Journal of Surgical Research **143,** 158–163 (2007)
doi:10.1016/j.jss.2007.03.040

How Much Do Standardized Forms Improve the Documentation of Quality of Care?

Janak A. Parikh, M.D.,*,†,[1] Irina Yermilov, M.D.,* Sushma Jain, M.P.H.,† Marcia L. McGory, M.D.,* Clifford Y. Ko, M.D., M.S.H.S.,*,† and Melinda A. Maggard, M.D., M.S.H.S.*

**Department of Surgery, Geffen School of Medicine at UCLA, Los Angeles, California, †Department of Surgery, Greater West Los Angeles VA Healthcare System, Los Angeles, California*

Submitted for publication January 9, 2007

Background. **Chart abstraction is a common method for measuring the quality of surgical care. In this study we examine how the use of standardized operative dictation and history forms improves documentation rates of bariatric quality measures.**

Materials and methods. **Two independent reviewers evaluated 201 patient charts from two multi-surgeon bariatric surgery practices for documentation of five intraoperative and seven preoperative bariatric quality measures. Group 1 used fully standardized templates to dictate or collect both, while Group 2 did not. Documentation rates were compared between the groups.**

Results. **Operative reports more consistently documented quality assessment information for cases where a dictation template was used *versus* where it was not (89% *versus* 58%, respectively, $P < 0.001$). The greatest discrepancies between the two groups were found in "exploration of the abdomen" (95% in Group 1 *versus* 43% in Group 2, $P < 0.001$) and in "evaluation of the gallbladder" (76% *versus* 28%, $P < 0.001$). In comparison, overall documentation rates for preoperative comorbidities were greater in both groups but remained higher for Group 1, who used fully standardized forms (98% *versus* 74%, $P < 0.001$). Group 1 had statistically significant higher rates of documentation for all seven comorbidities.**

Conclusions. **The use of standardized dictation templates and history forms is associated with significantly higher documentation rates of quality measures in bariatric surgery. The adoption of these methods into routine use will be needed to allow for wide scale quality assessment and improvement for surgical practices.**

Key Words: **bariatric; laparoscopic; Roux-en-Y; gastric bypass; quality; documentation; assessment; standardized.**

INTRODUCTION

Recently, an increased amount of attention is being paid to quality of care in surgery. The Center for Medicare and Medicaid Services is considering pay-for-performance (P4P) measures for surgical diseases. Currently, the Center for Medicare and Medicaid Services uses hospital level P4P measures for acute myocardial infarction, heart failure, and pneumonia. Examples of P4P measures for acute myocardial infarction include receipt of a beta-blocker, an aspirin, and an ace-inhibitor [1, 2]. The Hospital Quality Initiative requires all eligible hospitals to submit data on adherence to the quality indicators for these conditions. Hospitals that do not submit performance data will receive a 0.4 percentage point reduction in the annual payment updates [1–5]. Presently, physician-level P4P measures are being considered. Given the current push for these programs, it is important to accurately document adherence to quality indicators.

The reporting of adherence to these quality measures is typically performed by the individual hospital by abstracting the data from medical records. Chart abstraction is one of the most common methods for measuring processes of care [6–8]. The value of chart abstraction as a tool to measure quality in surgery is best illustrated by the Department of Veterans Affairs' National Surgical Quality Improvement Program (VA NSQIP). This program employs 88 full-time, trained surgical clinical nurse reviewers to ensure accurate

[1] To whom correspondence and reprint requests should be addressed at Department of Surgery, Geffen School of Medicine at UCLA, CHS, Rm 72-215, 10833 Le Conte Ave, Los Angeles, CA 90095. E-mail: jparikh@mednet.ucla.edu.

0022-4804/07 $32.00

collection of data, which are then transmitted to a national database [9, 10]. Several studies have validated the accuracy and utility of the VA NSQIP collected data [9–14]. The success of this VA program is likely due to (1) a skilled, trained full-time chart abstractor, (2) a computerized medical record with templates for operative notes, progress notes, orders, etc., and (3) on-site auditing of quality of reporting. As evidenced by the success of NSQIP, the use of chart abstraction provides accurate documentation, particularly in a controlled setting such as the VA where trained abstractors are used.

Despite its utility, however, chart abstraction has some drawbacks. First, it is time intensive and costly because it typically requires employment of an on-site nurse dedicated to collecting the data [15], as with VA NSQIP. Second, chart abstraction has primarily been validated only in the inpatient setting [16, 17], and its application in the outpatient setting is not as well-defined [17–19]. Lastly, the usefulness of chart abstraction depends on how clearly and thoroughly the data are documented in the patient's chart [20].

In an effort to improve documentation, there has recently been an increase in the use of standardized dictation templates, history forms, and clinical pathways in many surgical fields [21]. Standardized templates often consist of preprinted forms designed for a specific purpose (e.g., completing a history and physical) so that a provider may save time by checking a box, for instance, rather than writing out sentences. In a recent study by Laflamme *et al.*, operative notes created using an electronic template had faster turnaround times, increased compliance with national standards for operative note documentation, and lower costs [22, 23]. These tools attempt to facilitate correct documentation in a time efficient manner. How the use of standardized templates affects the documentation of quality of care, however, remains unknown.

In this study, we examine how the use of standardized templates affects the documentation of intraoperative quality measures and preoperative comorbidities in bariatric surgery by comparing two different styles of surgical practice. We looked at one group of surgeons employing fully standardized templates for both operative report dictation and patient history (which includes assessment of preoperative comorbidities), and another set of surgeons who do not use a standardized template for operative reports but some partial templates for patient history.

METHODS

A retrospective review of 2 multi-surgeon bariatric surgery practices for documentation of 5 intraoperative quality indicators and 7 preoperative comorbidities. Two independent researchers abstracted information on 201 patients, approximately half from a 4-surgeon private practice bariatric surgery group that utilized standard template operative reports and preoperative patient history forms for collecting comorbidity data (Group 1, $n = 108$). The remaining half of the charts were from an eight-surgeon academic group (Group 2, $n = 93$) who did not use a standard template operative report or a template for preoperative history to document comorbidities. However, four of the eight surgeons in Group 2 did use a template to obtain a preoperative history. Electronic and paper charts for all patients were reviewed. Cases included both open and laparoscopic Roux-en-Y gastric bypass (RYGB), laparoscopic adjustable bands, and revisions. Patients were selected randomly from each surgeon's practice. In other words, we made an effort to obtain an equal representation of cases for each surgeon.

The operative reports were screened for documentation of five intraoperative quality measures: (1) exploration of the abdomen, (2) examination of the gallbladder, (3) intra- or postoperative evaluation of the anastomosis for leak, (4) closure of the large bowel mesenteric defect or antecolic placement of Roux limb, and (5) closure of the small bowel mesenteric defect.

Group 1 used dictation templates to dictate the portions of the operative report that were similar from patient to patient. Many dictations were performed by the physician's assistant (PA), then reviewed and signed by the surgeon. The PA (four in total) who scrubbed on the case performed the dictation. For cases where the operation deviated from the standard, the dictation was performed by the surgeon, but this occurred rarely. Group 2 did not consistently use a dictation template or a PA to perform the dictations.

The preoperative patient history was examined for documentation of seven comorbidities: (1) diabetes (DM), (2) hypertension (HTN), (3) hyperlipidemia (LIPID), (4) sleep apnea (SA), (5) venous stasis disease (VSD), (6) degenerative joint disease (DJD), and (7) reflux (REFLUX). Preoperative patient histories were taken by several provider types in each practice.

The quality measures selected were previously developed by Maggard *et al.* using evidence in the literature and expert consensus [24]. A summary of the comparison between the groups with regards to the types of data collection forms and degree of standardization is provided in Table 1.

Operative Reports

A standardized abstraction tool that included possible responses of "Yes", "No", "Not reported", "Not applicable" or "Missing" was used. For operative report quality measures, "Yes" was recorded if the operative report confirmed the indicator was met. "No" was assigned if it was clear that the criteria were not met for the particular indicator. "Not reported" was recorded if we were unable to determine if the indicator was met or not. For example, if the surgeon did not comment on whether an anastomotic leak test was done, then "Not reported" was assigned. For the analysis, "Yes" and "No" were classified together, because the record contained data to determine whether the indicator was met or not, thus satisfying that the data were documented. In contrast, "Not reported" meant that the data were not sufficiently documented in the records. "Not applicable" was assigned if the patient was not eligible for the particular indicator; for example for a patient who had previously undergone cholecystectomy, the indicator specific to the gallbladder would not be eligible. For the purpose of our analysis, an assignment of "Yes" was required to satisfy the criterion that the operative indicator was documented. "Missing" was assigned when the operative report was not found in the medical record, either the paper chart or electronic records.

Preoperative Comorbidities

For assessing the preoperative comorbidities, we reviewed all history forms and patient intake forms. "Yes" was recorded if there was documentation that the patient suffered from the comorbidity, and "No" was recorded if there was documentation that the patient did not have the comorbidity. For example, the history or patient intake form must document "no diabetes" to get credit for assign-

TABLE 1

Data Form Documentation and Degree of Standardized Templates Used by Practice Group Types

Type of Data	Type of form where data was collected	Group 1	Group 2
Operative indicators	Operative report	Fully standardized template	No standardized template
Preoperative comorbidities	Patient intake	Fully standardized template	Variable standardization template
	Physician assistant history	Fully standardized template	Not applicable
	Surgeon history	Partially standardized template	No standardized template
	Internist history	No standardized template	No standardized template
	Anesthesia history	Partially standardized template	Partially standardized template

Practice Group 1: Uses primarily fully standardized operative and history templates.
Practice Group 2: Partially uses standardized operative and history templates.

ment of "No" during our abstraction process. "Not reported" was marked if there was no mention of the comorbidity. For the purpose of our analysis, an assignment of "Yes" or "No" satisfied the criterion that the comorbidity was documented. Patient charts that did not contain history and physical or patient intake form were categorized as "Missing".

During the abstraction process, we did not look for the use of specific key words in documentation of comorbidities, but rather gave credit for commenting on likely similar terms. For instance, if a provider notes that the patient has chronic joint pain, we gave credit for documenting DJD. However, we did not verify the presence or absence of a comorbidity based on laboratory results or imaging studies. All data are based on what was documented in a patient's chart.

Data Analysis

The overall documentation rate for intraoperative quality measures and preoperative comorbidities was calculated by obtaining a weighted average across all measures for each category. A pooled standard deviation was also obtained, as was an absolute difference between the two groups for each category. A two-sample test of proportions was then used to determine if the difference between the groups was statistically significant at an alpha level of 0.05.

To examine the difference in documentation of specific intraoperative quality measures between the two groups, we tabulated the number of "Yes" and "No" for each of the five individual measures. The rates were calculated based on the number of completed items for each measure. Therefore, the denominator varied from measure to measure and between groups. An absolute difference was calculated between the two groups and a *P*-value was obtained using a two sample test of proportions to determine if the absolute difference was statistically significant. A subgroup analysis was performed for documentation of comorbidities in Group 2 since four of the eight surgeons used a template to obtain patient histories. Missing items were excluded from the analysis.

Similar to the analysis of five individual intraoperative quality measures, we examined documentation rates of the seven preoperative comorbidities. Assignment of a "Yes" or "No" was evidence that the comorbidity status was documented. Missing items were excluded from the analysis and were only encountered in Group 2's documentation of REFLUX.

Inter-rater agreement was evaluated by cross validating approximately 10% of the charts, although a kappa statistic was not calculated. The details and scope of the project were reviewed and approved by the UCLA Institutional Review Board and the participating outside hospitals.

RESULTS

Patient Demographics

Patient characteristics were similar in both Group 1 and Group 2 (Table 2). Over 80% of patients in each practice were female with a median age of 45 y in Group 1 and 41 y in Group 2. Patients in Group 1 had a median BMI of 46 kg/m^2 compared with 51 kg/m^2 for Group 2. More patients in Group 1 underwent a laparoscopic RYGB (83.8%) compared with Group 2 (57.1%). Conversely, fewer patients in Group 1 underwent an open RYGB compared with patients in Group 2 (9.1% *versus* 39.6%, respectively).

Overall Documentation Rates

Group 1, which used primarily standardized dictation templates, demonstrated an 88% overall documentation rate for the collective group of intraoperative quality measures compared with 59% for the Group 2,

TABLE 2

Patient Demographics by Practice Group Type

Variables	Group 1	Group 2	*P* value
Age (median, std)	45 +/− 10.8 y	41 +/− 9.1 y	0.005
Female	83.3%	82.6%	0.895
BMI (median, std)	46 +/− 7.7 kg/m^2	51 +/− 9.1 kg/m^2	<0.001
Procedures			
Lap RYGB	83.8%	57.1%	<0.001
Open RYGB	9.1%	39.6%	<0.001
Revision RYGB	2.0%	3.3%	0.564
Adjustable band	5.1%	—	—

Group 1: Uses primarily fully standardized operative and history templates.
Group 2: Partially uses standardized operative and history templates.
y = years old; BMI = body mass index; Lap RYGB = Laparoscopic Roux-en-Y gastric bypass; Open RYGB = Open Roux-en-Y gastric bypass; Revision RYGB = Revision Roux-en-Y gastric bypass; Adjustable band = Laparoscopic or open adjustable band.

TABLE 3

Overall Documentation Rates by Practice Group

	Group 1		Group 2			
Type of data	Mean (std)	*n*	Mean (std)	*n*	Absolute difference	*P* value
Operative report*	88% (30%)	103	59% (44%)	93	29%	<0.001
Comorbidities†	98% (10%)	99	76% (31%)	89	22%	<0.001

* Total 5 operative report quality indicators.
† Total for 7 preoperative comorbidities

which did not use standardized dictation templates (Table 3). The absolute difference between the two groups for overall documentation of intraoperative quality measures was statistically significant ($P < 0.001$). Documentation rates for comorbidities as a whole were 98% for Group 1, who used fully standardized data collection forms, and 89% for Group 2, who used a variable degree of standardized forms. The absolute difference in documentation rates for preoperative comorbidities was statistically significant ($P < 0.001$).

Operative Report Documentation Rates

Group 1 performed significantly better than Group 2 in documenting four of the five intraoperative quality measures: exploration of the abdomen (95% *versus* 55%, $P < 0.001$), examination of the gallbladder (72% *versus* 19%, $P < 0.001$), evaluation of the anastomosis for leak (93% *versus* 70%, $P < 0.001$), and closure of the small bowel mesentery (94% *versus* 73%, $P < 0.001$), Table 4. Group 1 had a trend in better documentation of closure of the large bowel mesentery (87% *versus* 79%), but this difference was not statistically significant ($P = 0.113$).

Comorbidity Documentation Rates

Group 1 was statistically superior to Group 2 in documenting each of the seven comorbidities compared with Group 2, with the exception of SA (97% Group 1 *versus* 90% Group 2, $P = 0.055$), Table 5. The largest absolute difference between the groups occurred in documentation of VSD (95% Group 1 *versus* 5% Group 2, $P < 0.001$). Group 1 had a 100% documentation rate for DM, HTN, and REFLUX, whereas Group 2 was only able to achieve a maximum documentation rate of 91% (for both HTN and DJD). Subgroup analysis in Group 2 revealed no statistically significant difference in documentation of all comorbidities except REFLUX ($P = 0.04$) between the four providers in Group 2 who used a template *versus* the four that did not.

DISCUSSION

With recent discussion of P4P in surgery, the importance of documentation has taken center stage. The goal of this study was to determine how standardized forms can aid providers in documenting certain intraoperative and preoperative bariatric surgery-related quality measures. Routine use of standardized forms increased documentation of intraoperative measures by 29% and preoperative measures by 22%. When standardized templates are not used, documentation of items that are routine may suffer. For example, "exploration of the abdomen" is likely performed by most surgeons during bariatric cases. However, in our study, "exploration of the abdomen" was only documented in 55% of cases by Group 2 (compared with 95% by Group 1). Standardized forms may also help remind the provider to inquire about items that are uncommon. For instance, Group 1 had a much higher rate of documen-

TABLE 4

Documentation Rates by Intraoperative Quality Measure for Each Group

	Group 1*		Group 2			
Operative report measure	% Documented	*n*	% Documented	*n*	Absolute difference	*P* value
Exploration of abdomen	95%	107	55%	93	40%	<0.001
Evaluation of gallbladder	72%	108	19%	93	53%	<0.001
Evaluation of anastomosis for leak	93%	108	70%	93	23%	<0.001
Closure of large bowel mesentery	87%	95	79%	93	8%	0.113
Closure of small bowel mesentery	94%	95	73%	93	21%	<0.001

* Sample size differs for each indicator because of missing operative reports or for procedures where certain indicators were not applicable.

TABLE 5

Documentation Rates by Comorbidity Comparing Practice Groups

Comorbidity	Group 1		Group 2		Absolute difference	*P* value
	% Documented	*n*	% Documented	*n*		
Diabetes	100%	99	88%	92	12%	<0.001
Hypertension	100%	99	91%	92	9%	0.003
Dyslipidemia	97%	99	80%	92	17%	<0.001
Sleep apnea	97%	99	90%	92	7%	0.055
Venous stasis disease	95%	99	5%	92	90%	<0.001
Degenerative joint disease	98%	99	91%	92	7%	0.039
Reflux	100%	99	87%	70*	13%	<0.001

* Documentation of reflux was not evaluated for the first 22 patients in our study.

tation for VSD compared with Group 2. In our study, the smallest absolute difference in documentation rates between the two groups occurred for SA and DJD. Moreover, the lack of a well-established definition for VSD may have contributed to Group 2's poor documentation rate for this indicator. Definitions for DM and HTN, on the other hand, are clear and uniform, hence the higher documentation rates. This illustrates the need for uniform definitions for all bariatric-related comorbidities that are accepted and used by all bariatric surgeons.

The use of standardized templates increased the frequency of documentation of items indicative of good quality bariatric surgical care. A shortcoming of using standardized forms, however, is whether or not the recorded information is accurate. Although this is also true of items documented without the use of a standardized form, it may be easier to record inaccurately when using a standardized form. For example, if a standardized form has a series of check boxes, providers often draw a line through all of the boxes instead of checking one at a time. Additionally, providers completing standardized forms may feel obligated to mark each box even though they may not have had time to inquire about each item on the form during the patient visit. This behavior would be magnified if adherence to bariatric quality measures was a P4P measure. It is this type of phenomenon that may have been partly responsible for the extraordinarily high compliance rates with quality indicators seen among primary care physicians in Great Britain when that country introduced P4P measures in 2004 [25]. One may speculate that standardized templates may prompt healthcare providers to ask certain questions, thereby affecting quality of care, but this would have to be validated with a prospective observational study design.

Our study had several important limitations. First, we did not assess the accuracy of the information contained in the patient's chart. Therefore, we cannot comment on the validity of the information gathered from standardized forms. Second, we did not confirm in this report if improved documentation actually resulted in improved quality. However, we have plans to complete these analyses of the data in the future. Third, the abstractors in our study were not blinded to the study objective, which may lead to biases in data collection or assessment. The above-mentioned limitations result from the small number of patients, the use of templates not designed prior to the study, and the utilization of multiple abstractors. However, the data represents current practices from several institutions, including a large number of bariatric surgeons. Our study did allow us to demonstrate an increase in documentation with template use.

The need to measure the quality of care we deliver to our patients is pressing. In P4P systems, process measures are often used to represent provision of good quality care. Process measures are a priori steps that, when performed, result in good quality care. These measures are often developed by a panel of experts who review the evidence and reach a consensus on the importance of a measure. One of the true challenges in quality improvement is to measure and record outcomes once process measures have been implemented. This measurement is the only way to validate that each process actually improves quality in a quantifiable manner. In addition, this type of validation may require payors to independently and randomly audit patient charts. Validation is also necessary to ensure that documentation equates to actual adherence, which will be challenging without directly observing the doctor–patient interaction. As this may prove to be very resource intensive, validation may only be possible indirectly by measuring outcomes. Before elaborate systems of validation can be initiated, complete and consistent documentation is necessary from patient to patient. Importantly, future analysis of our data will include determining whether adherence to the indicators is associated with better outcomes such as lower complication rates and better weight loss. Currently, a lack of awareness of the need to document quality and no consensus on how and what to document are the challenges faced by the bariatric surgery community.

With rising healthcare costs and a move toward P4P, however, time is of the essence in finding solutions to these challenges.

REFERENCES

1. Medicare Pay-for-Performance Demonstration Shows Significant Quality of Care Improvement at Participating Hospitals. 2005, U.S. Department of Health and Human Services. Available at: http://www.cms.hhs.gov/media/press/release.asp?Counter = 1441.
2. The Premier Hospital Quality Incentive Demonstration: Clinical Conditions and Measures for Reporting. 2005, U.S. Department of Health and Human Services. Available at: http://www.cms.hhs.gov/quality/hospital/PremierMeasures.pdf
3. Centers for Medicare and Medicaid Services: Hospital Quality Initiative Overview. 2005. Available at: http://www.cms.hhs.gov/quality/hospital/overview.pdf
4. Agency for Healthcare Research and Quality Website. 2006, U.S. Department of Health and Human Services. Available at: http://www.ahrq.gov
5. Centers for Medicare and Medicaid Services Website. 2006, U.S. Department of Health and Human Services. Available at: http://www.cms.hhs.gov
6. McDonald CJ, Overhage JM, Dexter P, et al. A framework for capturing clinical data sets from computerized sources. Ann Intern Med 1997;127(8 Pt 2):675.
7. Gilbert EH, et al. Chart reviews in emergency medicine research: Where are the methods? Ann Emerg Med 1996;27:305.
8. Rubin HR, et al. Watching the doctor-watchers. How well do peer review organization methods detect hospital care quality problems? JAMA 1992;267:2349.
9. Dimick JB, et al. Hospital costs associated with surgical complications: A report from the private-sector National Surgical Quality Improvement Program. J Am Coll Surg 2004;199:531.
10. Khuri SF, et al. The Department of Veterans Affairs' NSQIP: The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 1998;228:491.
11. Hamel MB, et al. Surgical outcomes for patients aged 80 and older: Morbidity and mortality from major noncardiac surgery. J Am Geriatr Soc 2005;53:424.
12. Khuri SF. The NSQIP: A new frontier in surgery. Surgery 2005;138:837.
13. Khuri SF, et al. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. Ann Surg 2005;242:326;Discussion 341.
14. Dimick JB, et al. Who pays for poor surgical quality? Building a business case for quality improvement. J Am Coll Surg 2006; 202:933.
15. Allison JJ, et al. The art and science of chart review. Jt Comm J Qual Improv 2000;26:115.
16. Wu L, CM Ashton. Chart review. A need for reappraisal. Eval Health Prof 1997;20:146.
17. Peabody JW, et al. Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of three methods for measuring quality. JAMA 2000;283:1715.
18. Lawthers AG, et al. Designing and using measures of quality based on physician office records. J Ambul Care Manage 1995; 18:56.
19. Ashton CM, et al. The association between the quality of inpatient care and early readmission. Ann Intern Med 1995;122: 415.
20. Zanis DA, et al. A comparison of three methods of measuring the type and quantity of services provided during substance abuse treatment. Drug Alcohol Depend 1997;49:25.
21. DeOrio JK. Surgical templates for orthopedic operative reports. Orthopedics 2002;25:639.
22. Laflamme MR, et al. Efficiency comprehensiveness and cost-effectiveness when comparing dictation and electronic templates for operative reports. AMIA Annu Symp Proc 2005;425.
23. Pellegrini DP, S Raghu, DK Waechter, Structured electronic documentation (SED) cardiac surgery and PowerNote—a pilot project. AMIA Annu Symp Proc 2005;1076.
24. Maggard MA, et al. Quality indicators in bariatric surgery: Improving quality of care. Surg Obes Relat Dis 2006;2:423, 2006; discussion 429.
25. Doran T, et al. Pay-for-performance programs in family practices in the United Kingdom. N Engl J Med. 2006;355:375.