Surgical Endoscopy (2019) 33:679–683
https://doi.org/10.1007/s00464-019-06661-w

SAGES QUALITY, OUTCOMES, AND SAFETY

Developing minimally invasive procedure quality metrics: one step at a time

Thomas A. Aloia[1] · Timothy Jackson[2] · Amir Ghaferi[3] · Jonathan Dort[4] · Erin Schwarz[5] · John Romanelli[6]

Received: 27 December 2018 / Accepted: 14 January 2019 / Published online: 22 January 2019

Abstract

Background Despite extensive first-hand surgical experience, rank and file members of surgical societies are generally not trained in and have not therefore been included in surgical quality measure development. The purpose of this exercise was to determine if a structured quality metric design tool could bridge this gap, facilitating rapid development of focused quality metrics by minimally invasive surgeon attendees of the April 2018 SAGES Annual Meeting.

Methods Expert minimally invasive surgeons attended a 90-min workshop with didactic and interactive quality metric design sessions during the Annual Meeting. The interactive portion was formed around a novel structured quality measure development tool that graded presenting symptoms, short-term complications, and long-term disutility of care.

Results For first-time symptomatic inguinal hernia repair, first-time small to moderate size ventral hernia repair, and elective laparoscopic cholecystectomy, each workgroup was able to develop one quality, one short-term complication, and one long-term disutility metric.

Conclusions A structured quality metric design tool facilitates application of knowledge through rapid development of multifaceted, patient-centric outcomes measures by expert minimally invasive surgeons, otherwise not formally trained in metric development. The exercise also highlighted the need to rigorously define denominator populations and to guard against metric-driven undertreatment.

Keywords Surgical outcomes · Complications · Standard of care · Patient-reported outcomes

In 1965, Avedis Donabedian first proposed that medical care was subject to the concept that all results are dependent on and, importantly, can be measured by the three components *of structure, process, and outcomes* [1, 2]. Moving through these domains from structure to true patient outcomes presents different challenges, mainly based on the premise that structures and process are immediately measurable; however, true outcomes relate not only to the product of the structures and processes, but also to the interaction of the product with the end user. Although some have challenged the transferability of Donabedian's manufacturing paradigm to medicine and surgery, ultimately there are certainly more similarities than differences.

Medicine and surgery are currently awash with measurement. Unfortunately, the creators of these measures, mainly regulatory bodies, have launched into measurement of surgical outcomes without extensive surgeon or patient input. This has resulted in a confusing set of metrics, mainly consisting of process measures that seldom reflect the complete and/or complex nature of medical care, and even more rarely speak to true patient outcomes. Particularly when tied to reimbursement, surgeons rightly bristle at the application of these measures to surgical practice [3].

From a design thinking perspective, the goal of outcomes measurement is to quantify customer value [4]. Explicitly naming the customer, or patient in this scenario, is highly

✉ Thomas A. Aloia
taaloia@mdanderson.org

1 Department of Surgical Oncology, UT-MD Anderson Cancer Center, 1400 Herman Pressler Drive, Unit 1484, Houston, TX, USA

2 University Health Network, Toronto, ON, Canada

3 University of Michigan, Ann Arbor, MI, USA

4 Inova Fairfax Medical Campus, Falls Church, VA, USA

5 Society of American Gastrointestinal and Endoscopic Surgery, Los Angeles, USA

6 University of Massachusetts Medical School – Baystate Medical Center, Springfield, MA, USA

relevant, as the value of a surgical procedure should primarily, if not completely, be determined from the perspective of the patient. To this end, we propose a revised value formula, expanding on the classical definition of outcomes over cost, separating 'outcomes' into quality (the achievement of a positive result) versus safety (avoidance of a negative result/harm) (Fig. 1). This alteration of the formula conforms to the surgical consent discussion where risks and benefits are detailed and weighed in shared decision-making between surgeon and patient.

The purpose of this knowledge-to-action exercise was to improve the competence of rank and file surgical society meeting attendees in the application of this formula to guide development of outcomes measures for minimally invasive surgical procedures [5].

Materials and methods

At the April 2018 Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) Annual Meeting in Seattle, WA, an outcomes measure design workshop was sponsored by the Quality, Outcomes, and Safety Committee. The workshop consisted of three didactic lectures covering historical and modern concepts related to outcomes measurement, spanning Donabedian's structure–process–outcome rubric to current pay-for-performance measures. Subsequently, the session attendees, all experts in the field of minimally invasive surgery, participated in a structured quality metric design exercise.

$$\text{Value} = \frac{\text{Quality} - \text{Safety/Harm}}{\text{Cost}}$$

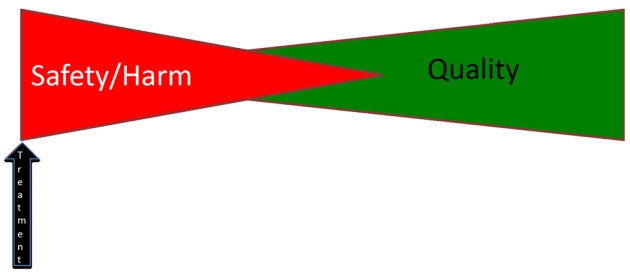

Fig. 1 A revised Value Formula to structure shared decision-making between surgeons and patients, accounting for the interaction between safety (the avoidance of harm) and quality (the achievement of an intended positive outcome)

The exercise included evaluation of both quality and safety/harm domains across the early postoperative phase and the long-term utility and disutility of the procedure. 4–10 experts completed the exercises for first-time symptomatic inguinal hernia repair, first-time small- to mid-size ventral hernia repair, and elective laparoscopic cholecystectomy.

Individual and group outputs were developed and recorded on a template form (Table 1). At the conclusion of the session, each group presented their recommendations and received feedback from other groups to refine the proposed metrics.

As the exercise was completed as a thought exercise, and there was no potential risk of harm to participants, Institutional Research Board approval was not necessary. Participants were informed at the start of the session that any potential output from the metric design exercise would be collated and tabulated into manuscript form, and that their participation in this session would be considered consent to use the material to disseminate via the said manuscript.

Results

Using this structured approach, each procedural-focused group was successfully able to reach consensus on outcomes measures within the timeframe of the session.

Regarding first-time elective laparoscopic inguinal hernia repair, the group determined that the primary surgically remediable symptom was preoperative groin pain. The corollary quality metric selected was resolution of preoperative groin area pain at four postoperative weeks, with metric failure defined as > 15% of patients with residual symptoms. The group calculated (frequency x grade) that urinary retention as the most impactful early postoperative complication/harm, yielding a metric of urinary retention requiring catheterization, and a metric failure rate of > 15% of patients. For long-term disutility, the group decided on symptomatic recurrence (defined as bulge or pain) at 1 year with > 5% indicating metric failure.

Regarding first-time elective incisional hernia repair (defect size < 10 cm × 20 cm), preoperative pain "interfering in the patient's life" was deemed the most relevant presenting symptom. Improvement over baseline in life interference from pain was, therefore, selected as the primary quality metric and the promise pain score was offered as a tool of measurement. Metric failure was defined as > 15% of patients not achieving this goal within 8 weeks of repair. The most impactful complication was determined to be wound infection, with a metric failure rate defined as > 5%. The long-term disutility selection was symptomatic hernia recurrence at 1 year, with a metric failure rate of 10%.

Regarding elective laparoscopic cholecystectomy, the remediable primary presenting symptom was determined

Table 1 A structured quality measure design tool for surgeons and patients to participate in the development of patient-centered true outcomes metrics

Characteristics of optimal quality measures		
• Relevance: Is it a meaningful measure that identifies potential for improvements?		
• Scientific soundness: Is it a scientifically valid, accurate, and reproducible measure? Is there clinical evidence to support its use? Can it provide a process–outcome link?		
• Feasibility: Is it fiscally and logistically workable? Can it be precisely specified and conducted within confidentiality parameters? Is it auditable?		
• Comprehensiveness: How extensive is the information yielded through the measure?		
Quality metric: Survival, resolution of symptoms, and/or degree of recovery		
1. Select a procedure		
2. List the most common diagnoses that indicate that procedure		
3. Symptom burden: Describe the most common symptoms patients with those diagnoses present with		
4. Life interference: Of the symptoms in #3, list 2–3 that are the most disabling to the patient (Ability to eat, walk, work, care for self, care for others, enjoy life)		
5. Circle the symptoms in #4 that improve/resolve with a "technically successful" procedure in more than 50% of cases		
6. Write a metric that addresses one of the circled items		
7. Define metric failure		
Safety metric: Short-term complication (including readmission)		
8. List the 3 most common 30-day surgical complications for the procedure	9. List the median complication grade that occurs with each complication 5-death; 4-organ failure, ICU; 3-rescue procedure (IR or OR); 2-medical management at bedside; 1-no specific intervention)	10. Multiply across, then circle the row with the highest number
1st most common (5pts)		
2nd most common (3 pts)		
3rd most common (1 pt)		
11. Write a metric that addresses the circled item		
12. Define metric failure		
Safety metric: Long-term disutility		
13. List the most common surgically induced disability present at 6–12 months postop		
14. Write a metric that addresses the item		
15. Define metric failure		

to be right upper quadrant pain with nausea. The resultant quality metric was resolution of opioid-dependent pain at two postoperative weeks, with failure defined as any narcotic consumption for abdominal pain at that timepoint. The most impactful immediate postoperative complication was determined to be bleeding or infection requiring a second intervention, such as reoperation or percutaneous drainage procedure, with a threshold for metric failure of > 5%. The most impactful long-term disutility was diarrhea. Interestingly, the group defined the corollary long-term quality metric as percent of cholecystectomy patients with diarrhea who were not offered therapy with cholestyramine. They commented that this indicated a failure to identify and treat a known and potentially remediable long-term disutility of the operation.

Discussion

Over the past two decades, there has been a steady increase in pressure on surgeons and hospitals to transparently report outcomes after surgical procedures. Patients, payors, regulators, and hospital systems all see value in the ability to identify surgical teams and surgeons who are able to safely perform procedures that meet the patients' intended quality goal(s). But accomplishing this level of transparency is easier said than done. Surgeons frequently push back on these efforts, in many cases with legitimate concerns that the data will not adequately adjust for risk, not account for unrecorded challenges, and/or fail to recognize uncontrollable aspects of recovery and outcomes.

This tension has led to the development of metrics that focus more on processes of care and less on actual patient outcomes. Compared to true patient outcomes, processes of care are more easily collected (i.e., *feasibility*) [6]. However, true patient outcomes better align with the three other qualities of a valuable outcomes measure: *relevance, scientific soundness*, and *comprehensiveness* [6]. Furthermore, most individual process measures only weakly correlate with ultimate outcomes [7]. By favoring ease of collection and simplicity, much of the power of the metric is lost. For these and other reasons, all parties are eager to move away from process measures and toward actual patient outcomes measures.

Unfortunately, surgeons have traditionally been on the sidelines of measure development, as few have been trained in the formal and rigorous processes of quality measure design. In contrast to this nidus of inexperience, almost all surgeons have comprehensive and unique perspectives on the relationships between complex interdisciplinary care models and surgical patient outcomes. The purpose of this educational exercise was to bridge the gap between the surgeon's deep knowledge perspectives and the quality measure design process.

Using a novel tool that separated outcomes into four domains based on the axes of safety/harm vs quality and short- vs long-term utility/disutility of care, groups of otherwise untrained surgeons were able to rapidly construct patient-centric surgical outcomes measure proposals for three common minimally invasive procedures. By creating quality measures based on frequency and impact of presenting symptoms and short-term harm measures based on frequency and grade of complications, they were rapidly able to hone in on "metrics that matter" [8, 9]. In contrast to process measures, when metrics include true outcomes they have the opportunity to drive a proactive approach to the issue. For example, by naming postoperative urinary retention as the most impactful outcome after inguinal hernia repair, patient-centric solutions including preoperative query regarding severity of nocturia, appropriate urologic referral, and perioperative medical management with alpha-1a adrenergic receptor antagonists are more likely to become hard-wired.

One obvious limitation of this exercise was the absence of patient input. The surgeons were asked to be surrogates for what patients see as value in each examined surgical procedure. This admittedly paternalistic approach was not, however, meant to be the final arbiter of patient value, but did produce a set of measures and a format of outcomes measure design that can easily be replicated with a patient experience/advisory group. Next steps for this particular project would be to host sessions with patients, dividing them into those who start the process amnestic to the physician expert opinion and those given the expert opinion findings up-front. Through this iterative interplay between technical (surgeon) and experiential (patient) experts and other methodologically rigorous techniques, [10] a final measure set can be developed.

Also of note, the exercise was expressly designed to focus on simple presentations. For example, re-operative and emergency surgeries are exponentially more risky than elective primary operations and these two risk factors account for the largest loss of specificity in risk models. Exclusion of these confounders, can reduce, if not eliminate concerns regarding the need for complex risk adjustment models. By engaging the far more common presentations of elective primary surgeries, the variability in outcome is narrowed and confidence in the measure increases. Further, the focus of measurement on standard presentations reduces the impulse to react to a surgical measure by limiting higher-risk patient access to surgical care.

As some surgeons legitimately practice in the setting of a high proportion of acute/emergent and/or re-operative cases, we argue for two strategies which increase fairness for these practitioners. First, limiting the scope of the metrics described here only to their primary elective cases, even if they represent a minority of their practice, reduces

metric-related anxiety by facilitating comparison of surgeons across like-risk cases. In general, we find that those surgeons with highly complex practices (e.g., complex redo abdominal wall reconstruction) perform very well on the simpler presentations of their field (e.g., small- to mid-size primary hernia repair). Second, the identical exercise as above can be performed for complex redo abdominal wall reconstruction, creating a second set of metrics that scale to complexity, allowing for expanded definitions of outcomes and moderation of failure rates to account for increased complexity. In contrast to a 'one metric fits all' approach, this proposed dual strategy manages the needs and goals of practitioners in multiple practice settings.

The perfect outcomes measure set in surgical procedures is likely to remain an elusive goal. For surgeons, the wrong answer is to disengage from the opportunity that this challenge represents. The wave of transparent reporting is upon us, we have seen the adverse impact of measures created outside of our purview and we do not want to be on the wrong side of history on this issue [11]. Further, of all surgical fields, minimally invasive surgery (MIS) is uniquely positioned to take advantage of the opportunity. Most all of the battles that MIS has fought (and won) compared to open techniques were by showing short-term patient outcomes benefits, including less pain, less disability, and faster return to normal activity. Indeed, the laparoscopic cholecystectomy battle of three decades ago is similar to the current argument over MIS versus open approach to liver resection [12]. Ultimately, (provided adequate oncologic benchmarks where applicable) patient-centered outcomes, and not process measures will determine the growth trajectory of MIS surgery. Tools like the one used in this exercise may help rank-in-file members of surgical societies to take a more active role in, and even ownership of, the surgical metric design space.

Lastly, the development of validated outcomes measures may provide a framework for vetting the value of new innovations. Here again, SAGES has been at the forefront of safe introduction of new technologies to surgery. Having a solid internal foundation of outcomes measures will facilitate the future introduction, objective testing beyond "expert opinion," and subsequent uptake of the next generation of these technologies [13].

Next Steps

In addition to cross validation of outcomes measure value with patients, surgical societies need to continue to develop a culture of outcomes reporting and the information technology systems that retrieve and integrate patient-reported outcomes data that accurately reflect postoperative recovery. In particular, structured and interoperable solutions to capture long-term disutility (in this case, 1-year hernia recurrence, diarrhea, life interference, etc.) are urgently needed. Only through accurate accounting of patient symptom, function, and life interference measures can we compare the recovery between approaches, techniques, and surgeons. Lastly, given the ability of this exercise to leverage already assembled surgical expertise for metric development, surgical societies should explore expansion of this pilot to addition annual meeting platforms.

Compliance with ethical standards

Disclosures Thomas A. Aloia, Timothy Jackson, Amir Ghaferi, Jonathan Dort, Erin Schwarz, and John Romanelli have no conflicts of interest or financial ties to disclose.

References

1. Ayanian JZ, Markel H (2016) Donabedian's lasting framework for health care quality. N. Engl J Med 375(3):205–207
2. Donabedian A (1966) Evaluating the quality of medical care. Milbank Mem Fund Q 44(Suppl):166–206
3. Khullar D. How can you tell whether your doctor is any good? The Washington Post. 2018 July 22, 2018;Sect. Health & Science
4. Martin R (2009) The Design of business. Harvard Business School Press, Boston
5. Field B, Booth A, Ilott I, Gerrish K (2014) Using the knowledge to action framework in practice: a citation analysis and systematic review. Implement Sci 9:172
6. Simpson L https://archive.ahrq.gov/chip/content/monitoring_evaluation/quality_stratagy_choice-chars_of_measures.htm
7. Dua A, Desai SS, Seabrook GR, Brown KR, Lewis BD, Rossi PJ et al (2014) The effect of surgical care improvement project measures on national trends on surgical site infections in open vascular procedures. J Vasc Surg 60(6):1635–1639
8. Teisberg EO, Wallace S (2009) Creating a high-value delivery system for health care. Semin Thorac Cardiovasc Surg 21(1):35–42
9. Wallace S, Teisberg E. Measuring what matters: Connecting excellence, professionalism, and empathy. http://braininjuryprofessional.com/2015/08/value-for-patients/
10. Alam R, Figueiredo SM, Balvardi S, Nauche B, Landry T, Lee L et al. Development of a patient-reported outcome measure of recovery after abdominal surgery: a hypothesized conceptual framework. Surg Endosc. 2018
11. Chou E, Abboudi H, Shamim Khan M, Dasgupta P, Ahmed K (2015) Should surgical outcomes be published? J R Soc Med 108(4):127–135
12. Abu Hilal M, Aldrighetti L, Dagher I, Edwin B, Troisi RI, Alikhanov R et al (2018) The southampton consensus guidelines for laparoscopic liver surgery: from indication to implementation. Ann Surg 268(1):11–18
13. Pennell CP, Hirst A, Sedrakyan A, McCulloch PG (2016) Adapting the IDEAL framework and recommendations for medical device evaluation: a modified delphi survey. Int J Surg (London England) 28:141–148