Published in final edited form as:

Gastroenterology. 2014 October ; 147(4): 725–729. doi:10.1053/j.gastro.2014.08.027.

Quality Measures And Their Emerging Importance For Trainees

Sameer D. Saini, MD, MS^{1,2}, Akbar K. Waljee, MD, MS^{1,2}, Philip Schoenfeld, MD, MSc, MSc(Epi)^{1,2}, Eve A. Kerr, MD, MPH^{1,2}, and Sandeep Vijan, MD, MS^{1,2}

¹Veterans Affairs Center for Clinical Management Research, Ann Arbor, MI

²Department of Internal Medicine, University of Michigan Medical School

“Not everything that counts can be counted, and not everything that can be counted counts.”

Healthcare in the United States is in a state of tremendous flux. As health-related expenditures have risen, often without the expected improvements in clinical outcomes, patients and payers have demanded greater value and accountability from the medical community. Many of these demands are codified in the Affordable Care Act (ACA), which contains numerous provisions to enhance value and accountability in healthcare. At the core of these provisions are efforts to measure and improve quality of care. Quality is the currency by which value is measured and accountability is issued. Indeed, the word “quality” appears more than 500 times in the ACA. Though researchers have studied quality for decades, only in the last several years have we seen payers taking a greater interest in quality measures for common GI conditions and procedures. With the endorsement of several GI quality measures by the National Quality Forum (NQF) in 2013, quality measurement will soon become the new normal for our specialty. No doubt, our trainees are entering a very different climate and practice environment than we entered. Unfortunately, available data suggest that trainees have limited knowledge of quality measures,¹ and many training programs do not include trainees in internal quality reporting programs. As educators and mentors, it is our duty to ensure that our trainees are adequately prepared to understand quality measures, how they should be optimally designed, how they can be used to improve clinical practice, and when their application might lead to unintended consequences. We believe that quality measurement and improvement is a natural and necessary compliment to the broader educational mission of GI fellowship training, lifelong learning, and providing the best possible care for patients.

What is quality, and how do we measure it?

The Institute of Medicine (IOM) has defined quality as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.” The delicate wording of this statement is worth examining more closely. First, it highlights the importance of health outcomes, a metric that is valued by patients, providers, and payers alike. Second, it acknowledges that

Address correspondence to: Sameer D. Saini, Veterans Affairs Medical Center, 2215 Fuller Road – 111D, Ann Arbor, MI 48105, Phone: 734-845-5865; sdsaini@umich.edu.

No relevant conflicts of interest exist.

the process of care is also critically important. The implication of these two statements is that quality has to do with not only the result or outcome of healthcare, but also the way in which that result is achieved. Indeed, some “good” outcomes occur not because of what our health system does, but in spite of it.

Avedis Donabedian, a physician and researcher at the University of Michigan School of Public Health, proposed a conceptual framework for quality measurement in 1966.² This framework, known as the “Donabedian Model”, is still used to this day. In Donabedian’s model, the physical and organizational **structure** of a healthcare environment supports **processes** of care that lead to desired health-related **outcomes** (Table 1). Each of these domains – structure, process, and outcome – contains elements that are determinants of quality, and some of these elements can be measured.

Structural measures of quality are those that assess the infrastructure and environment in which healthcare is delivered. Examples include the presence or absence of electronic report writing software, the use of state of the art endoscopic equipment, and the number of nurses available to staff rooms in the endoscopy unit. For the most part, structural measures of quality are not currently utilized in GI (Table 2). Notable exceptions include measures currently used as part of Joint Commission accreditation reviews, and upcoming measures to assess quality of ambulatory endoscopy units. While measures of structural quality are important tools in the quality measurement armamentarium, these measures are too far upstream to be of interest to most clinicians, who are busy with the day-to-day processes of direct patient care. That being said, without the appropriate structure, these processes of care will inevitably falter, ultimately leading to poorer outcomes.

Process measures are those that assess *how* care is delivered. Most commonly used GI quality measures are, in fact, process measures (Table 2). One example of a process measure is testing for latent infections before initiation of anti-TNF therapy in patients with inflammatory bowel disease (IBD). Process measures have several important advantages over structural measures of quality. First, they are more closely related to clinical outcomes than structural measures. Second, they focus on what high-performing clinicians and teams actually do – counseling patients with hepatitis C about avoidance of alcohol, recommending the appropriate follow-up interval after colonoscopy, performing preventive care in patients with IBD. Because they focus on clinical actions, process measures are often clearly *actionable*. That is, the remedy to low performance on a process measure is usually quite clear. For example, if a provider is not meeting a measure for counseling patients about alcohol avoidance, the potential solutions are obvious: the provider needs to be more diligent about counseling, or a system needs to be put into place to facilitate this important process.

Despite these advantages, process measures also have significant shortcomings. One downside is that they are not always easy to measure. Take, for instance, colonoscopy withdrawal technique, the process by which colonoscopy quality is determined. High quality withdrawal technique comprises four components: (1) adequate insufflation; (2) washing and suctioning of debris; (3) examination behind folds; and, (4) adequate time for mucosal inspection (withdrawal time). Withdrawal technique is “tightly linked” to clinically

meaningful outcomes.³ In other words, the link between withdrawal technique and outcomes (in this case, adenoma detection rate) has been confirmed by empiric data.⁴ However, of the four components of high quality withdrawal technique, only withdrawal time can be measured easily. But, by itself, this component is not tightly linked to outcomes.⁵ Withdrawal technique can be assessed formally, but this requires examination and rating of video recordings by trained observers, making it somewhat impractical for routine use.

Outcome measures are measures that assess clinical outcomes. One well-known example of an outcome measure is the adenoma detection rate (ADR). Outcome measures are appealing for a number of reasons. First, outcomes are ultimately what we are interested in as clinicians, and they are also of great importance to patients and to payers. Second, outcomes are relatively straightforward to measure. Third, measuring outcomes simply makes intuitive sense. But outcome measures also bring with them several important caveats that limit their utility. First, the ultimate outcome of interest may be difficult to measure reliably, forcing us to rely on intermediate or “surrogate” outcome measures. For example, post-colonoscopy colorectal cancer is a critically important outcome, but it occurs too infrequently to measure reliably at the level of the individual endoscopist.⁶ Fortunately, the ADR has been shown to be a predictor of post-colonoscopy colorectal cancer, making it a potentially useful intermediate outcome measure. Yet this raises another important limitation of outcome measures: outcomes are only partly under the control of the provider. The ADR is determined not only by the skill of the endoscopist, but also by the prevalence of adenomas in the population that is screened. Consider, for example, an outstanding endoscopist who screens a population with an adenoma prevalence of 20% (e.g., predominantly younger patients). If this provider detects 95% of all adenomas (miss rate = 5%), the ADR would be 19%. In contrast, a less skilled endoscopist might screen a population with an adenoma prevalence of 30%. If this provider only detects 70% of adenomas (miss rate = 30%), the ADR would be 21%. Thus, the less skilled provider, despite a substantially higher miss rate, would actually appear to be delivering higher quality care. This scenario might seem unlikely, but data suggest that adenoma prevalence varies even more widely than our posited numbers.⁷ One way to deal with this issue is to statistically “adjust” for differences in patient populations (“case mix”), a technique that can improve the reliability of outcome measures. For example, GI professional societies recommend different target ADRs for men versus women – an example of simple case mix adjustment. However, simple approaches are not always effective or accurate. On the other hand, more sophisticated approaches increase measure complexity, particularly if additional data collection is necessary.

It is also important to note that, unlike process measures, outcome measures are often not clearly actionable. For instance, an endoscopist with a low ADR does not necessarily know *why* his or her ADR is low. Potential causes of low ADR include not only poor endoscopist performance, but also random “noise” and case mix. Even if it were established that the endoscopist is the source of the problem, determining the specific target of remediation would likely require direct observation of the endoscopist’s technique – in other words, an attempt to improve the process of care.

Potential risks of quality measurement

Quality measures are powerful tools for improving quality of care, but like any set of tools, they can be used in a variety of ways, some of which might improve quality of care and some of which might not. Indeed, the history of quality measurement is replete with instances of unintended consequences. For instance, the current HEDIS measure for colorectal cancer screening encourages screening utilization in patients between ages 50 and 75. But evidence suggests that this measure discourages thoughtful clinical decision-making in older patients with significant comorbidity, with many undergoing screening even when the benefit is low due to limited life expectancy.⁸ Similarly, measures that encourage “treat to target” approaches for blood pressure control have been shown to encourage overtreatment and harm.⁹ Likewise, measures can inadvertently encourage providers to avoid challenging or difficult patients. For instance, one proposed measure for colonoscopy quality that is currently under consideration will assess the rate of post-procedure complications. Of course, individuals who older or are at higher risk of colorectal cancer are more likely to experience complications than those who are younger or at lower risk. Thus, one way to improve performance on this measure is to avoid performing colonoscopies on such patients, an unintended consequence that could actually reduce quality of care.

It is also important to acknowledge that quality measures can feel like an affront to our professionalism, in some instances promoting “one size fits all” medicine and discouraging us from practicing in a clinically nuanced way. A recent qualitative study suggested that measures often distract providers from patients’ concerns.¹⁰ Additionally, when measures are punitive rather than constructive, they can have a pernicious effect on our professional culture, discouraging the transparency and disclosure that are necessary for quality improvement.¹¹

Preparing our trainees for what lies ahead

Quality of care is clearly a complex topic, but it also has a sound conceptual framework and a wealth of prior experience to guide its implementation. In this context, we believe that GI training programs must adopt more formal approaches to educate our trainees on how quality is defined and measured, to actively engage them in efforts to improve quality of care, and to ultimately encourage them to take responsibility for the quality of care that they deliver. These formal approaches would incorporate didactic, problem-based learning, and practice-based components. But more importantly, they would promote a culture of continuous quality improvement among our trainees and faculty, one that encourages transparency, collaboration, and creative problem solving to improve clinical practice.

To achieve this goal, we must first ensure that our trainees are provided with ongoing feedback on their personal performance on commonly used GI quality measures. Many large GI practices now utilize feedback reports to their providers and, in some instances, public reporting of quality measures. The provision of feedback serves several purposes. First, this practice explicitly invites trainees into the culture of quality at the institution. For busy physicians and educators, it also reframes quality measurement as an educational endeavor rather than an institutional burden. Second, feedback provides trainees with an objective

assessment of their clinical performance across a variety of high-value clinical topics. Third, feedback presents trainees with a critical learning challenge: how do I interpret my own performance data? Examining a quality report often raises questions about measure definitions, specifications, and reliability. It also begs the question: what next? How should I respond to these data? The best measures are not only reliable and clinically relevant, but also actionable. Yet, not all measures are designed with this point in mind. The appropriate remedial action differs for process measures (where the remedial action is clear) versus outcome measures (where the remedial action is often unclear).

While reporting quality measures to our trainees creates a number of valuable learning opportunities, this is just a starting point. Indeed, data suggest that feedback alone has little effect on provider behavior.¹² Rather, for this approach to succeed, our trainees need a supportive environment and curricular infrastructure.¹³ Quality reports are often considered a private matter for the individual provider, but this cultural norm of privacy discourages collaborative efforts to improve quality and deters us from seeking help when we need it. Trainees could therefore benefit from regular, collaborative, safe forums to interpret and discuss their performance. As our health system becomes increasingly integrated and quality-focused, we will need an engaged and active provider citizenry that embraces continuous quality improvement in this way. At the same time, we should consider other creative, less traditional approaches. For instance, trainees could be encouraged to develop, with assistance, their own informal measures of quality, using these measures to assess their performance in areas of personal clinical interest. Such an approach will become increasingly feasible with the widespread use of electronic health records. Novel measures that are particularly useful can be shared with others. Ultimately, such approaches could meet the immediate and long-term needs of our trainees and our institutions while also fostering a culture of quality that will improve care for our patients.

Summary

Quality measurement is the “new normal” in gastroenterology. As we have argued, quality measurement presents a unique and valuable educational opportunity for our trainees. But to capitalize on this opportunity, and to train gastroenterologists who will thrive in our increasingly integrated, quality-oriented healthcare system, we need to engage trainees in formal educational programs on this topic. We also need to be thoughtful about not only the potential benefits, but also the potential harms of quality measurement. And as a profession, we need to be sure that we have a voice in defining and improving not only quality of care, but also the quality measures by which we are judged.

Acknowledgments

Drs. Saini's research is supported by a Veterans Affairs Health Services Research and Development (HSR&D) Career Development Award. Dr. Waljee's research is also supported by a VA HSR&D Career Development Award.

REFERENCES

1. Thompson JS, Lebwohl B, Syngal S, et al. Knowledge of quality performance measures associated with endoscopy among gastroenterology trainees and the impact of a web-based intervention. *Gastrointest Endosc.* 2012; 76:100–106. e1–e4. [PubMed: 22421498]
2. Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q.* 1966; 44(Suppl): 166–206. [PubMed: 5338568]
3. Kerr EA, Krein SL, Vijan S, et al. Avoiding pitfalls in chronic disease quality measurement: a case for the next generation of technical quality measures. *Am J Manag Care.* 2001; 7:1033–1043. [PubMed: 11725807]
4. Lee RH, Tang RS, Muthusamy VR, et al. Quality of colonoscopy withdrawal technique and variability in adenoma detection rates (with videos). *Gastrointest Endosc.* 2011; 74:128–134. [PubMed: 21531410]
5. Sawhney MS, Cury MS, Neeman N, et al. Effect of institution-wide policy of colonoscopy withdrawal time ≥ 7 minutes on polyp detection. *Gastroenterology.* 2008; 135:1892–1898. [PubMed: 18835390]
6. Corley DA, Jensen CD, Marks AR, et al. Adenoma detection rate and risk of colorectal cancer and death. *N Engl J Med.* 2014; 370:1298–1306. [PubMed: 24693890]
7. Saini SD, Schoenfeld P, Vijan S. Can the adenoma detection rate reliably identify low-performing endoscopists? Results of a modeling study. *Dig Dis Sci.* 2013; 58:1856–1862. [PubMed: 23456494]
8. Saini SD, Vijan S, Schoenfeld P, et al. Role of quality measurement in inappropriate use of screening for colorectal cancer: retrospective cohort study. *BMJ.* 2014; 348:g1247. [PubMed: 24574474]
9. Kerr EA, Lucatorto MA, Holleman R, et al. Monitoring performance for blood pressure management among patients with diabetes mellitus: too much of a good thing? *Arch Intern Med.* 2012;1–8. [PubMed: 22710831]
10. Powell AA, White KM, Partin MR, et al. Unintended consequences of implementing a national performance measurement system into local practice. *J Gen Intern Med.* 2012; 27:405–412. [PubMed: 21993998]
11. Bevan G, Hood C. What's measured is what matters: targets and gaming in the English public health system. *Public Administration.* 2006; 84:517–538.
12. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev.* 2012; 6:CD000259. [PubMed: 22696318]
13. Boonyasai RT, Windish DM, Chakraborti C, et al. Effectiveness of teaching quality improvement to clinicians: a systematic review. *JAMA.* 2007; 298:1023–1037. [PubMed: 17785647]

TABLE 1

The Donabedian Model

Example of Quality Measure	Adequate equipment and staffing of endoscopy unit	Quality of withdrawal technique	Adenoma detection rate	Interval cancer rate
Advantages	<ul style="list-style-type: none"> Relatively straightforward to measure Actionable (though this often requires system level changes) 	<ul style="list-style-type: none"> Under control of the provider (actionable) Often strong data linking process of care to relevant outcomes Reliability can be achieved with smaller numbers than with outcome measures No need for case mix adjustment (unlike outcome measures) 	<ul style="list-style-type: none"> Key measure of interest to patients, providers, and payers Relatively easy to measure Intuitive 	
Disadvantages	<ul style="list-style-type: none"> Limited data linking many structural measures to clinically meaningful outcomes 	<ul style="list-style-type: none"> In some instances, can be difficult to measure Data linking to outcome may be limited 	<ul style="list-style-type: none"> Require relatively large numbers for reliability Require case mix adjustment for reliability Not always actionable 	

Table 2

Fitting Selected AGA Quality Measures Into the Donabedian Model

Quality Measure Definition	Measure Type
Colonoscopy performance	
Rate of precancerous polyps and cancers identified during initial colon cancer screening exams (ADR)	Outcome
Rate of utilization of anesthesia professional for colonoscopy in ASA class I-II patients undergoing screening	Outcome
Rate of potentially avoidable complications during the “colonoscopy episode”	Outcome
Colorectal cancer screening and surveillance	
Completion of colorectal cancer screening test(s) in eligible population	Outcome
Colonoscopy interval for patients with a history of adenomatous polyps	Process
Inflammatory bowel disease	
Type, anatomic location and activity documented	Process
Preventive care: corticosteroid related iatrogenic injury – bone loss assessment	Process
Testing for latent TB before initiating anti-TNF therapy	Process
Assessment of hepatitis B status prior to initiating anti-TNF therapy	Process
Testing for <i>Clostridium difficile</i> in IBD patients who develop diarrhea	Process
Hepatitis C	
HCV genotype testing prior to treatment	Process
Hepatitis C: antiviral treatment prescribed	Process
Counseling regarding risk of alcohol consumption	Process
Counseling regarding use of contraception prior to antiviral therapy	Process
Hepatitis C: hepatitis B vaccination in patients with HCV	Process