NIH Public Access
Author Manuscript
JAMA Surg. Author manuscript; available in PMC 2014 September 13.

Published in final edited form as:
JAMA Surg. 2014 January ; 149(1): 10–16. doi:10.1001/jamasurg.2013.4109.

Composite measures for profiling hospitals on bariatric surgery performance

Justin B. Dimick, MD, MPH, **Nancy J. Birkmeyer, PhD**, **Jonathan F. Finks, MD**, **David A. Share, MD, MPH**, **Wayne J. English, MD**, **Arthur M. Carlin, MD**, and **John D. Birkmeyer, MD**
Michigan Bariatric Surgery Collaborative (MBSC) (JD, JB, NB, JF), University of Michigan, Ann Arbor, MI; Marquette General Hospital, Marquette, MI (WE); Henry Ford Health System, Detroit, MI (AC); and Blue Cross Blue Shield of Michigan, Detroit, MI (DS)

Abstract

Objective—We sought to develop a novel composite measure for profiling hospital performance with bariatric surgery.

Design, Setting, and Patients—Using clinical registry data from the Michigan Bariatric Surgery Collaborative (MBSC), we studied all patients undergoing bariatric surgery from 2008 to 2010. For gastric bypass surgery, we used empirical Bayes techniques to create a composite measure by combining several measures, including serious complications, reoperations, and readmissions; hospital and surgeon volume; and outcomes with other, related procedures. Hospitals were ranked based on 2008-09 and placed in one of 3 groups: 3-star (top third), 2-star (middle third), and 1-star (bottom third). We assessed how well these ratings predicted outcomes in the next year (2010), compared to other widely used measures.

Main Outcome Measures—Risk-adjusted serious complications.

Results—Composite measures explained a larger proportion of hospital-level variation in serious complication rates with gastric bypass than other measures. For example, the composite measure explained 89% of the variation compared to only 28% for risk-adjusted complication rates alone. Composite measures also appeared better at predicting future performance compared to individual measures. When ranked on the composite measure, 1-star hospitals (bottom 20%), had 2-fold higher serious complication rates (4.6% vs. 2.4%; OR 2.0; 95% CI, 1.1 to 3.5) compared to 3-star (top 20%) hospitals. Differences in serious complications rates between 1-star and 3-star hospitals were much smaller when hospitals were ranked using serious complications (4.0% vs. 2.7%; OR 1.6; 95% CI, 0.8-2.9) and hospital volume (3.3% vs. 3.2%; OR 0.85; 95% CI, 0.4 to 1.7)

Conclusions—Composite measures are much better at explaining hospital-level variation in serious complications and predicting future performance than other approaches. In this preliminary study, it appears that such composite measures may be better than existing alternatives for profiling hospital performance with bariatric surgery.

Correspondence: Justin B. Dimick, MD, MPH, Center for Healthcare Outcomes & Policy, 2800 Plymouth Road, Building 126, Room 137E, Ann Arbor, MI 48109, Phone: (734) 998-7470, Fax: (734) 998-7473, jdimick@umich.edu.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Introduction

There is growing enthusiasm for overhauling the current hospital accreditation program for bariatric surgery. The American College of Surgeons (ACS) and the American Society for Metabolic & Bariatric Surgery (ASMBS), the two leading professional organizations that offer accreditation, currently rely on a volume standard (>125 cases per year) among other structure and process measures [1, 2]. Recent studies, however, have shown that centers receiving accreditation, so-called “centers of excellence”, do not have better outcomes than other hospitals [3, 4]. As a result, there is mounting pressure to move beyond volume standards towards more direct measures of hospital outcomes.

However, the best approach for profiling hospitals based on outcomes is unclear. Many advocate directly measuring outcomes alone (e.g., serious morbidity, reoperation, etc…), but these measures may be too “noisy” to reliably reflect performance [5]. Small sample sizes and low event rates conspire to limit the precision of hospital outcome measures [6, 7]. Because of these limitations, there is growing use of composite scores to measure hospital performance [8, 9]. Composite measures combine multiple different quality indicators into a single score to increase the reliability of hospital performance assessment. Although composite measures have been applied to other medical and surgical conditions, their use has not been explored in bariatric surgery.

In this manuscript, we use data from the Michigan Bariatric Surgical Collaborative (MBSC) clinical registry to explore the value of composite measures of bariatric surgery performance. To create these composite measures, we combined multiple measures by weighting each according to its ability to predict serious morbidity [8, 9]. Because bariatric surgery accreditation is used for selective referral, for example with the Center for Medicare and Medicaid Service s (CMS) national coverage decision, we evaluated this composite measure by its ability to predict future performance compared to other widely used measures. The ability to predict future performance is an essential criterion for quality measures used for selective referral, since patients and payers make decisions about where to have surgery based on historical outcomes data.

Methods

Data Source and Study Population

This study is based on data from the Michigan Bariatric Surgery Collaborative (MBSC), a payer-funded quality improvement program that administers a prospective, externally audited clinical outcomes registry of patients undergoing bariatric surgery in Michigan. The MBSC is a consortium of 29 Michigan hospitals and 75 surgeons performing bariatric surgery and has been described in detail elsewhere [3, 10, 11]. Participation in the MBSC is voluntary and any hospital that performs at least 25 bariatric procedures per year is eligible to participate. The MBSC currently enrolls approximately 6000 patients annually into its clinical registry. Procedures meeting this definition include open and laparoscopic gastric bypass, adjustable gastric banding, sleeve gastrectomy, and biliopancreatic diversion with duodenal switch. Participating hospitals submit data on all patients undergoing primary and revisional bariatric surgery.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

For the MBSC clinical registry, data on patient characteristics, procedure type, processes of care, and postoperative outcomes are obtained by chart abstraction at the end of the perioperative period (in-hospital and up to 30 days after surgery). The medical records are reviewed by centrally trained data abstractors using a standardized and validated data collection instrument. Each hospital within the MBSC is audited annually by nurses from the coordinating center to verify the accuracy and completeness of its clinical registry data. For this study, we identified all patients undergoing a primary (non-revisional) bariatric surgical procedure during 2008 through 2010. We excluded patients undergoing a revisional procedure due to the heterogeneity of risk in this patient population.

Outcomes Assessment

The MBSC registry includes clinical data on 13 different types of perioperative complications. Complications were determined by documentation of the specific complication in the chart, including confirmatory radiographic imaging reports when available, as well as the treatment for the complication. In MBSC, complications are categorized according to severity as nonlife-threatening (grade 1), potentially life-threatening (grade 2), or life-threatening complications associated with permanent residual disability or death (grade 3).

For the purposes of this study, we included all serious complications (grade 2 or higher) in our primary outcome variable [11]. Grade 2 complications included abdominal abscess (requiring percutaneous drainage or reoperation), bowel obstruction (requiring reoperation), leak (requiring percutaneous drainage or reoperation), bleeding (requiring blood transfusion > 4 units, endoscopy, reoperation or splenectomy), wound infection or dehiscence (requiring reoperation), respiratory failure (requiring 2–7 days mechanical ventilation), renal failure (requiring in-hospital dialysis), venous thromboembolism (deep venous thrombosis or pulmonary embolism) and band-related problems requiring reoperation (port site infection, gastric perforation, band slippage and outlet obstruction). Grade 3 complications included myocardial infarction or cardiac arrest, renal failure requiring long-term dialysis, respiratory failure (requiring > 7 days mechanical ventilation or tracheostomy) and death.

Hospital Quality Rankings

Our goal was to compare the composite measure to several existing approaches for assessing hospital performance with bariatric surgery. Herein we describe the methods used to rank hospitals on the following measures: 1) hospital volume; 2) risk-adjusted complication rates; 3) risk-and reliability adjusted complication rates; and 4) composite measures.

Hospital volume—We calculated hospital volume by assessing the average number of cases done per year from 2009-2010. For hospitals that contributed data for less than two years, estimated their predicted annual volume by multiplying their monthly volume by the total number of months they submitted data. All included hospitals had data for more than 12 months.

Risk-adjusted complication rates—We used standard techniques for calculating risk-adjusted complication rates for each hospital [3, 11]. Data on patient characteristics included

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

demographics (age, gender, and payer type), height, weight, mobility limitations (requiring ambulation aids, nonambulatory, or bed-bound), smoking status and comorbid conditions. The height and weight were used to calculate body mass index (BMI), a ratio of the weight in kilograms to the height in meters squared. The definitions for most comorbidities included documentation of the condition and its treatment in the medical record. Comorbid conditions included pulmonary disease (asthma, obstructive/restrictive disorders, home oxygen use, Pickwickian syndrome), cardiovascular disease (coronary artery disease, dysrhythmia, peripheral vascular disease, stroke, hypertension, hyperlipidemia), sleep apnea, psychological disorders, prior venous thromboembolism (VTE), diabetes, chronic renal failure (requiring dialysis or transplant), liver disease (nonalcoholic fatty liver, cirrhosis, liver transplantation), urinary incontinence, gastroesophageal reflux disease, peptic ulcer disease, cholelithiasis, previous ventral hernia repair, and musculoskeletal disorders.

Logistic regression models were used to create models that included all significant patient-level covariates. The predicted probabilities of each patient were estimated from this model and then summed for each hospital to calculate the “expected” number of deaths. The observed number of deaths was then divided by the expected number to yield an “O/E ratio”. This ratio was then multiplied by the overall average to yield a risk-adjusted morbidity rate for each hospital.

Risk- and reliability adjusted complication rates—Reliability adjustment is an increasingly used technique to adjust hospital-specific outcomes for statistical “noise”. This is a based on the standard shrinkage estimator approach that places more weight on a hospital's own complication rate when it is measured reliably, but shrinks back towards the average complication rate when a hospital's own complication rate is measured with error (e.g. for hospitals with small numbers of patients undergoing the procedure) [12]. For this study, we performed reliability adjustment by generating empirical Bayes estimates of hospital-specific risk-adjusted complication rates for each hospital. To create these estimates, we first used a hierarchical model where the patient was the first level and the hospital was the second level. The dependent variable was complications at the patient level and the same risk-adjustment variables described above (in the prior section) were included at the first level as independent variables. The second level included only a hospital level random effect. The random effect in log(odds) was then predicted using empirical Bayes techniques. This was added back to the mean log(odds) of complications in the overall population and an inverse logit was performed to estimate the risk- and reliability adjusted complication rate for each hospital [13].

Composite measure—We developed a composite measure that incorporates information from multiple quality indicators to optimally predict “true” risk-adjusted complications for laparoscopic gastric bypass. In creating these measures, we considered several individual quality measures, including hospital volume, several risk-adjusted outcomes (mortality, complications, reoperation, readmission, and length of stay). We also considered risk-adjusted outcomes not only for the index operation but also for other, related procedures (e.g., laparoscopic banding and sleeve gastrectomy). Of note, all of the input measures were

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

risk-adjusted using clinical registry data, as described above in the section on risk-adjusted complication rates.

Our composite measure is a generalization of the reliability adjustment described above. While the simple shrinkage estimator described above is a weighted average of a single measure of interest and its mean, our composite measure is a weighted average of all available quality indicators. The weight on each quality indicator is determined for each hospital to minimize the expected mean squared prediction error, using an empirical Bayes methodology. Although the statistical methods used to create these measures are described in detail elsewhere [8, 14], we will provide a brief conceptual overview. The first step in creating the composite measure was to determine the extent to which each individual quality indicator predicts risk-adjusted complication rates for the index operation. To evaluate the importance of each potential input, we first estimated the proportion of systematic (i.e., nonrandom) variation in explained by each individual quality indicator (Table 1). We included any quality indicator in the composite measure that explained more than 10% of hospital variation in risk-adjusted complications during 2008-09.

Next, we calculated weights for each quality indicator. The weight placed on each quality indicator in our composite measure was based on 2 factors [14]. The first is the hospital-level correlation of each quality indicator with the complication rate for the index operation. The strength of these correlations indicates the extent to which other quality indicators can be used to help predict complications for the index operation. The second factor affecting the weight placed on each quality indicator is the reliability with which each indicator is measured. Reliability ranges from 0 (no reliability) to 1 (perfect reliability) [12]. The reliability of each quality indicator refers to the proportion of the overall variance that is attributable to true hospital-level variation in performance, as opposed to estimation error ("noise"). For example, in smaller hospitals, less weight is placed complication rates because they are less reliably estimated.

Analysis

We determined the value of different quality indicators by determining how well hospital performance rankings from 2008-09 based on each measurement approach predicted risk-adjusted serious complications in the next year (2010). For each operation, hospitals were ranked based on each quality measure (data from 2008-09) and were divided into three even sized groups (1-star, 2-star, and 3-star). The "worst" hospitals (bottom third) received a 1-star rating, the middle of the distribution (middle third) received a 2-star rating, and the "best" hospitals (top third) received a 3-star rating. We then assessed the ability of our composite measure to predict future performance compared to standard techniques for ranking hospitals, including hospital volume, risk-adjusted complication rates, risk- and reliability adjusted complication rates, and our composite measure. For these analyses, we evaluated the discrimination in future risk-adjusted complications, comparing the 1-star hospitals (bottom 20%) to the 3-star hospitals (top 20%) for each of the measures. These analyses were conducted using patient-level logistic regression models in the 2010 data. The dependent variable was one or more serious complications and the independent variables were patient characteristics used for risk-adjustment. Each quality indicator from 2008-09

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

data was then added as 3 dummy variables (1-star, 2-star, and 3-star), and we present the odds ratio and 95% confidence interval representing 3-star vs. 1-star hospitals.

We also assessed the ability of the composite measure and other quality indicators (assessed in 2008-09) to explain future (2010) hospital level variation in risk-adjusted serious complications. To avoid problems with "noise variation" in the subsequent time period, we determined the proportion of systemic hospital-level variation explained. We generated hierarchical models with one or more complications as the dependent variable (2009) and used them to estimate the hospital-level variance. We first used an "empty model" that contained only patient variables for risk-adjustment. We then entered each historical quality measure (assessed in 2008-09) into the model. We then calculated the degree to which the historical quality measures reduced the hospital-level variance, an approach described in our prior work [14]. All statistical analyses were conducted using STATA 11.0 (College Station, Texas).

Results

Inputs to the composite measure

There were several individual quality indicators that explained hospital-level variation in serious complications with bariatric surgery, varying from hospital volume which explained 11% to rates of prolonged length of stay which explained 65% (Table 1).

The weights applied to each quality indicator in the composite measure are shown in Table 2. The largest amount of weight (39%) is applied to hospital structural characteristics and the overall average—i.e., the target to which observed performance is anchored and "shrunk" back towards. Complications received the next highest weight (25%) followed by readmission (18%), prolonged length of stay (10%), emergency room visit (6%), reoperation (6%), and outcomes with other procedures, such as sleeve gastrectomy (2%) and laparoscopic banding (2%) (Table 2).

Patient characteristics

When hospitals were grouped according to the composite measures, there were not substantial differences in patient illness levels across the groups (Table 3). Although several individual characteristics differed significantly between groups, the expected mortality—a function of all patient characteristics combined—was exactly the same (3.3%) across centers (Table 3).

Evaluation of the composite measure

The composite measure predicted larger differences in future performance compared to the other quality measures (Figure). The spread from 3-star (top 20%) to 1-star (bottom 20%) hospitals in future risk-adjusted complication rates for each measure was as follows: hospital volume (3.3% to 3.2%), risk-adjusted complications (2.7% to 4.0%), risk- and reliability adjusted complications (2.4% to 4.1%), and the composite measure (2.4% to 4.6%). In the logistic regression model, that accounts for all patient characteristics, the composite measure

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

was the only quality indicator that predicted statistically significant differences between 1-star and 3-star hospitals (Table 3).

The composite measure also explained a higher proportion of hospital-level variation than the other quality measures (Table 4). This analysis confirmed the same trend shown above for discrimination among 1-star and 3-star hospitals. Hospital volume explained no variation (0%), risk-adjusted complications (28%) and risk- and reliability adjusted complications (47%) explained a larger fraction. However, the composite measure explained the largest proportion of hospital-level variation in serious complications (89%).

Discussion

In this paper, we demonstrate that a composite measure of bariatric surgery performance is superior to existing quality indicators at identifying the hospitals with the best outcomes. The composite measure described in this paper is created by combining multiple different outcomes (e.g., complications, reoperation, and prolonged length of stay), structural variables, and outcomes with other, related bariatric surgery procedures. We found that the composite measure was better t at predicting future performance and explained a higher proportion of hospital-level variation than the most widely used quality indicators for bariatric surgery.

Composite measures are becoming more widely used to profile hospital performance in other areas of surgery. For example, the Society of Thoracic Surgeons (STS), which maintains a clinical registry that captures nearly all cardiac surgery programs in the United States, uses a composite measure for profiling hospitals [15, 16]. The STS measure is a combination of processes of care and multiple outcomes (e.g., death and complications). The STS measure has one key conceptual difference compared to our approach – their measure is designed to reflect “global” quality across all domains of performance. Because of this objective, each domain (death, complications, and processes of care) is afforded equal weight. In contrast, our goal was to combine the quality signal from multiple measures to best predict a single “gold-standard” outcome, risk-adjusted serious complications[8].

In addition to the literature supporting the use of composite measures for assessing hospital performance, there is also a growing use of so-called “reliability adjustment”. These approaches address measurement problems in small hospitals by shrinking an observed outcome back towards the mean in the population. While there is consensus that this approach is better than using “noisy” outcome measures, there is a great deal of debate about which value should be the target for shrinkage[8, 17]. For example, with the Center for Medicare and Medicaid Services' (CMS) Hospital Compare measures for acute myocardial infarction, heart failure, and pneumonia the risk-adjusted mortality and readmission rates are shrunk towards the average rate for all hospitals. Many have challenged this approach because it assumes that small hospitals have average performance [17]. For any procedure or condition with a strong volume-outcome relationship, however, this assumption clearly does not hold.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The method used in this paper provides an alternative approach. Rather than making the assumption that all hospitals have average performance, we employ a flexible approach that incorporates information on hospital and surgeon characteristics including hospital volume. With such a strategy, the relationship between hospital characteristics and the outcome of interest is explicitly modeled and incorporated. For example, if hospital volume is an important predictor of outcomes, these methods would shrink the hospitals outcome rate towards the expected outcome for the hospitals volume group. Small hospitals with very low case loads, and very little “signal”, would therefore have an outcome rate very close to the rate at low volume hospitals. In the present study, hospital volume was not a strong predictor of complications and therefore did not receive much weight. However, in other studies of surgical procedures (e.g., esophageal and pancreatic cancer resections) hospital volume was an extremely important input to the composite measure [8, 9].

The present study has several limitations. First, our registry is limited to hospitals in Michigan, which may not be representative of the nation as a whole. Specifically, the relationship between hospital volume and outcomes may be stronger in states without a regional quality improvement program. With a national sample of hospitals, it is possible that our results would be different. However, this approach may change the relative weight placed on input measures, but it is unlikely to impact our main findings – that a composite measure that combines “signal” across measures is superior to individual measures alone. Second, the present study focuses on shortterm outcomes, such as perioperative safety. Quality in bariatric surgery in much more broad than short-term safety and should include measures of longer-term effectiveness, such as weight loss, comorbid disease resolution, and patient satisfaction. Unfortunately, however, long term follow up data is not widely available. Where it does exist it is often incomplete or inaccurate. Future accreditation efforts should emphasize the complete collection of these long term data. Finally, our study is limited because it includes a relatively small sample of hospitals. This sample size limitation prevents us from performing a bootstrapping (or resampling) analysis to directly compare the composite measure with the other measures. As such, our study should be viewed as preliminary and exploratory in nature, and should be replicated in a national cohort of bariatric surgery patients.

The findings of this study have important policy implications. Bariatric surgery is one of only a few surgical procedures for which accreditation is currently linked to insurance coverage. In 2006, CMS issued a national coverage decision for bariatric surgery that limited payment for surgery performed in hospitals accredited by the American College of Surgeons (ACS) and the American Society for Metabolic and Bariatric Surgery (ASMBS). Many private payers have since linked coverage or created tiered networks that steer patients towards ACS and ASMBS accredited centers. However, recent evidence from both clinical registry and administrative data has shown that accredited centers do not have better performance when compared to non-accredited centers. Better measures of hospital quality are needed to ensure that selective referral efforts of CMS and private payers are having the intended effect—steering patients to safer hospitals. More reliable measures are also needed for benchmarking and quality improvement in bariatric surgery. ACS and ASMBS are currently moving away from the “center of excellence” model and developing a national outcomes feedback program. Reliable outcome measures are needed to give hospitals and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

surgeons a true sense of where they stand compared to their peers. This study provides preliminary data that empirically weighted composite outcomes measures may be better than existing alternatives for both selective referral and outcomes feedback programs.

Acknowledgments

Disclosures: Drs. Dimick, Staiger, and Birkmeyer are consultants and have an equity interest in ArborMetrix, Inc, a venture-capital backed company that provides software and analytics for measuring and improving hospital quality and efficiency. The company had no role in the study herein.

Funding: This study was supported by a career development award to Dr. Dimick from the Agency for Healthcare Research and Quality (K08 HS017765) and a research grant from the National Institute of Diabetes and Digestive and Kidney Diseases (R21 DK084397). The views expressed herein do not necessarily represent the views of the United States Government.

References

1. Pratt GM, McLees B, Pories WJ. The ASBS Bariatric Surgery Centers of Excellence program: a blueprint for quality improvement. Surg Obes Relat Dis. 2006; 2(5):497–503. discussion 503. [PubMed: 17015199]
2. Schirmer B, Jones DB. The American College of Surgeons Bariatric Surgery Center Network: establishing standards. Bull Am Coll Surg. 2007; 92(8):21–7. [PubMed: 17715581]
3. Birkmeyer NJ, et al. Hospital complication rates with bariatric surgery in Michigan. JAMA. 2010; 304(4):435–42. [PubMed: 20664044]
4. Livingston EH. Bariatric surgery outcomes at designated centers of excellence vs nondesignated programs. Arch Surg. 2009; 144(4):319–25. discussion 325. [PubMed: 19380644]
5. Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA. 2004; 292(7):847–51. [PubMed: 15315999]
6. Dimick JB, Welch HG. The zero mortality paradox in surgery. J Am Coll Surg. 2008; 206(1):13–6. [PubMed: 18155563]
7. Drye EE, Chen J. Evaluating quality in small-volume hospitals. Arch Intern Med. 2008; 168(12): 1249–51. [PubMed: 18574078]
8. Dimick JB, et al. Composite measures for rating hospital quality with major surgery. Health Serv Res. 2012; 47(5):1861–79. [PubMed: 22985030]
9. Dimick JB, et al. Composite measures for predicting surgical mortality in the hospital. Health Aff (Millwood). 2009; 28(4):1189–98. [PubMed: 19597221]
10. Finks JF, et al. Predicting risk for venous thromboembolism with bariatric surgery: results from the Michigan Bariatric Surgery Collaborative. Ann Surg. 2012; 255(6):1100–4. [PubMed: 22566018]
11. Finks JF, et al. Predicting risk for serious complications with bariatric surgery: results from the Michigan Bariatric Surgery Collaborative. Ann Surg. 2011; 254(4):633–40. [PubMed: 21897200]
12. Dimick JB, Staiger DO, Birkmeyer JD. Ranking hospitals on surgical mortality: the importance of reliability adjustment. Health Serv Res. 2010; 45(6 Pt 1):1614–29. [PubMed: 20722747]
13. Dimick JB, et al. Reliability adjustment for reporting hospital outcomes with surgery. Ann Surg. 2012; 255(4):703–7. [PubMed: 22388108]
14. Staiger DO, et al. Empirically derived composite measures of surgical performance. Med Care. 2009; 47(2):226–33. [PubMed: 19169124]
15. Peterson ED, et al. ACCF/AHA 2010 Position Statement on Composite Measures for Healthcare Performance Assessment: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures (Writing Committee to develop a position statement on composite measures). Circulation. 2010; 121(15):1780–91. [PubMed: 20351232]
16. O'Brien SM, et al. Quality measurement in adult cardiac surgery: part 2--Statistical considerations in composite measure scoring and provider rating. Ann Thorac Surg. 2007; 83(4 Suppl):S13–26. [PubMed: 17383406]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

17. Silber JH, et al. The Hospital Compare Mortality Model and the Volume-Outcome Relationship. Health Services Research. 2010; 45(5):1148–1167. [PubMed: 20579125]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

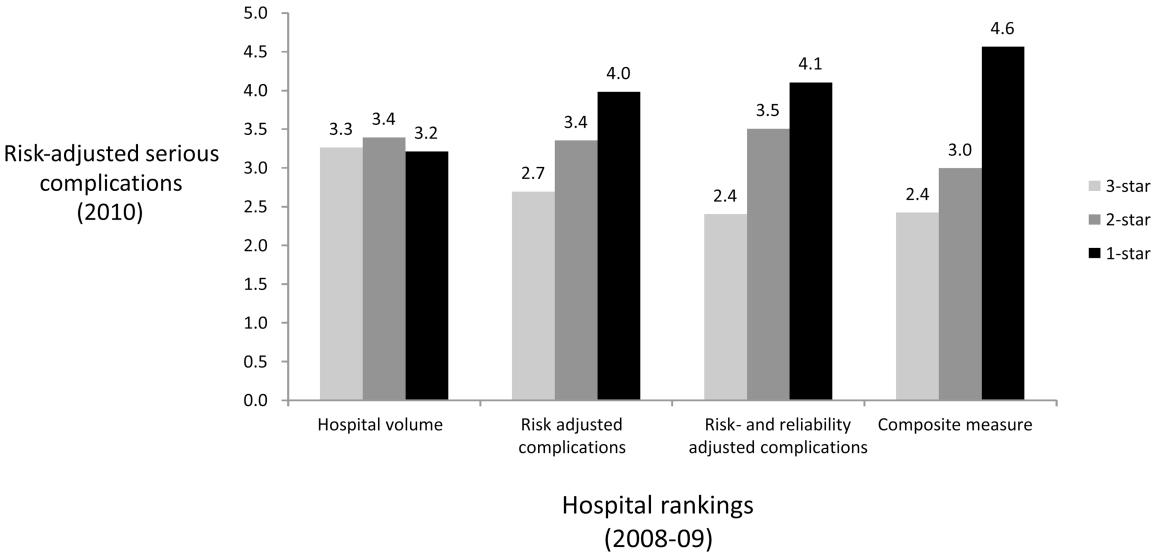

Figure.

Future risk-adjusted mortality rates (2010) for 3-star (top 20%), 2-star (middle 60%), 1-star (bottom 20%) hospitals as assessed using hospital volume, risk-adjusted complications, risk- and reliability adjusted complications, and the composite measure from the prior year (2008).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Individual hospital quality indicators are shown, along with the proportion of non-random hospital-level variation in laparoscopic gastric bypass complications explained by each.

Hospital quality indicators (2008-09)	Proportion of hospital-level variation explained (2008-09)
Hospital volume	11%
Prolonged length of stay	65%
Readmission	30%
Emergency department revisit rate	35%
Reoperation	45%
Complication rates with sleeve gastrectomy	10%
Complication rates with gastric banding	40%

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Weights placed each individual quality indicator in the composite measure of performance for laparoscopic gastric bypass.

Hospital quality indicators	Weight afforded in the composite measure
Expected complication rate given hospital structural characteristics and overall average complication rate	39%
Complication rate with laparoscopic gastric bypass	25%
Readmission with laparoscopic gastric bypass	18%
Prolonged length of stay with laparoscopic gastric bypass	10%
Emergency department revisit rate with laparoscopic gastric bypass	6%
Reoperation with laparoscopic gastric bypass	6%
Complication rates with sleeve gastrectomy	2%
Complication rates with gastric banding	2%

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Patient characteristics for the 1-star, 2-star, and 3-star hospitals according to hospital rankings created using the composite measure.

Patient characteristics	3-star hospitals N=1,033	2-star hospitals N=968	1-star hospitals N=941
Expected complication rate (mean)	3.3%	3.3%	3.3%
Age >50 (%)	37%	42%	36%
Male (%)	21%	21%	21%
Body Mass Index (mean)	47.7	49.4	48.6
History of thromboembolism (%)	3.2%	3.9%	3.6%
Smoking history (%)	38%	42%	46%
Pulmonary disease (%)	34%	28%	26%
Cardiac disease (%)	6.4%	5.9%	5.8%
Mobility limitations (%)	9.0%	5.5%	8.8%
Sleep apnea (%)	50%	50%	51%
Hypertension (%)	65%	57%	57%
Diabetes (%)	41%	35%	36%
Hyperlipidemia (%)	57%	59%	55%
Liver disease (%)	3.0%	5.3%	2.6%
Gastroesophageal reflux disease (%)	56%	48%	57%
Peptic ulcer disease (%)	5.3%	1.9%	4.3%
Musculoskeletal disorder (%)	80%	77%	82%
Prior hernia repair (%)	2.8%	2.2%	1.6%

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Relative ability of different quality measures from 2008-09 to discriminate future (2010) risk-adjusted complications and explain variation in hospital-level complication rates.

	Risk-adjusted complications in 2010		
Hospital Rankings (2008-09)	**Odds ratio (95% CI)**	**Odds ratio (95% CI)**	**Proportion of hospital-level variation explained**
Quality measure	**2-star vs. 3-star hospitals**	**1-star vs. 3 star hospitals**	
Hospital volume	0.93 (0.47 to 1.85)	0.85 (0.43 to 1.68)	0%
Risk-adjusted complication rates	1.27 (0.66 to 2.44)	1.56 (0.84 to 2.91)	28%
Risk- and reliability adjusted complication rates	1.37 (0.71 to 2.61)	1.65 (0.88 to 3.10)	47%
Composite measure	1.26 (0.7 to 2.25)	1.99 (1.14 to 3.47)	89%