Are Patient-reported Outcomes Correlated With Clinical Outcomes After Surgery?

A Population-based Study

Jennifer F. Waljee, MD, MS, Amir Ghaferi, MD,† Ruth Cassidy, MS,‡ Oliver Varban, MD,§ Jonathan Finks, MD,¶ Kevin C. Chung, MD,|| Noelle E. Carlozzi, PhD,** and Justin B. Dimick, MD, MPH††*

Objective: To evaluate the extent to which patient-reported outcomes (PROs) (eg, health-related quality of life) are distinct from clinical outcomes following bariatric surgery.

Background: Hospital quality measurement often focuses on traditional clinical outcomes (eg, complications). However, PROs may provide a unique perspective regarding performance, particularly for common, low-risk procedures.

Methods: We used data from 11,420 patients who underwent bariatric surgery (2008–2012) from the Michigan Bariatric Surgery Collaborative (39 hospitals). We included both short-term (30-day complication rates) and long-term (1-year weight loss and comorbidity resolution) outcomes. For PROs, we used health-related quality of life assessed by the Health and Activities Limitations Index (HALex) and Bariatric Quality of Life (BQL) index preoperatively and at 1 year. We used multivariable linear regression to determine the association between these PROs and both short and long-term clinical outcomes, adjusting for patient factors and the type of surgical procedure.

Results: After adjustment for risk factors and surgical procedure, hospital rankings based on PROs (either the average change in HALex or BQL scores) were not correlated with hospital rankings based on complications. In contrast, both PRO measures were correlated with weight loss. Specifically, the average change in HALex score ($R^2 = 0.24$, $P < 0.002$) and average change in BQL score ($R^2 = 0.44$, $P < 0.001$) were correlated with hospital average percent excess. One PRO measure—BQL score—was correlated with a decline in the need for medications due to associated comorbidities ($R^2 = 0.16$, $P < 0.01$). After accounting for short and long-term clinical outcomes, between 15% and 44% of the variation in PROs remained unexplained at the hospital level.

Conclusions: Patient-reported outcomes are not correlated with early perioperative events, but are correlated with measures of clinical effectiveness after bariatric surgery. A comprehensive approach to surgical quality should incorporate both clinical events and self-reported measures of health status throughout the short and long-term recovery period.

Keywords: bariatric surgery, patient-reported outcomes, quality of life

(*Ann Surg* 2016;264:682–689)

Clinical outcomes, such as perioperative morbidity and mortality, are commonly used to benchmark hospital performance.[1–4] Perioperative events are easily captured in clinical registries, have high face validity for surgeons in practice, and can be quantified and categorized with ease. However, for many surgical procedures, such as bariatric surgery, hernia repair, or laparoscopic cholecystectomy, complications may be rare, and may not entirely reflect treatment effectiveness.[5,6] Alternatively, patient-reported measures of function, disability, and health status may offer a unique and more reliable assessment of provider quality and performance.

Despite a growing interest in using patient-reported outcomes (PROs) for assessing quality, important questions remain.[7,8] First, collecting patient-reported outcomes data is labor-intensive and costly, and the extent to which these measures add new information to existing performance metrics based on clinical events is not known.[9] For example, hospital-level variation in PROs may be largely explained by the occurrence of short-term clinical outcomes (such as postoperative complications), or long-term clinical effectiveness (such as disease-free survival or symptom resolution). If PROs do not provide useful information beyond existing metrics, it may not be cost-effective to direct scarce healthcare resources toward measuring PROs on a population-based level. Conversely, PROs may represent a unique domain of quality, distinct from clinical outcomes. As such, capturing PROs would be a valuable addition to quality assessment for those procedures in which clinical outcomes are less reliable. Therefore, understanding whether PROs overlap with, or diverge from, clinical outcomes, is critical to their application in quality assessment and improvement strategies.

We sought to evaluate whether PROs, specifically self-reported health-related quality of life (HRQOL), are unique domains of quality when compared with short-term and long-term clinical outcomes after bariatric surgery. We used data from the Michigan Bariatric Surgery Collaborative—a statewide consortium of hospitals performing bariatric surgery. We believe bariatric surgery is an ideal operation for this question because it is one of the most commonly performed major elective abdominal procedures in the United States with a profound effect on overall QOL, psychosocial functioning, and medical comorbidities.[10–13] The purpose of this study is to: (1) describe the relationship between short-term clinical outcomes (the occurrence of perioperative complications) and HRQOL at 1 year after bariatric surgery; and (2) describe the correlation between long-term clinical outcomes (percentage excess body weight loss and comorbidity resolution) and HRQOL at 1 year after bariatric surgery.

From the *Department of Surgery, Center for Health Outcomes and Policy, University of Michigan, Ann Arbor, MI; †Department of Surgery, Center for Health Outcomes and Policy, University of Michigan, Ann Arbor, MI; ‡Michigan Bariatric Surgery Collaborative, Center for Health Outcomes and Policy, University of Michigan, Ann Arbor, MI; §Department of Surgery, Center for Health Outcomes and Policy, University of Michigan, Ann Arbor, MI; ¶Department of Surgery, Center for Health Outcomes and Policy, University of Michigan, Ann Arbor, MI; ||Department of Surgery, University of Michigan, Ann Arbor, MI; **Center for Clinical Outcomes Development and Application, Department of Physical Medicine & Rehabilitation, University of Michigan, Ann Arbor, MI; and ††Department of Surgery, Center for Health Outcomes and Policy, University of Michigan, Ann Arbor, MI.

Reprints: Justin B. Dimick, MD, MPH, 1500 East Medical Center Drive, 2131 Taubman Center, Ann Arbor, MI 48109. E-mail: jdimick@med.umich.edu.

Funding: This research was supported by a Mentored Clinical Investigator Award to Dr. Waljee through the Agency for Healthcare Research and Quality (1K08HS023313–01).

The authors report no conflicts of interest.

ISSN: 0003-4932/14/26105-0821

DOI: 10.1097/SLA.0000000000001852

METHODS

Data Source and Study Population

We examined data collected from participants in the Michigan Bariatric Surgical Collaborative (MBSC) who underwent a primary (nonrevisional) bariatric surgical procedure from January 1, 2008, through December 31, 2012. The MBSC is a statewide consortium of over 39 hospitals and 160 surgeons (>45,000 patients), and maintains an externally audited prospective clinical registry.[14] Hospitals who perform ≥25 bariatric surgical procedures per year are eligible to voluntarily participate in MBSC.

Patient-reported Outcome Measures

Patients in the MBSC registry complete 2 measures of HRQOL preoperatively and at 1 year postoperatively. Generic HRQOL is measured using the Health and Activities Limitations Index (HALex), which captures perceived health and activity limitation into a single composite score.[15] Scores range from 0.10 (poor health) to 1.0 (superior health). Patients also complete the Bariatric Quality of Life (BQL) index, a 19-item tool that includes well-being, social functioning, physical functioning, and symptoms related to obesity.[16] The BQL contains 30 questions, with scores ranging from 14 (lowest QOL) to 78 (highest QOL).

Clinical Outcomes

As a measure of short-term clinical outcomes, we examined the occurrence of perioperative complications within 30 days of surgery through review of the medical record by trained MBSC chart abstractors.[17,18] Complications were identified by determined review of the chart, with confirmation by radiographic imaging and treatment when possible. Complications were categorized by severity in the following manner: nonlife-threatening, potentially life-threatening, or life-threatening complications associated with permanent residual disability or death[17] (Table 1). We then classified hospitals into 4 evenly distributed quantiles by the occurrence of postoperative complications. As a measure of long-term clinical outcomes, we determined the percent excess weight loss (PEWL) at 1 year, either through patient surveys or review of the medical record. As with complications, hospitals were ranked into 4 evenly distributed quartiles by the average PEWL at 1 year after surgery.

Additionally, we included the resolution of comorbid conditions at 1 year after bariatric surgery. We specifically examined the average decrease in the number of medications required for comorbid conditions related to morbid obesity, which we obtained by patient self-report.

Independent Variables

We obtained data regarding patient sociodemographic and clinical characteristics through review of the medical records. Demographic variables included patient age, sex, race, annual household income, education level, marital status, and payer type. Clinical variables included patient height and weight, which were used to calculate body mass index (BMI), a ratio of the weight in kilograms to the height in meters squared. Additionally, we determined patient smoking status and the presence of comorbid conditions by examining documentation in the medical record: pulmonary disease (asthma, obstructive/restrictive disorders, home oxygen use, Pickwickian syndrome); cardiovascular disease (coronary artery disease, dysrhythmia, peripheral vascular disease, stroke, hyperlipidemia); hypertension; sleep apnea; psychological disorders; prior venous thromboembolism (VTE); diabetes; chronic renal failure (requiring dialysis or transplant); liver disease (nonalcoholic fatty liver, cirrhosis, liver transplantation); gastroesophageal reflux disease; peptic ulcer disease; previous ventral hernia repair; and musculoskeletal disorders.

TABLE 1. Postoperative Complications After Bariatric Surgery

	N (%)
Any complication	836 (7.32%)
Nonlife-threatening	700 (6.13%)
Wound infections requiring antibiotics or exploration	229 (2.01%)
Anastomotic strictures requiring endoscopic dilation	115 (1.01%)
Bleeding requiring blood transfusion of ≤4 units, or requiring endoscopy and ≤4 units	191 (1.67%)
Pneumonia (hospital-acquired)	79 (0.69%)
Urinary tract infection (hospital-acquired)	93 (0.81%)
Clostridium difficile colitis (hospital-acquired)	24 (0.21%)
Ulcer	38 (0.33%)
Band-related problems not requiring reoperation: port site infection, band slippage, or outlet obstruction	11 (0.10%)
Potentially life-threatening	218 (1.91%)
Abdominal abscess requiring percutaneous drainage or reoperation	44 (0.39%)
Bowel obstruction requiring reoperation	42 (0.37%)
Anastomotic leak (RYGB/sleeve) or gastric perforation (band)	66 (0.58%)
Bleeding requiring blood transfusion >4 units, reoperation or splenectomy	52 (0.46%)
Wound infection or dehiscence requiring reoperation	14 (0.12%)
Respiratory failure requiring 2–7 d mechanical ventilation, no tracheostomy	9 (0.08%)
Renal failure requiring in-hospital dialysis	1 (0.01%)
Deep venous thrombosis or pulmonary embolism	43 (0.38%)
Band-related problems requiring reoperation	5 (0.04%)
Life-threatening complications associated with permanent residual disability	27 (0.24%)
Myocardial infarction or cardiac arrest	8 (0.07%)
Renal failure requiring long-term dialysis	7 (0.06%)
Respiratory failure requiring >7 d mechanical ventilation or tracheostomy	18 (0.16%)
Death	0 (0.00%)

Analysis

Our purpose was to determine the relationship between measures of HRQOL and both short and long-term clinical outcomes. To examine these relationships, we used hierarchical linear regression models to estimate the mean improvement in HRQOL (calculated as the difference between preoperative and postoperative HALex or BQL score) after bariatric surgery at the patient and hospital level. Separate models were created for HALex and BQL. We selected risk-adjustment variables for the multivariable regression models used to estimate the change in HALex or BQL scores using forward stepwise selection to include those patient sociodemographic and clinical variables that were significant at an alpha level of 0.10. We then assumed fixed coefficients for the patient-specific covariates in the hierarchical models for improvement in HALex or BQL. Models included a random hospital-specific intercept to account for the clustering of patients within hospitals. We included the occurrence of any complication and the PEWL after surgery as additional independent variables in these models. Finally, we performed reliability adjustment to account for the proportion of variation that may be due to chance alone.[19–22] To do this, we used empirical Bayes techniques to adjust the average improvement in HRQOL observed after bariatric surgery (measured by the HALex and BQL) according to the number of observations at each hospital. We calculated the correlation coefficient (r^2) directly from the hierarchical models to determine the relationship between change in HALex or BQL score and hospital ranking by complication rate or

PEWL. Statistical analyses were conducted using STATA, version 11.0 (StataCorp).

RESULTS

In this cohort of 11,420 patients who underwent bariatric surgery during the study period, the majority of patients were female (79.8%), white (84.1%), and had achieved some level of college education (45.3%). (Table 2) The majority of patients reported an annual household income of $45,000 or greater, and reported being married/partnered (69%). The average preoperative BMI was 46.6 kg/m^2, and the majority of patients underwent Roux-en-Y laparoscopic gastric bypass (56.8%).

After bariatric surgery, only 7.3% of patients suffered complications, and nearly all complications were nonlife-threatening (Table 1). Fortunately, only 1.9% of the cohort suffered complications that were potentially life-threatening, or associated with permanent disability (0.24%). Figure 1 illustrates the relationship between PROs as measured by the HALex and BQL and 30-day hospital complication rates categorized as quantiles. PROs demonstrated little correlation with complication rates when measured either by the HALex (HALex score;: $R^2 = 0.03$, $P = 0.32$) or BQL (BQL score: $R^2 = 0.003$, $P = 0.74$), after adjusting for patient factors and surgical procedure.

We then examined the association between PROs and percent excess weight loss at 1 year (Fig. 2). We observed that overall HRQOL measured by the HALex was significantly correlated with PEWL, and greater weight loss was associated with improved overall QOL at the hospital level (HALex score: $R^2 = 0.24$, $P < 0.002$). Similarly, obesity-specific HRQOL across hospitals was also correlated with PEWL after controlling for patient factors and surgical procedure, and BQL scores increased with increasing PEWL (BQL score: $R^2 = 0.44$, $P < 0.001$).

We estimated comorbidity resolution using the average number of medications consumed for condition related to morbid obesity, and examined the relationship between comorbidity resolution and PROs at the hospital level (Fig. 3). We observed that overall HRQOL was not significantly correlated with comorbidity resolution (HALex score: $R^2 = 0.05$, $P = 0.09$). In contrast, obesity-specific HRQOL was correlated with comorbidity resolution, and improved QOL scores at the hospital level were commensurate with a reduction in medications required for obesity-related conditions (BQL score: $R^2 = 0.16$, $P < 0.011$).

TABLE 2. Characteristics of Patients Undergoing Bariatric Surgery From 2008 to 2012 Enrolled in the Michigan Bariatric Surgery Collaborative (MBSC) (n = 11,420)

			N	%
Demographic factors	Age (mean ± SD)		48.1 ± 11.4	
	Sex	Males	2313	20.3
		Females	9107	79.8
	Ethnicity	White	9609	84.1
		Black	991	8.7
		Other	820	7.2
	Education	Eighth grade or less	48	0.4
		Some high school	350	3.1
		High school graduate or GED	2369	20.9
		Some college or technical school	5132	45.3
		College graduate	2159	19.0
		Graduate degree	1281	11.3
	Annual income	Less than $10,000 per year	633	5.7
		Between $10,000 and $24,999	1435	13.0
		Between $25,000 and $44,999	2733	24.8
		Between $45,000 and $75,00	3242	29.4
		More than $75,000	2995	27.1
	Private insurance		8426	73.8
	Marital status	Married or living with significant other	7833	69.0
		Widowed	319	2.8
		Divorced	1364	12.0
		Separated	177	1.6
		Single	1664	14.7
Preoperative risk factors	BMI		47.6 ± 8.3	
	Smoking history		4521	39.6
	Pulmonary disease		2912	25.5
	Cardiac disease		6629	58.1
	Sleep apnea		5173	45.3
	Hypertension		6409	56.1
	Diabetes		4036	35.4
	Liver disease		493	4.3
	Gastroesophageal reflux		5731	50.2
	Peptic ulcer disease		349	3.1
	Musculoskeletal disorder		8853	77.5
	Prior hernia repair		348	3.1
Surgical procedure	Roux-en-Y laparoscopic bypass		6488	56.8
	Laparoscopic gastric banding		2359	20.7
	Sleeve gastrectomy		2573	22.5

GED indicates general equivalency diploma.

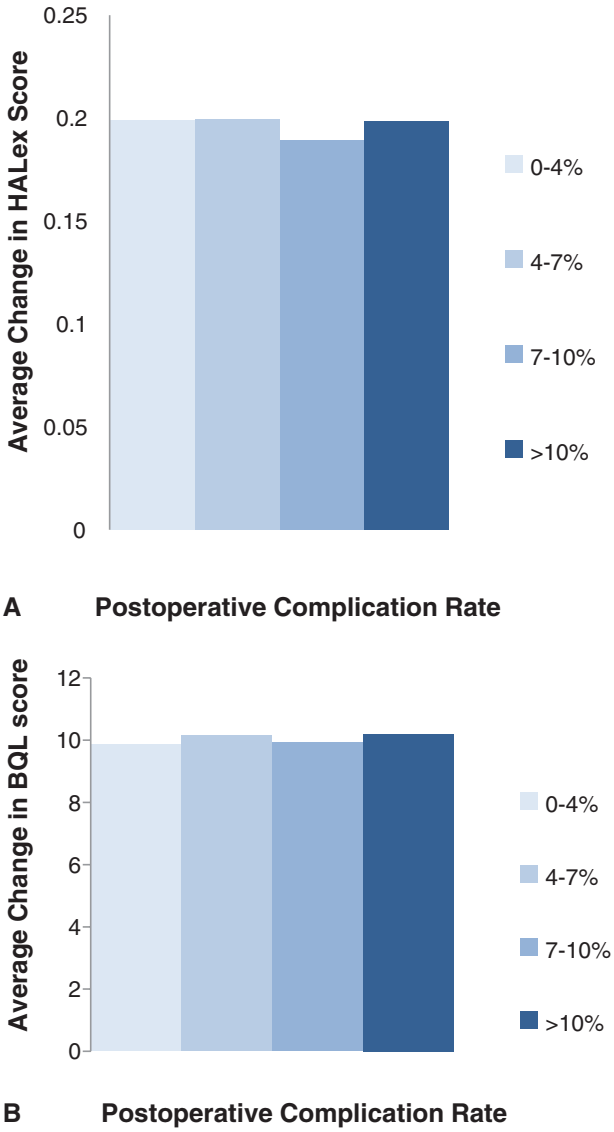

FIGURE 1. Patient-reported outcomes across hospitals ranked by the occurrence of 30-day postoperative complications (adjusted for patient clinical and sociodemographic factors, and reliability). A, HALex score: $R^2 = 0.03$, $P = 0.32$. B, BQL score: $R^2 = 0.003$, $P = 0.74$.

Finally, we used hierarchical modeling to examine the proportion of variation in PROs that is influenced by patient characteristics, surgical procedure, and clinical outcomes (Table 3). The majority of variation in HALex and BQL scores was explained by surgical procedure (HALex: 50%; BQL: 67%). Additionally, PEWL explained 37% of the variation in HALex score and 80% of the variation in BQL score, and comorbidity resolution accounted for 23% and 12%, respectively. We observed that only 4% of the variation in HALex score and 5% of the variation in BQL score was explained by patient demographic characteristics. When all of these factors were combined together in a single model, 44% of the variation in overall HRQOL and 15% of the variation in obesity-specific HRQOL was left unexplained.

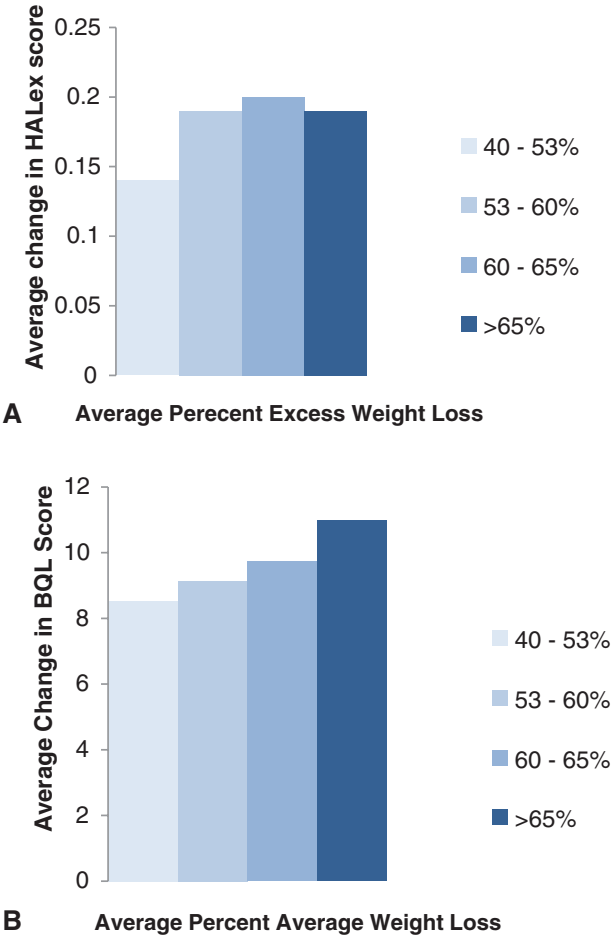

FIGURE 2. Patient-reported outcomes across hospitals ranked by percentage of excess weight lost at 1 year after bariatric surgery (adjusted for patient clinical and sociodemographic factors, and reliability). A, HALex score: $R^2 = 0.24$, $P = 0.002$. B, BQL score: $R^2 = 0.44$, $P < 0.001$.

DISCUSSION

The majority of existing surgical quality assessment and improvement programs focus largely on clinical events, such as measures of perioperative safety and clinical effectiveness. These indicators are commonly captured in clinical registries, but may not fully capture the potential benefits of surgical procedures directed toward symptom resolution and improving HRQOL. In this study of patients undergoing bariatric surgery in Michigan, we found that PROs vary across hospitals, and are, at least to some degree, distinct from both short and long-term clinical outcomes. Most patients experience improved QOL after surgery, and measures of clinical effectiveness, such as weight loss and comorbidity resolution, are more predictive of HRQOL than perioperative complications. Nonetheless, some variation in PROs remains unexplained after accounting for clinical factors, suggesting that these measures may represent a unique aspect of surgical quality.

Several studies have examined the potential for PROs to serve as indicators of provider performance. For example, in the United Kingdom, PROs are routinely collected from patients since 2009 through the National PROM Programme as a quality assessment and

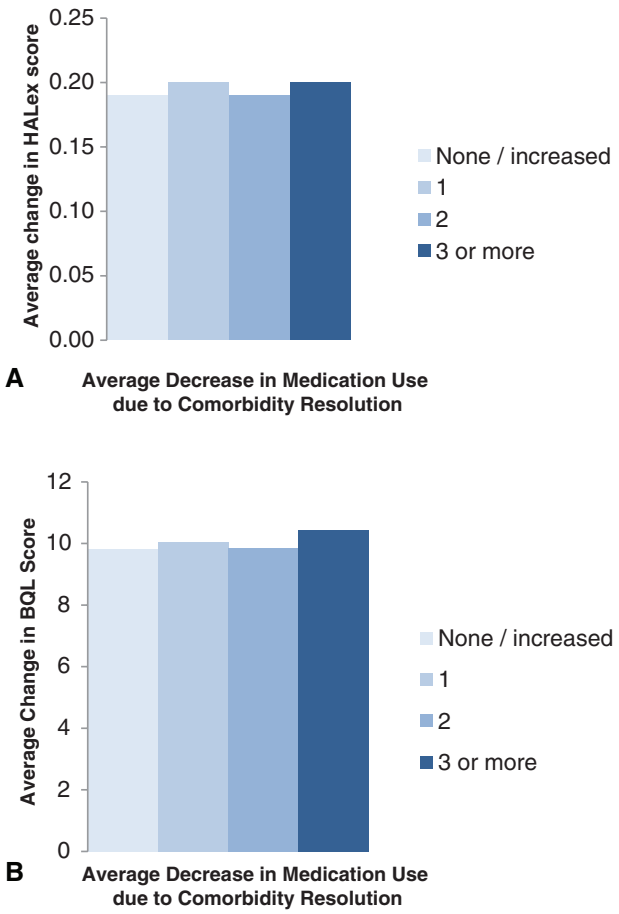

FIGURE 3. Patient-reported outcomes across hospitals ranked by average decrease in medications required for comorbid conditions (adjusted for patient clinical and sociodemographic factors, and reliability). A, HALex score: $R^2 = 0.05$, $P = 0.09$. B, BQL score: $R^2 = 0.16$, $P = 0.011$.

improvement strategy directed toward both hospitals and clinicians.[23–25] Evidence from this effort demonstrates that PROs vary across providers, and are dependent on the measure and domain used to profile physicians. Moreover, PROs provide a sensitive assessment of provider performance, although nonresponders differ from responders in ways that may influence provider ranking by these metrics.[8,26,27] Finally, PROs provide an opportunity for improved population-based cost-effectiveness analyses using outcomes germane to procedures performed for symptomatology and improving QOL.[25,28] As we observed in our cohort, measurable variation exists across providers, which may provide an opportunity to distinguish clinical performance. Furthermore, in our study, we observed that commonly used metrics such as perioperative complications explain little of the variation in long-term PROs. It is likely that the sequelae of clinical complications, such as a surgical site infection or anastomotic leak, have resolved by 1 year, and PROs obtained at this time do not impact the self-reported health status. It is also possible that patient perception of health status is influenced by other factors beyond short-term complications of care. For example, patients are often satisfied despite the occurrence of complications, and patient satisfaction is more heavily influenced by the patient–provider relationship than the occurrence of adverse or unexpected events.[29–31]

We also observed that PROs are correlated with long-term outcomes, and self-reported health status may be a better measure of treatment effectiveness than other available metrics. For example, after joint replacement procedures, implant failure rates and functional assessment parameters are commonly used to describe successful or failed outcomes. However, previous studies have demonstrated that patient satisfaction does not correlate with these parameters, and pain and disability may persist even without implant failure.[32–34] In this context, the occurrence of postoperative complications may be more indicative of perioperative safety measures and patient selection, whereas PROs may better reflect the durability of clinical outcomes. For example, chronic groin pain can occur in up to 30% of patients undergoing otherwise successful inguinal hernia repair.[35] As such, PROs can provide a markedly different assessment of treatment success compared with clinical outcomes.

Our study has several notable limitations. First, although the BQL and HALex have been used to describe HRQOL among patients

TABLE 3. Proportion of Variation in PROs at the Hospital Level Explained by Clinical and Patient-level Factors*

		Proportion of Hospital-level Variation of Difference in HRQOL Measured by the HALex Score (%)	Proportion of Hospital-level Variation of Difference in HRQOL Measured by the BQL Score (%)
Sociodemographic factors	Age	0.03	0.11
	Sex	0.32	0.02
	Race	0.49	2.09
	Education level	1.73	0.59
	Income	0.03	0.40
	Insurance status	2.66	0.99
	All	4.04	5.09
Surgical procedure		49.87	66.48
Postoperative complications		0.14	1.24
Percent excess weight loss after 1 year		37.19	79.54
Decrease in the average number of medications required due to comorbidity resolution		22.92	12.23
All factors		56.17	85.01
Unexplained		43.83	14.99

*Sociodemographic factors were added both individually and as a group (“All”) to the model.

with morbid obesity who have undergone bariatric surgery, their ability to discriminate clinically meaningful differences over time is uncertain. As research on PROs moves forward, evaluating clinically meaningful differences will be essential. Second, this cohort only includes patients who completed health assessments at 1 year after surgery, which comprised 32% of all patients in the MBSC registry, leaving the potential for a response bias. Patients who were lost to follow-up may differ in important ways with respect to HRQOL that we were unable to measure. However, whereas this may impact the baseline PRO measures, it is unlikely to impact the correlation between clinical events and PROs. Third, the MBSC registry captures only short-term clinical complications, rather than longer-term complications (eg ulceration, bowel obstruction), which may have the potential to influence PROs at 1 year. Finally, this analysis is drawn from patients undergoing procedures at hospital participating in the MBSC, and there may be aspects of hospitals that engage in quality collaborative improvement that are inherently different compared with hospitals that are not involved in these endeavors.

Nonetheless, our findings have important implications for efforts to integrate PROs into surgical quality assessment and improvement strategies. PROs are aspects of health status that cannot be measured without self-report, and, unlike aspects of patient experience, are analogous to physiologic or survival outcomes with respect to treatment effectiveness. Variation in PROs across hospitals remains after accounting for clinical events, suggesting the PROs should be considered for inclusion into quality assessment strategies. However, several unanswered questions remain. Measuring surgical quality may be a dynamic process, and the optimal time to capture specific metrics of performance is not known. For example, early events, such as clinical complications, may be ideal to identify provider safety and the ability to rescue patients from adverse events.[36] In contrast, long-term outcomes may be more indicative of treatment effectiveness. Future work that examines the trajectory of patients based on preoperative self-reported health status may provide important information on prognosis after surgery and opportunities to optimize fragile or vulnerable patients preoperatively. Furthermore, we observed that clinical outcomes correlated differently with generic and condition-specific measures of HRQOL. This highlights the importance of maintaining a comprehensive approach to capturing PROs in the context of hospital performance, as overall HRQOL may reflect different aspects of quality compared with HRQOL specific to a condition or procedure of interest.[23] Finally, achieving adequate survey response rates is challenging, even in clinical research trials that include generous incentives and intensive recruitment and retention strategies. Nonresponder bias is not random, and relates to systematic differences in patient risk, healthcare engagement, and experience. Therefore, any hospital-quality initiative centered on PROs must account for this bias, and its influence on performance assessment. As such, using PROs as quality metrics may not be feasible in all practice settings due to financial and logistical barriers. Therefore, identifying efficient and practical strategies to implement PROs in environments that will derive the greatest benefit is critical to leverage scarce resources effectively.

In conclusion, PROs are distinct from clinical outcomes, and represent a potential indicator of performance that can be targeted to improve quality of care. Future studies that examine the influence of measurement techniques, case mix, and disease characteristics on PROs will inform efforts to integrate these critical outcomes into quality assessment strategies.

REFERENCES

1. Cevasco M, Ashley SW. Quality measurement and improvement in general surgery. *Permanente J*. 2011;15:48–53.
2. Hawn MT, Vick CC, Richman J, et al. Surgical site infection prevention: time to move beyond the surgical care improvement program. *AnnSurg*. 2011;254:494–499 [discussion 499–501].
3. Hofer TP, Hayward RA. Can early re-admission rates accurately detect poor-quality hospitals? *Med Care*. 1995;33:234–245.
4. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med*. 1997;127(8 Pt 2):666–674.
5. Shih T, Dimick JB. Reliability of readmission rates as a hospital quality measure in cardiac surgery. *Ann Thorac Surg*. 2014;97:1214–1218.
6. Birkmeyer JD, Dimick JB, Birkmeyer NJ. Measuring the quality of surgical care: structure, process, or outcomes? *J Am Coll Surg*. 2004;198: 626–632.
7. Black N. Patient reported outcome measures could help transform healthcare. *BMJ*. 2013;346:f167.
8. Varagunam M, Hutchings A, Black N. Do patient-reported outcomes offer a more sensitive method for comparing the outcomes of consultants than mortality? A multilevel analysis of routine data. *BMJ Qual Saf*. 2014.
9. Franklin PD, Harrold L, Ayers DC. Incorporating patient-reported outcomes in total joint arthroplasty registries: challenges and opportunities. *Clin Orthopaed Related Res*. 2013;471:3482–3488.
10. Nguyen NT, Masoomi H, Magno CP, et al. Trends in use of bariatric surgery, 2003–2008. *J Am Coll Surg*. 2011;213:261–266.
11. Carlin AM, Zeni TM, English WJ, et al. The comparative effectiveness of sleeve gastrectomy, gastric bypass, and adjustable gastric banding procedures for the treatment of morbid obesity. *Ann Surg*. 2013;257:791–797.
12. Nguyen NT, Nguyen B, Gebhart A, et al. Changes in the makeup of bariatric surgery: a national increase in use of laparoscopic sleeve gastrectomy. *J Am Coll Surg*. 2013;216:252–257.
13. Aftab H, Risstad H, Sovik TT, et al. Five-year outcome after gastric bypass for morbid obesity in a Norwegian cohort. *Surg Obesity Related Dis*. 2014;10: 71–78.
14. Birkmeyer NJ, Dimick JB, Share D, et al. Hospital complication rates with bariatric surgery in Michigan. *JAMA*. 2010;304:435–442.
15. Erickson P. Evaluation of a population-based measure of quality of life: the Health and Activity Limitation Index (HALex). *Qual Life Res*. 1998;7: 101–114.
16. Weiner S, Sauerland S, Fein M, et al. The Bariatric Quality of Life index: a measure of well-being in obesity surgery patients. *Obes Surg*. 2005;15: 538–545.
17. 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;254:633–640.
18. Finks JF, English WJ, Carlin AM, et al. Predicting risk for venous thromboembolism with bariatric surgery: results from the Michigan Bariatric Surgery Collaborative. *Ann Surg*. 2012;255:1100–1104.
19. Dimick JB, Ghaferi AA, Osborne NH, et al. Reliability adjustment for reporting hospital outcomes with surgery. *Ann Surg*. 2012;255:703–707.
20. Smith KA, Sussman JB, Bernstein SJ, et al. Improving the reliability of physician ''report cards''. *Med Care*. 2013;51:266–274.
21. Hayward RA, Heisler M, Adams J, et al. Overestimating outcome rates: statistical estimation when reliability is suboptimal. *Health Serv Res*. 2007;42:1718–1738.
22. Hashmi ZG, Dimick JB, Efron DT, et al. Reliability adjustment: a necessity for trauma center ranking and benchmarking. *J Trauma Acute Care Surg*. 2013;75:166–172.
23. Gutacker N, Bojke C, Daidone S, et al. Hospital variation in patient-reported outcomes at the level of EQ-5D dimensions: evidence from England. *Med Decis Making*. 2013;33:804–818.
24. Gutacker N, Bojke C, Daidone S, et al. Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. *Health Econ*. 2013;22: 931–947.
25. Vogl M, Leidl R, Plotz W, et al. Comparison of pre- and post-operative health-related quality of life and length of stay after primary total hip replacement in matched English and German patient cohorts. *Qual Life Res*. 2014;24: 513–520.
26. Hutchings A, Neuburger J, Grosse Frie K, et al. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. *Health Qual Life Outcomes*. 2012;10:1–8.
27. Neuburger J, Hutchings A, van der Meulen J, et al. Using patient-reported outcomes (PROs) to compare the providers of surgery: does the choice of measure matter? *Med Care*. 2013;51:517–523.

28. Appleby J, Poteliakhoff E, Shah K, et al. Using patient-reported outcome measures to estimate cost-effectiveness of hip replacements in English hospitals. *J R Soc Med*. 2013;106:323–331.
29. Wang R, Wood DP Jr, Hollenbeck BK, et al. Risk factors and quality of life for post-prostatectomy vesicourethral anastomotic stenoses. *Urology*. 2012;79:449–457.
30. Yao SL, Lu-Yao G. Population-based study of relationships between hospital volume of prostatectomies, patient outcomes, and length of hospital stay. *J Natl Cancer Inst*. 1999;91:1950–1956.
31. Hu JC, Elkin EP, Pasta DJ, et al. Predicting quality of life after radical prostatectomy: results from CaPSURE. *J Urol*. 2004;171(2 Pt 1):703–707 [discussion 707–708].
32. Waljee JF, Chung KC. Objective functional outcomes and patient satisfaction after silicone metacarpophalangeal arthroplasty for rheumatoid arthritis. *J Hand Surg Am*. 2012;37:47–54.
33. Price AJ, Longino D, Rees J, et al. Are pain and function better measures of outcome than revision rates after TKR in the younger patient? *Knee*. 2010;17:196–199.
34. Noble PC, Conditt MA, Cook KF, et al. The John Insall Award: patient expectations affect satisfaction with total knee arthroplasty. *Clin Orthopaed Related Res*. 2006;452:35–43.
35. Sun P, Pandian TK, Abdelsattar JM, et al. Reoperation for groin pain after inguinal herniorrhaphy: does it really work? *Am J Surg*. 2016;211:637–643.
36. Ghaferi AA, Birkmeyer JD, Dimick JB. Complications, failure to rescue, and mortality with major inpatient surgery in medicare patients. *Ann Surg*. 2009;250:1029–1034.

DISCUSSANTS

P. Kuo (Maywood, IL):

I thank the authors for sending me their article so early and congratulate them on a nice study.

In this article, the authors examined patient-reported outcomes using a cohort of patients who underwent bariatric surgery during the period of 2009 to 2013 from the Michigan Bariatrics Surgery collaborative of 39 hospitals. They find that hospital rankings based on patient-reported outcomes do not correlate with rankings based on complications. In addition, the patient-reported outcome measures correlate with weight loss. They conclude that patient-reported outcomes correlate with measures of clinical effectiveness after bariatric surgery. They propose that a comprehensive approach to surgical quality should incorporate both the clinical events and self-reported measures.

I have few questions for the authors. Some of them are methodologic.

1. You use 2 patient-reported outcomes measures. Do you have any thoughts about the collinearity of the 2 measures; why did you divide the hospitals into quantiles as opposed to considering the hospitals as a deciles or other method of binning?

2. Did you include zip code among the patient sociodemographic and clinical characteristics? The cohort only includes 32% of all that patients in the registry. Was there any effort to contact patients in the nonreporting group?

I will skip over my third question and go to the fourth.

4. What do you think the role of preoperative patient education and management of expectation plays in this finding? Given that bariatric surgery typically requires a 6-month waiting period and extensive patient education, does this influence patient-reported outcomes? Is there any possibility of reporting the comparison with another group that does not have the benefit of the extensive preoperative period? Perhaps trauma patients?

Lastly, and probably this is my most important question, at least for me.

5. Philosophically speaking, I personally worry about the popular societal trend in which opinion becomes fact. For example, if most of the world suddenly believes that the sun revolves around the earth, what happens then?

In closing, I thank the authors and society for the privilege of commenting on this article.

Response From J.F. Waljee:

Thank you so much for your comments regarding our work. In this study, we used 2 discrete measures to assess overall quality of life and bariatric-specific quality of life. Each was included in separate analytic models to avoid overlap and collinearity. We used a 2-level regression model to control for patient-related risk factors, and examine the correlation between PROs and clinical outcomes entirely at the hospital level. As such, we have reported risk-adjusted correlation coefficients that are bi-directional, and we would not expect a difference in their magnitude if clinical events were changed to the outcome and PROs were included as the predictor variable. We agree that understanding the relationship between PROs and healthcare costs has important implications when considering prioritizing and improving care based on value. In future work, linking claims data with clinical registry outcomes may provide greater clarity on this relationship. With respect to categorization of the clinical outcomes, we used quartiles to achieve evenly spaced groups across the 39 included hospitals. However, our findings remained similar when these were examined by quintiles and tertiles, and have opted to present these as quartiles for ease of presentation. Within MBSC, sociodemographic characteristics are captured using age, sex, ethnicity, education level, annual income, insurance type, and marital status. Given that these are patient-level covariates, we felt that these factors provide a more sensitive assessment of socioeconomic status compared with zip-code, for which there may be wide variation in these variables within a single geographic region.

The point raised regarding responder bias and rates is critical. Our data represent a real-world assessment of the challenges of collecting PROs in clinical practice. Standard protocols include regular phone calls and follow-up reminders to contact patients. However, only a proportion of patients who undergo surgery follow-up at 1 year, and not all hospital sites routinely schedule 1-year follow-up with patients. Therefore, taken together, these rates represent 32% of all patients who initially undergo surgery.

Although nonresponse is a source of bias, it also represents an important implication if PROs are to be used for the purposes of quality assessment. Achieving adequate survey response rates is challenging, even in clinical research trials that include generous incentives and intensive recruitment and retention strategies. Nonresponder bias is not random, and relates to systematic differences in patient risk, healthcare engagement, and experience. Therefore, any hospital-quality initiative centered around PROs must account for this bias, and its influence on performance assessment. As such, using PROs as quality metrics may not be feasible in all practice settings due to financial and logistical barriers. Therefore, identifying efficient and practical strategies to implement PROMs in environments that will derive the greatest benefit is critical to leverage scarce resources effectively.

We also agree that the baseline level of patient-reported function may provide important information regarding the potential to derive benefit from surgeries performed for symptomatology and quality of life. In this analysis, we accounted for baseline differences in health-related quality of life at the hospital level by using a standardized measure (the difference in preoperative and postoperative score divided by the preoperative score) across hospitals. Using this approach, we were able to more closely examine the gain/improvement in quality of life after bariatric surgery. Although beyond the scope of this study, future work that examines the trajectory of patients based on preoperative self-reported health status may provide important information on prognosis after surgery and opportunities to optimize fragile or vulnerable patients preoperatively.

Finally, in recent years, aspects of patient experience during healthcare episodes have been increasingly used as a measure of hospital performance. For example, the Consumer Assessment of Healthcare Providers and Systems program collects self-reported data regarding patient satisfaction with inpatient hospital care, such as provider communication, environmental conditions, and quietness. Although self-reported, these measures reflect experiences rather than discrete measures of health status, such as functional status, mobility, and pain. In this context, PROs are quite distinct from aspects of satisfaction with experience, which may reflect an amalgam of patient and provider interactions and customer service-related phenomena, such as convenience in parking or expenses incurred. By definition, PROs are aspects of health status that cannot be measured without self-report, and are more similar to physiologic or survival outcomes with respect to treatment effectiveness compared with measure of patient experience.

R. Mullins (Portland, OR):

I have nothing to disclose. Congratulations on a very nice outcome study. And like many outcome studies, it looks at patients who had an operation. This operation was elective.

Should not you really start your outcome assessment in the clinic where the doctor says, ''I will operate an you," and says to another patient, ''I just don't think you are going to benefit."

Is not that an important indicator of quality that they can find the patient who is going to have a dissatisfied outcome and say, ''Well, we shouldn't operate on you"?

Response From J.F. Waljee:

Thank you for your question, which raises an important point regarding our findings. Just as many clinical aspects of postoperative recovery are dynamic, such as the resolution of pneumonia or surgical site infections, patient-reported outcomes are expected to change over time as well. Moreover, patient-reported outcomes could play a critical role in improving our ability to predict patients who will derive the greatest benefit after surgery, and who may not. Currently, the logistical barriers and expense of collecting PROs from all patients at all time points preclude their use for this purpose, but may be enhanced going forward with dispersion of electronic medical records and tablet-based methods leveraging modern psychometric theory.

P. Schauer (Cleveland, OH):

I have no disclosures. Congratulations on an excellent study, and thank you for attempting to bring some clarity on this disparity between what patients think is a good outcome and what surgeons think are good outcomes. I think this type of analysis is extremely important.

If I interpret your data correctly, there seems to be fairly good correlation with patient-reported outcome and the longer-term clinical benefits of surgery, but not so much good correlation with these early perioperative events or complications.

I wonder if this is a fact that patients, perhaps a year out, are forgetting about some of these early complications, these strictures and DVTs that are probably resolved, and do you think that the timing of this survey, a year out vs perhaps in the early perioperative period, influences the patient's decision when they balance the risks and the benefits of a procedure they received?

Response From J.F. Waljee:

Thank you for your question, which raises an important point regarding our findings. Just as many clinical aspects of postoperative recovery are dynamic, such as the resolution of pneumonia or surgical site infections, patient-reported outcomes are expected to change over time as well. Moreover, patient-reported outcomes could play a critical role in improving our ability to predict patients who will derive the greatest benefit after surgery, and who may not. Currently, the logistical barriers and expense of collecting PROs from all patients at all time points preclude their use for this purpose, but may be enhanced going forward with dispersion of electronic medical records and tablet based methods leveraging modern psychometric theory.

J. Vetto (Portland, OR):

Nothing to disclose. I would like to get back to one of your limitation slides. You mentioned at the very end of your limitations that perhaps the disease had some kind of impact. In fact, you study a disease, morbid obesity, where people generally get better after operation. That is, their reported improvements, and the improvements we measure, align. But what about diseases where people, unfortunately, get worse, at least initially, after operation? I refer specifically, for example, to a cancer patient. Someone with breast cancer or melanoma, or sarcoma comes in relatively asymptomatic, and we induce arm swelling, leg swelling, chest wall pain, and other disabilities. In those cases, do not you think that physician and patient reported outcomes may not align? So maybe, the disease that you study answers the question posed in the title of your article.

Response From J.F. Waljee:

Thank you for this question, and I agree that it is challenging, and likely impossible, to identify a single, universal instrument that captures all aspects of recovery germane to both patients and surgeons for all conditions. Furthermore, recent studies from the United Kingdom suggest that provider ratings vary by instrument selection, which has important implications for using PROs as quality metrics. More generic measures of health status may be less responsive to clinical change over time after a surgical procedure. However, they may also highlight elements of care that are different and perhaps better managed by some providers rather than others, such as pain and mobility. In this context, selecting instruments that elicit those elements of recovery that provide the most meaningful information regarding recovery will be essential when integrating PROs for this purpose.