SPECIAL ARTICLE

Surgical Skill and Complication Rates after Bariatric Surgery

John D. Birkmeyer, M.D., Jonathan F. Finks, M.D., Amanda O'Reilly, R.N., M.S., Mary Oerline, M.S., Arthur M. Carlin, M.D., Andre R. Nunn, M.D., Justin Dimick, M.D., M.P.H., Mousumi Banerjee, Ph.D., and Nancy J.O. Birkmeyer, Ph.D., for the Michigan Bariatric Surgery Collaborative

ABSTRACT

BACKGROUND

Clinical outcomes after many complex surgical procedures vary widely across hospitals and surgeons. Although it has been assumed that the proficiency of the operating surgeon is an important factor underlying such variation, empirical data are lacking on the relationships between technical skill and postoperative outcomes.

METHODS

We conducted a study involving 20 bariatric surgeons in Michigan who participated in a statewide collaborative improvement program. Each surgeon submitted a single representative videotape of himself or herself performing a laparoscopic gastric bypass. Each videotape was rated in various domains of technical skill on a scale of 1 to 5 (with higher scores indicating more advanced skill) by at least 10 peer surgeons who were unaware of the identity of the operating surgeon. We then assessed relationships between these skill ratings and risk-adjusted complication rates, using data from a prospective, externally audited, clinical-outcomes registry involving 10,343 patients.

RESULTS

Mean summary ratings of technical skill ranged from 2.6 to 4.8 across the 20 surgeons. The bottom quartile of surgical skill, as compared with the top quartile, was associated with higher complication rates (14.5% vs. 5.2%, $P<0.001$) and higher mortality (0.26% vs. 0.05%, $P=0.01$). The lowest quartile of skill was also associated with longer operations (137 minutes vs. 98 minutes, $P<0.001$) and higher rates of reoperation (3.4% vs. 1.6%, $P=0.01$) and readmission (6.3% vs. 2.7%) ($P<0.001$).

CONCLUSIONS

The technical skill of practicing bariatric surgeons varied widely, and greater skill was associated with fewer postoperative complications and lower rates of reoperation, readmission, and visits to the emergency department. Although these findings are preliminary, they suggest that peer rating of operative skill may be an effective strategy for assessing a surgeon's proficiency.

From the Center for Healthcare Outcomes and Policy and Department of Surgery, University of Michigan, Ann Arbor (J.D.B., J.F.F., A.O., M.O., J.D., M.B., N.J.O.B.), the Department of Surgery, Henry Ford Health System, Detroit (A.M.C.), and Forest Health Medical Center, Ypsilanti (A.R.N.) — all in Michigan. Address reprint requests to Dr. Birkmeyer at the Center for Healthcare Outcomes and Policy, University of Michigan, N. Campus Research Complex, 2800 Plymouth Rd., Bldg. 16-136E, Ann Arbor, MI 48109, or at jbirkmey@umich.edu.

N Engl J Med 2013;369:1434-42.
DOI: 10.1056/NEJMsa1300625

A CONSIDERABLE BODY OF RESEARCH SUGgests that some surgeons have better results than others. Early studies of coronary-artery bypass surgery showed wide variation in risk-adjusted patient mortality across surgeons; studies of other procedures and other outcomes have shown similar variation among surgeons.[1-3] Efforts to reduce such variation have focused primarily on improving perioperative care. For example, the Surgical Care Improvement Project and related pay-for-performance programs have provided financial incentives to increase surgeons' compliance with evidence-based practices related to prophylaxis against surgical-site infection and venous thromboembolism. As of this writing, however, there is little evidence that such initiatives have improved outcomes overall or have reduced the variation in outcomes across surgeons.[4-6]

In many procedures, the technical skill of the operating surgeon may be a more important determinant of outcomes than perioperative care. A high level of surgical skill may be essential in preventing intraoperative problems such as bleeding or tissue devascularization and may be associated with more precise reconstruction in cardiovascular or gastrointestinal surgery, possibly reducing the risk of anastomotic complications (e.g., thrombosis and leak, respectively). A high level of skill may also be associated with shorter operations, which are important in light of research linking prolonged operating times to increased risks of certain types of complications such as infection and venous thromboembolism.[7-12] The importance of surgical skill may be inferred from studies assessing potential surrogate variables, including procedure volume and subspecialty training. As of this writing, however, few studies have directly assessed the technical skill of practicing surgeons, and to our knowledge none have linked the level of surgical skill to clinical outcomes.

We conducted a population-based study of complications after gastric bypass surgery, a common but complex procedure for which outcomes have been shown to vary widely according to the surgeon. We first rated the technical skill of 20 practicing bariatric surgeons, as judged anonymously by their peers. We then examined the relationship between the surgeons' technical skill and their risk-adjusted complication rates.

METHODS

STUDY OVERVIEW

This study was based on an analysis of data from the Michigan Bariatric Surgery Collaborative (MBSC), a payer-funded consortium of hospitals and surgeons performing bariatric surgery in Michigan.[13,14] The MBSC includes all 40 Michigan hospitals engaged in bariatric surgery and all 75 surgeons operating at those facilities. Participating hospitals submit data on all patients undergoing bariatric surgery (approximately 6000 new patients each year).

As described in greater detail elsewhere, the MBSC clinical registry includes a wide range of information on demographic variables, coexisting conditions, the perioperative process of care, complications at 30 days, and longer-term outcomes, including weight loss and health status.[15] All variable definitions are standardized, and data collection is conducted by centrally trained data abstractors. Each participating hospital is visited annually by external auditors from the central coordinating center to verify the accuracy and completeness of the submitted data.

The objective of this study was to assess the relationship between the technical skill of 20 MBSC surgeons and postoperative complications in 10,343 patients undergoing laparoscopic gastric bypass between August 28, 2006 (the start date of the MBSC registry), and August 1, 2012. The study was approved by the institutional review board of the University of Michigan.

PARTICIPATING SURGEONS AND RATERS

We recruited 20 surgeons from a total of 63 Michigan surgeons who were performing laparoscopic gastric bypass when the study was initiated. Participation was voluntary and based entirely on self-selection. Participating surgeons were asked to submit one video of themselves performing laparoscopic gastric bypass. In addition to being the most common bariatric procedure, laparoscopic gastric bypass is relatively complex and is associated with higher complication rates than gastric banding and sleeve gastrectomy. Surgeons were asked to submit a "representative" case (the choice of which was left entirely to them) that was stripped of any patient identifiers.

To minimize the burden on the raters, each of the 20 study videos was edited to remove all but the critical components of the procedure:

creation of the gastric pouch, the gastrojejunostomy, and the jejunojejunostomy. Edited videos, which ranged from approximately 25 to 40 minutes in length, were then distributed electronically for rating by peer surgeons. One video was distributed approximately every 2 weeks between July 2011 and June 2012. All MBSC surgeons other than the 20 surgeons who participated in the study were invited to rate each video. Follow-up reminders were sent until each video had been rated by at least 10 peers. Overall, 33 surgeons from 24 hospitals served as raters. Of these surgeons, 15 performed 78% of the ratings. Both the surgeon raters and the investigators were unaware of the identity of the surgeon in each video and the outcome of the procedure.

A video showing two levels of surgical skill is available at NEJM.org

Surgeons rated each video using an instrument modified slightly from the Objective Structured Assessment of Technical Skills,[16] which is used commonly in assessing the skill of surgical trainees[17] and was recently validated for use among practicing bariatric surgeons[18] (see the video, available with the full text of this article at NEJM.org). In addition to providing a summary judgment of overall skill, surgeon raters assessed each video in five domains of technical skill: gentleness, tissue exposure, instrument handling, time and motion, and flow of operation. Each domain of surgical skill was rated on a scale of 1 to 5 (with 1 indicating the skill expected of a general-surgery chief resident and 5 indicating the skill of a master bariatric surgeon). A score of 3 reflected the skill of an average practicing bariatric surgeon. Surgeon raters judged the videos and applied these rating criteria independently. We made no attempt to train the raters or otherwise establish rating norms.

OUTCOMES

The primary outcome was the occurrence of any postoperative complication. Surgical complications included a surgical-site infection (a wound infection treated with antibiotics, wound opening, or both), a wound infection or dehiscence (requiring reoperation), an abdominal abscess (requiring percutaneous drainage or reoperation), a leak (requiring percutaneous drainage or reoperation), an anastomotic stricture (requiring dilation), a bowel obstruction (requiring reoperation), and bleeding (requiring a blood transfusion, reoperation, or a splenectomy). Medical complications included pneumonia (requiring treatment with antibiotics only), respiratory failure (requiring intubation or tracheostomy), renal failure (requiring dialysis), venous thromboembolism (deep-vein thrombosis or pulmonary embolism), myocardial infarction, cardiac arrest, and death. We also examined 30-day rates of death, unplanned reoperation, readmission, and emergency department visits.

STATISTICAL ANALYSIS

Our exposure variable, technical skill, was determined at the surgeon level. For statistical modeling and inferences, outcomes were assessed at the patient level. All analyses accounted for clustering of outcomes from multiple patients treated by a single surgeon.

As our summary skill measure for each of the 20 study surgeons, we used the average rating of all peer surgeons who reviewed each video. In turn, each peer rating was determined according to a simple average of the six measures on the review instrument (one for each of the five domains of technical skill, plus an overall skill rating). The five domains of technical skill were highly correlated with one another and with the overall skill ratings, with correlation coefficients ranging from 0.85 to 0.97. For this reason, our main results were not affected by whether we used a simple average of component measures (our primary analysis) or alternative weighting schemes.

Since the pool of raters varied somewhat from video to video, we considered the possibility of rating bias because some raters were tougher graders than others. To explore this potential effect, we calculated average z scores for each peer rater who reviewed multiple videos. For each video, a z score was calculated as the difference between the score of each rater and the average score of all raters, divided by the standard deviation. By this measure, no rater's score was significantly different from the mean. In a sensitivity analysis, we repeated our main analysis while adjusting for raters' average z scores. These results were nearly identical to those of our main analysis and are not presented here.

For statistical inference, we used surgeon-specific ratings in assessing relationships between surgical skill and outcomes. For presentation purposes, however, surgeons were categorized into three groups according to quartiles of skill ratings: the bottom group (first quartile), middle group (second and third quartiles), and top group (fourth quartile). On the basis of models described and validated in our previous study,[19]

multivariate logistic regression was used for risk adjustment in assessing associations between surgical skill and the occurrence of complications. Variables used in this model included age, male sex, previous venous thromboembolism, mobility limitations, coronary artery disease, and pulmonary disease. In examining rates of specific complications, other covariates were added to the models on the basis of the results of forward stepwise logistic multivariable analyses ($P<0.05$ for the inclusion criterion). To account for the clustering of data from multiple patients treated by a single surgeon, we used generalized estimating equations to obtain robust sandwich estimates for the standard errors. In calculating risk-adjusted rates, we obtained the predicted probability of the events for each patient and summed these probabilities to get the expected number of events for each quartile of surgical skill. For each quartile, risk-adjusted rates were calculated as the overall crude rate of the event multiplied by the ratio of the number of observed events to the number of expected events (i.e., the observed-to-expected ratio).

In a sensitivity analysis (see the Supplementary Appendix, available at NEJM.org), we rated a video of a second operation from each surgeon in the best and worst quartiles of skill. As shown in Figure S1a in the Supplementary Appendix, mean ratings for the first and second videos were highly correlated ($R=0.85$, $P=0.002$). In a second sensitivity analysis, we asked five non-Michigan surgeons with no connection to the study surgeons to rate the original videos of surgeons in the lowest and highest skill quartiles. Perhaps not surprisingly, the non-Michigan surgeons were somewhat harsher judges, particularly of surgeons in the bottom quartile. Nonetheless, mean ratings from Michigan and non-Michigan surgeons were highly correlated ($R=0.94$, $P<0.001$), and there was no overlap among mean ratings of surgeons in the lowest and highest quartiles (Fig. S2a in the Supplementary Appendix). Our main conclusions about relationships between surgical skill and outcomes were unchanged by these additional data and analyses.

RESULTS

SKILL RATINGS AND CHARACTERISTICS OF SURGEONS AND PATIENTS

Summary ratings of the technical skill of the 20 surgeons varied substantially, from 2.6 to 4.8 (Fig. 1). The 5 surgeons in the bottom quartile had a mean rating of 2.9, as compared with a mean rating of 4.4 for the 5 surgeons in the top quartile. Ratings in each of the five domains of technical skill varied to a similar degree, as did peers' ratings of each surgeon's overall technical skill (Table 1).

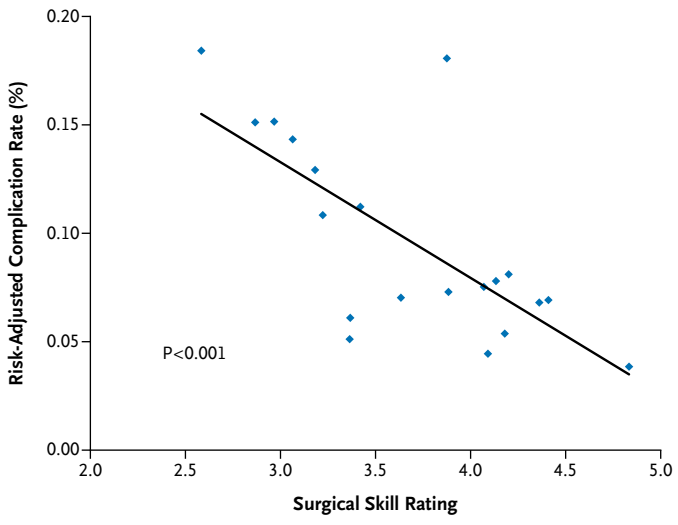

Figure 1. Relationship between Summary Peer Rating of Technical Skill and Risk-Adjusted Complication Rates after Laparoscopic Gastric Bypass.
Each diamond in the scatter plot represents 1 of 20 practicing bariatric surgeons.

Surgical skill was not related to years in bariatric surgery practice, status with respect to completion of a fellowship in advanced laparoscopy or bariatric surgery, or practice at a teaching or nonteaching hospital. Surgical skill was, however, strongly related to procedure volume (Table 1). As compared with surgeons in the top quartile of summary skill ratings, surgeons in the bottom quartile had lower mean annual volumes of laparoscopic gastric bypass procedures (53 vs. 157, $P=0.005$) and of any bariatric procedure (106 vs. 241, $P=0.02$). Surgeons in the bottom quartile of skill, as compared with those in the top quartile, also had longer average operating times for laparoscopic gastric bypass (137 minutes vs. 98 minutes, $P<0.001$).

There were no clinically important differences in patient age, sex, or body-mass index or in the rate of private insurance coverage across quartiles of surgical skill (Table 2). There were several statistically significant differences in the prevalence of specific coexisting conditions, but there was no consistent relationship between

Table 1. Characteristics of Surgeons, Patient Volume, and Surgery, According to Peer Rating of Surgical Skill.*

Variable	Level of Surgical Skill			P Value
	Quartile 1	Quartile 2 or 3	Quartile 4	
Surgeons (no.)	5	10	5	
Mean peer rating of technical skill*				
Gentleness	3.3	3.9	4.4	
Time and motion	2.6	3.4	4.3	
Instrument handling	2.9	3.7	4.4	
Flow of operation	3.1	3.8	4.5	
Tissue exposure	3.0	3.9	4.4	
Overall technical skill	2.7	3.6	4.4	
Summary rating	2.9	3.7	4.4	
Patients (no.)				
Undergoing laparoscopic gastric bypass	1459	4520	4364	
Undergoing any bariatric procedure	2978	7491	6665	
Mean annual procedure volume (no.)†				
Laparoscopic gastric bypass procedures	53	96	157	0.005
Any bariatric procedure	106	155	241	0.02
Mean operating room times (min)‡				
Laparoscopic gastric bypass procedures	137	123	98	<0.001
Any bariatric procedure	110	111	85	<0.001
Mean duration of bariatric surgery practice (yr)	11	9	11	0.44
Completion of fellowship in advanced laparoscopy or bariatric surgery (%)	20	44	20	0.56
Practicing at teaching hospital (%)	60	70	40	0.63

* Each domain of surgical skill was rated on a scale of 1 to 5 (with 1 indicating the skill expected of a general-surgery chief resident and 5 indicating the skill of a master bariatric surgeon). A score of 3 reflected the skill of an "average" practicing bariatric surgeon. The summary rating was the mean of the ratings for the six domains.
† The difference across the quartiles was significant (P=0.005 for laparoscopic gastric bypass procedures and P=0.02 for any bariatric procedure).
‡ The difference across the quartiles was significant (P<0.001 for both procedures).

surgical skill and these conditions. For example, among patients treated by surgeons in the bottom quartile, as compared with those treated by surgeons in the top quartile, there was a higher prevalence of cardiovascular disease, sleep apnea, and mobility disorders but a lower prevalence of smoking and musculoskeletal disorders.

SURGICAL SKILL AND COMPLICATION RATES

Technical skill was related to overall complication rates associated with individual surgeons (Fig. 1). The bottom quartile of technical skill was associated with higher overall complication rates than the top quartile (14.5% vs. 5.2%, P<0.001) (Fig. 2). Patients treated by each of the five surgeons in the bottom quartile of skill had complication rates that were higher than the overall average. Lower skill ratings were associated with higher rates of both surgical and medical complications. Among specific complications, the bottom quartile of technical skill, as compared with the top quartile, was associated with significantly higher rates of surgical-site infections (4.60% vs. 1.04%, P=0.001) and pulmonary complications (3.91% vs. 0.71%, P=0.004) (Table 3). The bottom skill quartile, as compared with the top quartile, was also associated with higher mortality (0.26% vs. 0.05%, P=0.01) (Table 3).

Finally, surgical skill was associated with rates of reoperation, readmission, and return visits to the emergency department. As compared with the top quartile of skill, the bottom quartile was

Table 2. Characteristics of Patients Undergoing Laparoscopic Gastric Bypass, According to Peer Rating of Surgical Skill.

Patient Characteristic	Level of Surgical Skill			
	Quartile 1 (5 surgeons, 1459 patients)	Quartile 2 or 3 (10 surgeons, 4520 patients)	Quartile 4 (5 surgeons, 4364 patients)	P Value
Demographic				
Mean age (yr)	44.6	46.7	45.1	<0.001
Male sex (%)	18.8	21.2	20.2	0.13
Private insurance (%)	66.4	69.0	67.0	0.06
Clinical				
Mean body-mass index*	49.2	48.9	48.5	0.003
Medical history (%)				
Musculoskeletal disorder	77.3	78.0	90.0	<0.001
Cardiovascular disease	62.4	57.2	53.0	<0.001
Hyperlipidemia	57.0	54.1	52.1	0.004
Gastroesophageal reflux disease	57.3	46.0	46.7	<0.001
Psychological conditions	46.5	49.4	47.1	0.04
Sleep apnea	54.3	46.1	43.4	<0.001
Smoking	37.0	45.9	42.4	<0.001
Diabetes	37.2	38.1	35.1	0.01
Cholelithiasis	25.0	29.8	26.7	<0.001
Lung disease	34.5	25.0	28.5	<0.001
Urinary incontinence	20.7	21.4	19.1	0.03
Mobility problems	8.3	5.8	4.1	<0.001
Liver disorder	2.9	6.0	5.0	<0.001
Venous thromboembolism	4.8	4.4	3.5	0.04
Peptic ulcer disease	5.8	3.3	2.1	<0.001

* The body-mass index is the weight in kilograms divided by the square of the height in meters.

associated with significantly higher rates of reoperation (3.4% vs. 1.6%, P=0.01), readmission within 30 days (6.3% vs. 2.7%, P<0.001), and return visits to the emergency department (10.2% vs. 3.8%, P=0.004) (Fig. 3).

DISCUSSION

In this preliminary study, 20 practicing bariatric surgeons in Michigan agreed to have their technical skills rated anonymously by their peers. Although very few were rated as having below-average skill, there was nonetheless a wide range of technical-skill scores between the highest and lowest ratings. Previous studies of surgical skill, most of which were from the graduate medical education field, focused primarily on methods of measuring and discriminating among the skills of surgical trainees.[16,17,20-22] Such research has shown that trainees, as a group, have lower technical-skill ratings than attending surgeons, a finding that is often considered to be the standard for validation of instruments used to rate surgical skill. Other research has described the "learning curve" after surgical training.[23] In contrast, our study showed wide variation in technical skill among fully trained, practicing surgeons.

The clinical importance of such variation is evident in our second major finding — that surgical skill is a strong predictor of clinical outcomes. As compared with patients treated by surgeons with high skill ratings, patients treated by surgeons with low skill ratings were at least twice as likely to die, have complications, undergo reoperation, and be readmitted after hospital discharge. Although our study shows the impor-

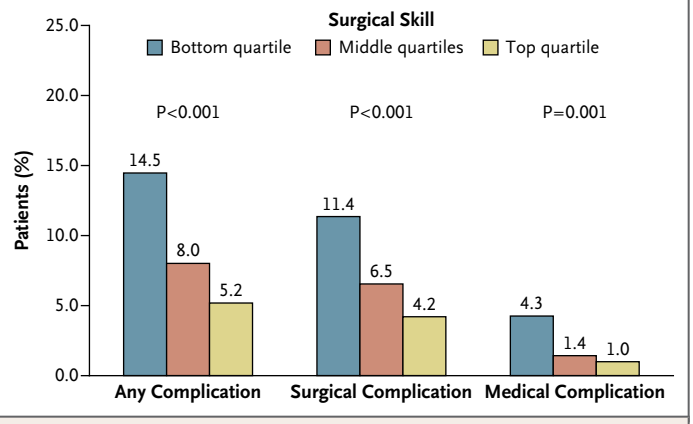

Figure 2. Risk-Adjusted Complication Rates with Laparoscopic Gastric Bypass, According to Quartile of Surgical Skill.

tance of surgical skill directly, the extensive literature on the relationship between procedure volume and clinical outcomes perhaps presages our findings. Although surgical volume is less important for simple procedures, for complex procedures, including cardiovascular surgery, certain cancer resections, and bariatric surgery, higher surgical volume has been associated with lower morbidity and mortality.[15,24,25] Procedure volume may be linked to other delivery-system factors associated with surgical outcomes, but it has been assumed that volume is important in large part because it serves as a proxy for operative proficiency. Thus, our finding that surgeons with high skill ratings had higher caseloads than surgeons with low skill ratings is probably not a coincidence.

Table 3. Risk-Adjusted Rates of Complications after Laparoscopic Gastric Bypass, According to Peer Rating of Surgical Skill.

Variable	Level of Surgical Skill			
	Quartile 1	Quartile 2 or 3	Quartile 4	P Value
		percent		
Surgical complications				
Leak or perforation	0.69	0.66	0.56	0.44
Obstruction	4.75	1.93	1.61	0.01
Infection	4.60	2.27	1.04	0.001
Hemorrhage	2.93	2.31	1.98	0.17
Medical complications				
Venous thromboembolism	0.26	0.49	0.28	0.81
Cardiac complication	0.06	0.13	0.10	0.98
Renal failure	0.33	0.10	0.10	0.07
Pulmonary complication	3.91	0.96	0.71	0.004
Death	0.26	0.09	0.05	0.01

On clinical grounds, few surgeons would be surprised that technical skill is an important determinant of outcomes in patients who have undergone laparoscopic gastric bypass. The procedure is technically complex and performed in patients with morbid obesity, for whom surgical exposure is often challenging. Moreover, a high proportion of complications after gastric bypass, including anastomotic problems, bleeding, and bowel obstruction, occur at the surgical site and thus are directly related to operative technique. The link to surgical skill may be less obvious in the case of other complications, including infection and complications that do not occur at the surgical site. However, previous studies have suggested that rates of such complications increase with longer operating times and a longer duration of general anesthesia.[7-12] In our study, operations performed by surgeons in the bottom quartile of technical skill were approximately 40% longer than those performed by surgeons in the top quartile. Such findings suggest that the technical skill of surgeons may affect clinical outcomes in many ways.

Our findings should be considered in light of several limitations of the study, particularly those related to our measurement of surgical skill. The participation of surgeons was voluntary, and their skill was rated on the basis of a single, self-selected video operation. Although a small group of raters reviewed the majority of videos, no two surgeons were rated by exactly the same set of peer reviewers. In addition, we made no attempt to train the peer reviewers and calibrate the very subjective process of judging a surgeon's technical skill. Together, such limitations would tend to decrease the reliability of our skill measures, misclassify surgeons' true skill, and ultimately bias our findings toward the null hypothesis (i.e., no effect of surgical skill on outcomes). Stated differently, better methods for measuring surgical skill, as reviewed elsewhere,[17,20-22] might have strengthened our main conclusion that skill is a powerful predictor of clinical outcomes.

We also did not examine the extent to which skills ratings could be influenced by other members of the surgical team. Although the operating surgeon does most of the work during the

critical components of a laparoscopic gastric bypass, the extent to which the first assistant makes the operation easier could influence the rating of the operating surgeon's skill. Similarly, we did not examine the role, if any, played by surgical residents in each case. It is worth noting, however, that surgeons at teaching hospitals were represented in both the top and bottom quartiles of skill.

Our findings may have direct implications for organizations charged with ensuring the proficiency of practicing surgeons, including the surgical subspecialty boards of the American Board of Medical Specialties (ABMS). The limitations of traditional methods for ensuring physician competence, including continuing medical education, are widely recognized.[26] In 2011, the ABMS introduced stricter standards for maintenance of board certification, including requirements that surgeons systematically track their performance. Although requirements for maintenance of certification are not currently tied to specific performance standards, they could move toward that goal as clinical registries evolve and support more rigorous benchmarking in the various surgical specialties. Even if current practical barriers were overcome, however, outcomes-based assessment of the proficiency of surgeons would be hindered by concerns about the imprecision of measures and risk adjustment. Although our findings would first need to be confirmed across other procedures and specialties, they suggest that peer assessment of a surgeon's operative skill may be a more practical, more direct, and ultimately more informative test for assessing the surgeon's proficiency than other measures.

Variation in surgical skill and outcomes may never be eliminated. As with musicians or athletes, some surgeons may simply be more talented than others. At the same time, coaching and deliberate practice are mainstays in both music and athletics, and they are considered to be effective in improving proficiency among persons at any skill level. The feasibility and potential benefits of one-on-one surgeon coaching have been described.[27,28] In Michigan, bariatric surgeons now watch each other operate during site visits to other programs and can watch online the videos of surgeons with superior skill and outcomes. They will soon be receiving anonymous, constructive feedback from their peers on strategies for refining their technique. The effectiveness of such strategies for enhancing operative proficiency remains to be established. To the extent that practicing surgeons are still "teachable," however, our findings suggest a potential opportunity for improving surgical outcomes.

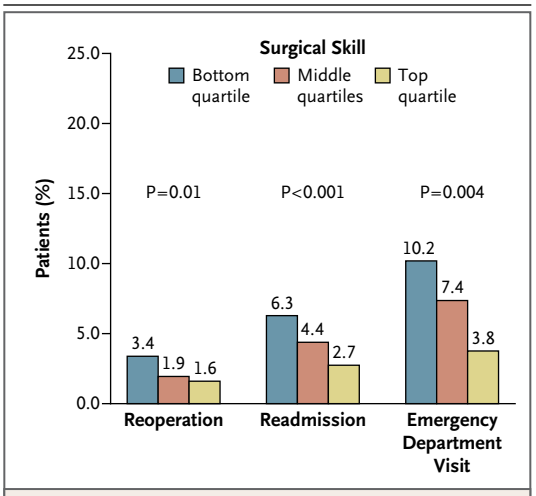

Figure 3. Rates of Reoperation, Readmission, and Visits to the Emergency Department after Laparoscopic Gastric Bypass, According to Quartile of Surgical Skill.

Supported by Blue Cross Blue Shield of Michigan and Blue Care Network.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

REFERENCES

1. Hall BL, Hamilton BH. New information technology systems and a Bayesian hierarchical bivariate probit model for profiling surgeon quality at a large hospital. Q Rev Econ Finance 2004;44:410-29.
2. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State: an analysis of risk factors and hospital mortality rates. JAMA 1990;264:2768-74.
3. O'Connor GT, Plume SK, Olmstead EM, et al. A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. JAMA 1991;266:803-9.
4. Hawn MT, Vick CC, Richman J, et al. Surgical site infection prevention: time to move beyond the Surgical Care Improvement Program. Ann Surg 2011;254:494-9.
5. Nicholas LH, Osborne NH, Birkmeyer JD, Dimick JB. Hospital process compliance and surgical outcomes in Medicare beneficiaries. Arch Surg 2010;145:999-1004.
6. Stulberg JJ, Delaney CP, Neuhauser DV, Aron DC, Fu P, Koroukian SM. Adherence to Surgical Care Improvement Project measures and the association with postoperative infections. JAMA 2010;303:2479-85.
7. Campbell DA Jr, Henderson WG, Englesbe MJ, et al. Surgical site infection prevention: the importance of operative duration and blood transfusion — results of the first American College of Surgeons–National Surgical Quality Improvement Program Best Practices Initiative. J Am Coll Surg 2008;207:810-20.
8. Chan MM, Hamza N, Ammori BJ. Duration of surgery independently influenc-

es risk of venous thromboembolism after laparoscopic bariatric surgery. Surg Obes Relat Dis 2013;9:88-93.
9. Kessler S, Kinkel S, Käfer W, Puhl W, Schochat T. Influence of operation duration on perioperative morbidity in revision total hip arthroplasty. Acta Orthop Belg 2003;69:328-33.
10. Leong G, Wilson J, Charlett A. Duration of operation as a risk factor for surgical site infection: comparison of English and US data. J Hosp Infect 2006;63:255-62.
11. Procter LD, Davenport DL, Bernard AC, Zwischenberger JB. General surgical operative duration is associated with increased risk-adjusted infectious complication rates and length of hospital stay. J Am Coll Surg 2010;210:60-5.
12. Tan T-W, Kalish JA, Hamburg NM, et al. Shorter duration of femoral-popliteal bypass is associated with decreased surgical site infection and shorter hospital length of stay. J Am Coll Surg 2012;215:512-8.
13. Birkmeyer NJO, Share D, Campbell DA Jr, Prager RL, Moscucci M, Birkmeyer JD. Partnering with payers to improve surgical quality: the Michigan plan. Surgery 2005;138:815-20.
14. Share DA, Campbell DA, Birkmeyer N, et al. How a regional collaborative of hospitals and physicians in Michigan cut costs and improved the quality of care. Health Aff (Millwood) 2011;30:636-45.
15. Birkmeyer NJ, Dimick JB, Share D, et al. Hospital complication rates with bariatric surgery in Michigan. JAMA 2010;304:435-42.
16. Martin JA, Regehr G, Reznick R, et al. Objective Structured Assessment of Technical Skill (OSATS) for surgical residents. Br J Surg 1997;84:273-8.
17. Reznick RK, MacRae H. Teaching surgical skills — changes in the wind. N Engl J Med 2006;355:2664-9.
18. Zevin B, Bonrath EM, Aggarwal R, Dedy NJ, Ahmed N, Grantcharov TP. Development, feasibility, validity, and reliability of a scale for objective assessment of operative performance in laparoscopic gastric bypass surgery. J Am Coll Surg 2013;216:955-65.
19. Finks JF, Kole KL, Yenumula PR, et al. Predicting risk for serious complications with bariatric surgery: results from the Michigan Bariatric Surgery Collaborative. Ann Surg 2011;254:633-40.
20. Gofton WT, Dudek NL, Wood TJ, Balaa F, Hamstra SJ. The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence. Acad Med 2012;87:1401-7.
21. van Hove PD, Tuijthof GJM, Verdaasdonk EGG, Stassen LPS, Dankelman J. Objective assessment of technical surgical skills. Br J Surg 2010;97:972-87.
22. Vassiliou MC, Feldman LS, Fraser SA, et al. Evaluating intraoperative laparoscopic skill: direct observation versus blinded videotaped performances. Surg Innov 2007;14:211-6.
23. Carty MJ, Chan R, Huckman R, Snow D, Orgill DP. A detailed analysis of the reduction mammaplasty learning curve: a statistical process model for approaching surgical performance improvement. Plast Reconstr Surg 2009;124:706-14.
24. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. N Engl J Med 2003;349:2117-27.
25. Livingston EH. Relationship between surgeon volume and adverse outcomes after Roux-en-Y gastric bypass in Longitudinal Assessment of Bariatric Surgery (LABS) study. Surg Obes Relat Dis 2010;6:587-8.
26. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? JAMA 1999;282:867-74.
27. Gawande A. Personal best. The New Yorker. October 3, 2011:44-53.
28. Hu Y-Y, Peyre SE, Arriaga AF, et al. Postgame analysis: using video-based coaching for continuous professional development. J Am Coll Surg 2012;214:115-24.