200; discussion 12-1-2.

5.20 Hip Fracture Mortality Rate (IQI 19)

Hip fractures, which are a common cause of morbidity and functional decline among elderly persons, are associated with a significant increase in the subsequent risk of mortality.

Relationship to Quality	Better processes of care may reduce mortality for hip fracture, which represents better quality.
Benchmark	State, regional, or peer group average.
Definition	Number of deaths per 100 discharges with principal diagnosis code of hip fracture.
Numerator	Number of deaths (DISP=20) among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	All discharges, age 18 years and older, with a principal diagnosis code for hip fracture. Exclude cases: • missing discharge disposition (DISP=missing) • transferring to another short-term hospital (DISP=2) • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 15 (newborns and other neonates)
Type of Indicator	Provider Level, Mortality Indicator for Inpatient Conditions

Summary of Evidence

Complications of hip fracture and other comorbidities lead to a relatively high mortality rate, and evidence suggests that some of these complications are preventable. Hip fracture mortality rate is measured with good precision, although some of the observed variance does not reflect true differences in performance. About 89% of hip fracture patients are elderly.

Patient age, sex, comorbidities, fracture site, and functional status are all predictors of functional impairment and mortality. Administrative data may not contain sufficient information for these risk factors.

Limitations on Use

Thirty-day mortality may be somewhat different than in-hospital mortality, leading to information bias. Mortality rates should be considered in conjunction with length of stay and transfer rates. Risk adjustment for clinical factors (or at a minimum APR-DRGs) is recommended. Limited evidence exists for the construct validity of this indicator.

Details

Face validity: Does the indicator capture an aspect of quality that is widely regarded as important and subject to provider or public health system control?

Hip fractures are associated with a significant increase in the subsequent risk of mortality, which persists for a minimum of 3 months among the oldest and most impaired individuals.[146] [147] Elderly patients often have multiple comorbidities and pre-fracture functional impairments. As a result, they are at significant risk of postoperative complications, which—if not recognized and effectively treated—can lead to life-threatening problems.

Precision: Is there a substantial amount of provider or community level variation that is not attributable to random variation?

The largest published study of in-hospital mortality reported a rate of 4.9% in 1979-88, which suggests that mortality rates are likely to be relatively reliable at the hospital level.[148] Empirical evidence shows that this indicator is precise, with a raw provider level mean of 14.4% and a standard deviation of 16.0%.[149]

---

[146] Forsen L, Sogaard AJ, Meyer HE, et al. Survival after hip fracture: short- and long-term excess mortality according to age and gender. Osteoporos Int 1999;10(1):73-8.
[147] Wolinsky FD, Fitzgerald JF, Stump TE. The effect of hip fracture on mortality, hospitalization, and functional status: a prospective study. Am J Public Health 1997;87(3):398-403.
[148] Myers AH, Robinson EG, Van Natta ML, et al. Hip fractures among the elderly: factors associated with in-hospital mortality. Am J Epidemiol 1991;134(10):1128-37.
[149] Nationwide Inpatient Sample and State Inpatient Databases. Healthcare Cost and Utilization Project. Agency

Relative to other indicators, a higher percentage of the variation occurs at the provider level, rather than the discharge level. The signal ratio (i.e., the proportion of the total variation across providers that is truly related to systematic differences in provider performance rather than random variation) is moderate, at 54.3%, indicating that some of the observed differences in provider performance likely do not represent true differences.

Minimal bias: Is there either little effect on the indicator of variations in patient disease severity and comorbidities, or is it possible to apply risk adjustment and statistical methods to remove most or all bias?

Demographic predictors of in-hospital or 30-day mortality include age, male sex, and prior residence in a nursing home. Fracture site may be a significant predictor for long-term outcomes. Comorbidity predictors include malnutrition; venous, digestive, and cardiovascular diseases; neoplasms, disorientation or delirium, chronic obstructive pulmonary disease, the number of chronic medical conditions, prior hospitalization within 1 month, and the American Society of Anesthesiology physical status score.

Empirical analyses confirm that this indicator has some potential bias, and risk adjustment with age and sex and APR-DRGs is highly recommended. Chart review may identify differences in functional status or other clinical factors not accounted for in discharge data.

Construct validity: Does the indicator perform well in identifying true (or actual) quality of care problems?

One study demonstrated that Medicare patients with poor "process of care" had similar risk-adjusted 30-day mortality rates as patients with good process of care.[150] Nevertheless, there is substantial evidence that at least two major causes of death among hip fracture patients are partially preventable: pulmonary emboli and acute myocardial infarction.[151] Very little evidence supports an association between hospital volume and mortality following hip fracture repair.

Empirical evidence shows that hip fracture repair mortality is positively related to pneumonia, stroke, gastrointestinal hemorrhage, and congestive heart failure mortality.[152]

Fosters true quality improvement: Is the indicator insulated from perverse incentives for providers to improve their reported performance by avoiding difficult or complex cases, or by other responses that do not improve quality of care?

All in-hospital mortality measures may encourage earlier post-operative discharge. Thirty-day mortality for hip fracture is substantially higher than in-hospital mortality in the largest published studies, suggesting that a relatively modest decrease in mean length of stay could significantly decrease inpatient mortality. Another potential effect would be to avoid operating on high-risk patients, although this seems unlikely.

Prior use: Has the measure been used effectively in practice? Does it have potential for working well with other indicators?

In-hospital mortality following hip fracture repair has not been widely used as a quality indicator, although it is included within a University Hospital Consortium indicator (mortality for DRG 209).

---

for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/data/hcup/

[150]Kahn KL, Rogers WH, Rubenstein LV, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. JAMA 1990;264(15):1969-73.

[151]Perez JV, Warwick DJ, Case CP, et al. Death after proximal femoral fracture—an autopsy study. Injury 1995;26(4):237-40.

[152]Nationwide Inpatient Sample.

5.3 Decubitus Ulcer (PSI 3)

Definition	Cases of decubitus ulcer per 1,000 discharges with a length of stay greater than 4 days.
Numerator	Discharges with ICD-9-CM code of decubitus ulcer in any secondary diagnosis field among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	All medical and surgical discharges 18 years and older defined by specific DRGs. Exclude cases: • with length of stay of less than 5 days • with ICD-9-CM code of decubitus ulcer in the principal diagnosis field or in a secondary diagnosis field if present on admission, if known • MDC 9 (Skin, Subcutaneous Tissue, and Breast) • MDC 14 (pregnancy, childbirth, and puerperium) • with any diagnosis of hemiplegia, paraplegia, or quadriplegia • with an ICD-9-CM diagnosis code of spina bifida or anoxic brain damage • with an ICD-9-CM procedure code for debridement or pedicle graft **before or on the same day** as the major operating room procedure (surgical cases only) • admitted from a long-term care facility (SID Admission Source=3) • transferred from an acute care facility (SID Admission Source=2)
Type of Indicator	Provider level
Empirical Performance	Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to flag cases of in-hospital decubitus ulcers. Its definition is limited to decubitus ulcer as a secondary diagnosis to better screen out cases that may be present on admission. In addition, this indicator excludes patients who have a length of stay of 4 days or less, as it is unlikely that a decubitus ulcer would develop within this period of time. Finally, this indicator excludes patients who are particularly susceptible to decubitus ulcer, namely patients with major skin disorders (MDC 9) and paralysis.

Panel Review

The overall usefulness of this indicator was rated as very favorable by panelists. Concerns regarding the systematic screening for ulcers and reliability of coding, especially for early stage ulcers, brought into question that assertion. Therefore, this indicator appears to be best used as a rate-based indicator. Panelists suggested that patients admitted from a long-term care facility be excluded, as these patients may have an increased risk of having decubiti present on admission.

Panelists noted that hospitals that routinely screen for decubitus ulcers as part of a quality improvement program might have an artificially high rate of ulcers compared to other hospitals, which may cause this indicator to be somewhat biased.

This indicator includes pediatric patients. Pressure sores are very unusual in children, except among the most critically ill children (who may be paralyzed to improve ventilator management) and children with chronic neurological problems. Age stratification is recommended.

Literature Review

Coding validity. No evidence on validity is available from CSP studies. Geraci et al. confirmed only 2 of 9 episodes of pressure ulcers reported on discharge abstracts of Veterans Affairs (VA) patients hospitalized in

1987-89 for congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), or diabetes.[45] The sensitivity for a nosocomial ulcer was 40%. Among Medicare hip fracture patients, Keeler et al. confirmed 6 of 9 reported pressure ulcers, but failed to ascertain 89 additional cases (6% sensitivity) using ICD-9-CM codes.[46] In the largest study to date, Berlowitz et al. found that the sensitivity of a discharge diagnosis of pressure ulcer among all patients transferred from VA hospitals to VA nursing homes in 1996 was 31% overall, or 54% for stage IV (deep) ulcers.[47] The overall sensitivity increased modestly since 1992 (26.0%), and was slightly but statistically significantly better among medical patients than among surgical patients (33% versus 26%).

Construct validity. Needleman and Buerhaus found that nurse staffing was inconsistently associated with the occurrence of pressure ulcers among medical patients, and was independent of pressure ulcers among major surgery patients.[48] As was expected, nursing skill mix (RN hours/licensed nurse hours) was significantly associated with the pressure ulcer rate.[49] Total licensed nurse hours per acuity-adjusted patient day were inconsistently associated with the rate of pressure ulcers.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Decubitus Ulcer generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is high, relative to other indicators, at 85.6%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.0147, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.01067. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Decubitus Ulcer is high, indicating that the measure is biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al.[50] as part of the Complications Screening Program (CSP 6, "cellulitis or decubitus ulcer"). Needleman and Buerhaus identified decubitus ulcer as an "outcome potentially sensitive to nursing".[51] The American Nurses Association, its State associations, and the California Nursing Outcomes Coalition have identified the total prevalence of inpatients with Stage I, II, III, or IV pressure ulcers as a "nursing-sensitive quality indicator for acute care settings."[52]

[45] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. Med Care 1997;35(6):589-602.
[46] Keeler E, Kahn K, Bentow S. Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare Provider Analysis and Review file. Springfield, VA: NTIS; 1991.
[47] Berlowitz D, Brand H, Perkins C. Geriatric syndromes as outcome measures of hospital care: Can administrative data be used? JAGS 1999;47:692-696.
[48] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.: 230-88-0021.
[49] Lichtig LK, Knauf RA, Hilholland DK. Some impacts of nursing on acute care hospital outcomes. J Nurs Adm 1999;29(2):25-33.
[50] Iezzoni LI, Daley J, Heeren T, Foley SM, Risher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.
[51] Needleman et al. 2001.
[52] Nursing-Sensitive Quality Indicators for Acute Care Settings and ANA's Safety & Quality Initiative. In: American Nurses Association; 1999.

5.4 Failure to Rescue (PSI 4)

Definition	Deaths per 1,000 patients having developed specified complications of care during hospitalization.
Numerator	Discharges with a disposition of "deceased" among cases meeting the inclusion and exclusion rules for the denominator.
Denominator	Discharges 18 years and older with potential complications of care listed in failure to rescue definition (i.e., pneumonia, DVT/PE, sepsis, acute renal failure, shock/cardiac arrest, or GI hemorrhage/acute ulcer). Exclude cases: • age 75 years and older • neonatal patients in MDC 15 • transferred to an acute care facility (SID Discharge Disposition = 2) • transferred from an acute care facility (SID Admission Source = 2) • admitted from a long-term care facility (SID Admission Source=3) Additional exclusion criteria specific to each diagnosis.
Type of Indicator	Provider level
Empirical Performance	Population Rate (2003): 127.687 per 1,000 population at risk Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to identify patients who die following the development of a complication. The underlying assumption is that good hospitals identify these complications quickly and treat them aggressively.

Failure to Rescue may be fundamentally different than other indicators reviewed in this report, as it may reflect different aspects of quality of care (effectiveness in rescuing a patient from a complication versus preventing a complication). This indicator includes pediatric patients. It is important to note that children beyond the neonatal period inherently recover better from physiological stress and thus may have a higher rescue rate.

Panel Review

Panelists expressed concern regarding patients with "do not resuscitate" (DNR) status. In cases where this DNR status is not a direct result of poor quality of care, it would be contrary to patient desire and poor quality of care to rescue a patient. In addition, very old patients—or patients with advanced cancer or HIV—may not desire or may be particularly difficult to rescue from these complications. As a result, this indicator definition was modified to exclude those patients age 75 years and older. In addition, panelists suggested the exclusion of patients admitted from long-term care facilities.

Panelists noted that several adverse incentives may be introduced by implementing this indicator. In particular, since some type of adjustment may be desirable, this indicator may encourage the upcoding of complications and comorbidities to inflate the denominator or manipulate risk adjustment. Others noted that this indicator could encourage irresponsible resource use and allocation, although this is likely to be a controversial idea. Finally, panelists emphasized that this indicator should be used internally by hospitals, as it is not validated for public reporting.

Literature Review

Construct validity. Silber and colleagues have published a series of studies establishing the construct validity of failure-to-rescue rates through their associations with hospital characteristics and other measures of hospital

performance. Among patients admitted for cholecystectomy and transurethral prostatectomy, failure to rescue was independent of severity of illness at admission, but was significantly associated with the presence of surgical house staff and a lower percentage of board-certified anesthesiologists.[53] The adverse occurrence rate was independent of this hospital characteristic. In a larger sample of patients who underwent general surgical procedures, lower failure-to-rescue rates were found at hospitals with high ratios of registered nurses to beds.[54] Failure rates were strongly associated with risk-adjusted mortality rates, as expected, but not with complication rates.[55]

More recently, Needleman and Buerhaus confirmed that higher registered nurse staffing (RN hours/adjusted patient day) and better nursing skill mix (RN hours/licensed nurse hours) were consistently associated with lower failure-to-rescue rates, even using administrative data to define complications.[56]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Failure to Rescue generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 66.6%, suggesting that observed differences in risk-adjusted rates may reflect true differences across hospitals.

The signal standard deviation for this indicator is also high, relative to other indicators, at 0.04617, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.01450. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Failure to Rescue is high, indicating that the measures are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Silber et al. as a more powerful tool than the risk-adjusted mortality rate to detect true differences in patient outcomes across hospitals.[57] The underlying premise was that better hospitals are distinguished not by having fewer adverse occurrences but by more successfully averting death among (i.e., rescuing) patients who experience such complications. More recently, Needleman and Buerhaus adapted Failure to Rescue to administrative data sets, hypothesizing that this outcome might be sensitive to nurse staffing.[58]

[53] Silber JH, Williams SV, Krakauer H, Schwartz JS. Hospital and patient characteristics associated with death after surgery. A study of adverse occurrence and failure to rescue. Med Care 1992;30(7):615-29.
[54] Silber J, Rosenbaum P, Ross R. Comparing the contributions of groups of predictors: Which outcomes vary with hospital rather than patient characteristics? J Am Stat Assoc 1995;90:7-18.
[55] Silber JH, Rosenbaum PR, Williams SV, Ross RN, Schwartz JS. The relationship between choice of outcome measure and hospital rank in general surgical procedures: Implications for quality assessment. Int J Qual Health Care 1997;9(3):193-200.
[56] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes In Hospitals. Boston MA: Health Resources and Services Administration: 2001 February 28. Report No.:230-99-0021.
[57] Silber et al. 1992.
[58] Needleman et al. 2001.

5.16 Postoperative Pulmonary Embolism or Deep Vein Thrombosis (PSI 12)

Definition	Cases of deep vein thrombosis (DVT) or pulmonary embolism (PE) per 1,000 surgical discharges with an operating room procedure.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM codes for deep vein thrombosis or pulmonary embolism in any secondary diagnosis field
Denominator	All surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure. Exclude cases: • with preexisting (principal diagnosis or secondary diagnosis present on admission, if known) deep vein thrombosis or pulmonary embolism where a procedure for interruption of vena cava is the only operating room procedure • where a procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure *Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.* • MDC 14 (Pregnancy, Childbirth and the Puerperium)
Type of Indicator	Provider level
Empirical Performance	Population Rate (2003): 9.830 per 1,000 population at risk Bias: Substantial bias; should be risk-adjusted
Risk Adjustment	Age, sex, DRG, comorbidity categories

Summary

This indicator is intended to capture cases of postoperative venous thromboses and embolism—specifically, pulmonary embolism and deep venous thrombosis. This indicator limits vascular complications codes to secondary diagnosis codes to eliminate complications that were present on admission. It further excludes patients who have principal diagnosis of DVT, as these patients are likely to have had PE/DVT present on admission.

Panel Review

Panelists rated the overall usefulness of this indicator relatively highly as compared to other indicators. They noted that preventative techniques should decrease the rate of this indicator. This indicator includes pediatric patients. In the absence of specific thrombophilic disorders, postoperative thromboembolic complications in children are most likely to be secondary to venous catheters rather than venous stasis in the lower extremities.

Because the risk for DVT/PE varies greatly according to the type of procedure performed, panelists suggested that this indicator be adjusted or stratified according to surgical procedure types.

Literature Review

Coding validity. Geraci et al. confirmed only 1 of 6 episodes of DVT or PE reported on discharge abstracts of VA patients for CHF, COPD, or diabetes; the sensitivity was 100%.[96] Among Medicare hip fracture patients, by contrast, Keeler et al. confirmed 88% of reported PE cases, and failed to ascertain just 6 cases (65% sensitivity) using ICD-9-CM codes.[97] For DVT, they found just 1 of 6 cases using ICD-9-CM codes (but no false positive codes). Other studies have demonstrated that ICD-9-CM codes for DVT and PE have high predictive value when listed as the principal diagnosis for readmissions after major orthopedic

[96] Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. In-hospital complications among survivors of admission for congestive heart failure, chronic obstructive pulmonary disease, or diabetes mellitus. J Gen Intern Med 1995;10(6):307-14.

[97] Keeler E. Kahn K, Bentow S. Assessing quality of care for hospitalized Medicare patients with hip fracture using coded diagnoses from the Medicare Provider Analysis and Review File. Springfield, VA: NTIS;1991.

surgery (100%) or after inferior vena cava filter placement (98%).[98] However, these findings do not directly address the validity of DVT/PE as a secondary diagnosis among patients treated by anticoagulation.

Construct validity. Explicit process of care failures in the CSP validation study were relatively frequent among both major surgical and medical cases with CSP 22 (72% and 69%, respectively), after disqualifying cases in which DVT/PE was actually present at admission.[99] Needleman and Buerhaus found that nurse staffing was independent of the occurrence of DVT/PE among both major surgical or medical patients.[100] However, Kovner and Gergen reported that having more registered nurse hours and non-RN hours was associated with a lower rate of DVT/PE after major surgery.[101]

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Postoperative PE or DVT generally performs well on several different dimensions, including reliability, bias, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 72.6%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is lower than many indicators, at 0.00633, indicating that the systematic differences (signal) among hospitals is low and less likely associated with hospital characteristics. The signal share is lower than many indicators, at 0.00511. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The project team assessed the effect of age, gender, DRG, and comorbidity risk adjustment on the relative ranking of hospitals compared to no risk adjustment. They measured (1) the impact of adjustment on the assessment of relative hospital performance, (2) the relative importance of the adjustment, (3) the impact on hospitals with the highest and lowest rates, and (4) the impact throughout the distribution. The detected bias for Postoperative PE or DVT is high, indicating that the measures likely are biased based on the characteristics observed. (It is possible that characteristics that are not observed using administrative data may be related to the patient's risk of experiencing an adverse event.) Risk adjustment is important for this indicator.

Source

This indicator was originally proposed by Iezzoni et al. as part of the Complications Screening Program (CSP 22, "venous thrombosis and pulmonary embolism")[102] and was one of AHRQ's original HCUP Quality Indicators for major surgery and invasive vascular procedure patients.[103] A code that maps to this indicator in the final AHRQ PSI was proposed by Miller et al. as one component of a broader indicator ("iatrogenic conditions").[104]

[98] White RH, Romano P, Zhou H, Rodrigo J, Barger W. Incidence and time course of thromboembolic outcomes following total hip or knee arthroplasty. Arch Intern Med 1998;158(14):1525-31.

[99] Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Does the Complications Screening Program flag cases with process of care problems? Using explicit criteria to judge processes. Int J Qual Health Care 1999;11(2):107-18.

[100] Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. Nurse Staffing and Patient Outcomes in Hospitals. Boston, MA: Health Resources Services Administration; 2001 February 28. Report No.:230-99-0021.

[101] Kovner C, Gergen PH. Nurse staffing levels and adverse events following surgery in U.S. hospitals. Image J Nurs Sch 1998;30(4):315-21.

[102] Iezzoni LI, Daley J, Heeren T, Foley SM, Fisher ES, Duncan C, et al. Identifying complications of care using administrative data. Med Care 1994;32(7):700-15.

[103] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

[104] Miller M, Elixhauser A, Zhan C, Meyer G. Patient safety indicators: Using administrative data to identify potential patient safety concerns. Health Services Research 2001;36(6 Part II):110-132.

5.25 Obstetric Trauma—Vaginal Delivery with Instrument (PSI 18)

Definition	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 instrument-assisted vaginal deliveries.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for 3rd and 4th degree obstetric trauma in any diagnosis or procedure field.
Denominator	All vaginal delivery discharges with any procedure code for instrument-assisted delivery.
Type of Indicator	Provider level
Empirical Performance	Bias: Did not undergo empirical testing of bias
Risk Adjustment	Age, comorbidity categories

Summary

This indicator is intended to flag cases of potentially preventable trauma during vaginal delivery with instrument.

Panel Review

The overall usefulness of an Obstetric trauma indicator was rated as favorable by panelists. After initial review, the indicator was eventually split into three separate Obstetric Trauma indicators: Vaginal Delivery with Instrument, Vaginal Delivery without Instrument, and Cesarean Delivery.

Literature Review

Coding validity. In a stratified probability sample of vaginal and Cesarean deliveries, the weighted sensitivity and predictive value of coding for third- and fourth-degree lacerations and vulvar/perineal hematomas (based on either diagnosis or procedure codes) were 89% and 90%, respectively.[158] The authors did not report coding validity for third- and fourth-degree lacerations separately. The project team was unable to find other evidence on validity from prior studies.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Obstetric Trauma—Vaginal Delivery with Instrument generally performs well on several different dimensions, including reliability, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is moderately high, relative to other indicators, at 69.9%, suggesting that observed differences in risk-adjusted rates likely reflect true differences across hospitals.

The signal standard deviation for this indicator is also high, relative to other indicators, at 0.09794, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is high, relative to other indicators, at 0.05539. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The bias for Obstetric Trauma—Vaginal Delivery with Instrument was not measured, since adequate risk adjustment was not available.

Source

An overlapping subset of this indicator (third- or fourth-degree perineal laceration) has been adopted by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) as a core performance measure for "pregnancy and related conditions" (PR-25). Based on expert consensus panels, McKesson Health Solutions included the JCAHO indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications

Measures Module. Fourth Degree Laceration, one of the codes mapped to this PSI, was included as one component of a broader indicator ("obstetrical complications") in AHRQ's original HCUP Quality Indicators.[132]

[132] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.

5.26 Obstetric Trauma—Vaginal Delivery without Instrument (PSI 19)

Definition	Cases of obstetric trauma (3rd or 4th degree lacerations) per 1,000 vaginal deliveries without instrument assistance.
Numerator	Discharges among cases meeting the inclusion and exclusion rules for the denominator.with ICD-9-CM code for 3rd and 4th degree obstetric trauma in any diagnosis or procedure field.
Denominator	All vaginal delivery discharges. Exclude cases: • with instrument-assisted delivery.
Type of Indicator	Provider level
Empirical Performance	Population Rate (2003): 46.340 per 1,000 population at risk Bias: Did not undergo empirical testing of bias
Risk Adjustment	Age, comorbidity categories

Summary

This indicator is intended to flag cases of potentially preventable trauma during a vaginal delivery without instrument.

Panel Review

The overall usefulness of an Obstetric Trauma Indicator was rated as favorable by panelists. After initial review, the indicator was split into three separate Obstetric Trauma indicators: Vaginal Delivery with Instrument, Vaginal Delivery without Instrument, and Cesarean Delivery.

Literature Review

Coding validity. In a stratified probability sample of vaginal and Cesarean deliveries, the weighted sensitivity and predictive value of coding for third- and fourth-degree lacerations and vulvar/perineal hematomas (based on either diagnosis or procedure codes) were 89% and 90%, respectively.[158] The authors did not report coding validity for third- and fourth-degree lacerations separately. The project team was unable to find other evidence on validity from prior studies.

Empirical Analysis

The project team conducted extensive empirical analyses on the PSIs. Obstetric Trauma—Vaginal Delivery without Instrument generally performs well on several different dimensions, including reliability, relatedness of indicators, and persistence over time.

Reliability. The signal ratio—measured by the proportion of the total variation across hospitals that is truly related to systematic differences (signal) in hospital performance rather than random variation (noise)—is high, relative to other indicators, at 86.4%, suggesting that observed differences in risk-adjusted rates reflect true differences across hospitals.

The signal standard deviation for this indicator is also high, relative to other indicators, at 0.04314, indicating that the systematic differences (signal) among hospitals is high and more likely associated with hospital characteristics. The signal share is lower than many other indicators, at 0.02470. The signal share is a measure of the share of total variation (hospital and patient) accounted for by hospitals. The lower the share, the less important the hospital in accounting for the rate and the more important other potential factors (e.g., patient characteristics).

Minimum bias. The bias for Obstetric Trauma—Vaginal Delivery without Instrument was not measured, since adequate risk adjustment was not available.

Source

An overlapping subset of this indicator (third- or fourth-degree perineal laceration) has been adopted by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO) as a core performance measure for "pregnancy and related conditions" (PR-25). Based on expert consensus panels, McKesson Health Solutions included the JCAHO indicator in its CareEnhance Resource Management Systems, Quality Profiler Complications Measures Module. Fourth-Degree Laceration, one of the codes mapped to this PSI, was included as one component of a broader indicator ("obstetrical complications") in AHRQ's original HCUP Quality Indicators.[133]

[133] Johantgen M, Elixhauser A, Bali JK, Goldfarb M, Harris DR. Quality indicators using hospital discharge data: State and national applications. Jt Comm J Qual Improv 1998;24(2):88-195. Published erratum appears in Jt Comm J Qual Improv 1998;24(6):341.