

## Measuring Surgical Quality

Doctors are developing new ways to test—and improve—operating room performance

Any patient scheduled for surgery hopes, and maybe assumes, that his surgeon will do a high-quality job. Surgeons know better. Nearly three decades of research have made clear that some hospitals and surgeons have significantly better outcomes than others.

Exactly how to measure the quality of a surgeon's skills, however, is up for debate. Surgical volume—the number of operations of a specific kind performed at a hospital or by an individual surgeon—is known to be a good marker for quality. But it's not perfect. For example, looking only at hospitals that perform at least 125 bariatric surgeries per year, a recent review found that the rate of serious complications ranged from less than 1 percent to more than 10 percent.

In a 2018 article in the *Annual Review of Medicine*, surgeon Justin Dimick, a national leader in surgical outcomes research at the University of Michigan, and his colleague Andrew Ibrahim discuss the challenges of measuring surgical quality and offer some suggestions for better ways to accomplish that important task. We asked Dimick, currently working on an NIH-funded study of federal policies aimed at improving surgical quality, how surgeons should be evaluated. This conversation has been edited for length and clarity.

### **Should patients be concerned about physician quality?**

Yes, because some doctors have better patient outcomes than others. It's obvious to physicians that there's a lot of performance variation. It may be less obvious to patients.

## **Who measures physician quality, and why?**

There are a lot of different stakeholders measuring physician quality. Some payers, including the Medicare program, have physician performance metrics that are built into what are called pay-for-performance programs. They pay doctors incrementally more for high quality, and they penalize them for poor quality.

Some health plans give physician ratings that are based on performance measures. Other groups measure performance for public reporting, meaning it is available to consumers on various websites.

The vast majority of performance data are mostly used by physicians and hospitals that see their quality scores and then launch improvement projects around weak areas. Those data don't necessarily get into the hands of patients to help them make decisions.

## **Why is that?**

There is no consensus among physicians on the best way to measure quality—let alone communicate it to patients—because there are different ways to look at physician quality.

There's the technical domain, which is how often physicians diagnose correctly vs. incorrectly and how well they treat once they make a diagnosis. There are many challenges to measuring technical quality, and a lot of approaches have been tried over the years.

The patient's perspective—how comprehensive her doctor was or how patient his doctor was—is getting a lot of attention as a way to evaluate a physician's performance. Some health systems are starting to post patients' ratings and reviews of their physicians online.

Patient-reported experience is important, but we know it is only partially related, or completely unrelated, to the technical quality of care delivered. In other words, you can have a physician who has good technical skill, but it's not an enjoyable

process. Likewise, you could have a very pleasant physician who makes the wrong diagnosis.

### **Why is it so hard to measure a physician's technical skill?**

There are three types of measures—structural, process, and outcomes—and each type has some limitations.

Structural measures—for example, where you work—don't directly measure how an individual physician performs but might still be an indicator. *U.S. News and World Report* publishes hospital ratings every year. If you believe that high-quality hospitals employ high-quality physicians, that can be a structural marker. Another structural marker could be the volume of patients with a certain condition that a doctor sees. We all recognize that practice makes perfect, and it shouldn't be any different for physicians.

Process measures are used to check whether a physician did the right thing for a patient's specific condition. For example, if a patient has diabetes, did the doctor check his blood glucose levels periodically? Process measures do not indicate whether the patient's health improves under the physician's care.

Then the final thing is outcome measures, which means how patients fared. These are probably the most direct measures of physician performance, but they are the hardest to get right because patients are so different from each other.

---

**“Right now, the direct-to-consumer ratings are essentially useless.”**  
—Justin Dimick

---

Some physicians treat patients who are sicker than others. Directly measuring outcomes has to take into account those differences in risk of complications. An 80-year-old overweight patient with heart failure and diabetes is more likely to have surgical complications than a young, healthy athlete, regardless of the surgeon's skill.

Also, the outcomes that are most important are sometimes the hardest to measure because of the sample size. Probably the best example is mortality after surgery. For a lot of complex surgeries, mortality is so rare that it can't be

measured reliably. You can calculate the rates—how many of a surgeon’s patients died after a certain type of operation—but hardly anyone is different than the average on a statistical test because there’s not a large enough sample size.

### **So then what’s a better way to measure surgical quality?**

Video evaluation of a surgeon’s actual performance. Lots of surgeons in Michigan worked together on a project that demonstrated this a few years ago. Twenty bariatric surgeons submitted a video of themselves performing a minimally invasive gastric bypass, and at least 10 other surgeons rated each video on various aspects of technical skill. Some surgeons got an average rating of 4.8 on a 5-point scale; others were as low as 2.6.

The raters largely agreed on their ratings of individual surgeons, showing that this is a reliable way to measure what it means to be a better surgeon. We’ve also shown that the rating matters in terms of a surgeon’s patient outcomes. The lowest-rated surgeons were associated with higher complication rates, higher rates of hospital readmission, higher rates of reoperation, and higher mortality.

### **What did the surgeons think of this approach?**

When we showed the video of the highest-performing surgeon at one of our meetings, someone stood up and said, “You know what? I felt really good about what I did in the operating room until I saw that video. Now I realize I have a lot of room to improve.”

Just the fact that people came to the table and sat down and learned from each other shows that surgeons can generally handle the scrutiny with humility and with a will to improve.

Surgeons can accurately rate the skill of their colleagues by observing their technique in the operating room, studies show. Credit: Thomas Eakins/Wikimedia Commons

After this study, a surgical coaching program was launched where the highest-skilled people coached the people who were in lower ratings. There have been

studies that show surgeons perceive that coaching affects their practice in a positive way, but no studies yet that show whether it actually improves performance. We're still in the early stages of this research, so we may have something soon.

### **Do you expect video assessment to become the standard way to measure surgical quality?**

There's a lot more science to be done. It was demonstrated in one operation, but for less technically complicated operations, there may be only a very small relationship, or no relationship, between peer ratings and patient outcomes. So you can't generalize from this one study.

One would say there's a lot of enthusiasm and a lot of scientific work being done to develop the evidence base. and we need to move forward with measures like this.

### **How might video assessments actually be used in practice?**

The American Society of Colorectal Surgeons is piloting a technical-skills component to becoming board-certified in colorectal surgery. Some interesting results from that pilot showed that the book knowledge—the test that you take to be board-certified—is completely uncorrelated to the technical-skill knowledge.

There's one professional society in Japan that accredits surgeons for laparoscopic gastric cancer surgery, and they use a video to decide if a person should be accredited or not. That's an interesting way to use these videos so they actually have teeth, assuming that a surgeon's accreditation status is available to patients and/or is related to payment. That's hopefully the direction in which we will be moving once the science catches up with the practice.

### **Any possibility these video assessments could be available to consumers?**

I think that would be one of the most helpful ways these data could be used.

Right now, the direct-to-consumer ratings are essentially useless. You can go online to find a doctor who has lots of patients who like him or her, and one or

two patients who say a lot of negative things. Does that mean they are a bad doctor? No. Those ratings are obviously not scientifically based.

Now, if you had a website that collected videos of surgeons' performance and ranked them so you could see that a certain surgeon is a verified expert based on her technical skill, would that be useful to you? Absolutely.

I think we're just short one entrepreneur to see this is a good idea and create a panel of surgeons to do the assessments. Surgeons could submit a video to get rated and, even if only the highest-rated surgeons were identified to the public, that's still useful: This is a top-rated surgeon in your area. Boom.

**Readers may please note that D. L. Shah Trust brings out two e-journals on a fortnightly basis. These are mailed to those persons or institutions who are desirous of receiving them: These two e-journals are:**

- 1. Safety Info**
- 2. Quality Info**

**If you or your friends or colleagues wish to receive these journals, you may send us an e-mail requesting for the same. There is no charge for these journals. Our e-mail address is:**

**[dlshahtrust@yahoo.co.in](mailto:dlshahtrust@yahoo.co.in)   [haritaneja@hotmail.com](mailto:haritaneja@hotmail.com)   [dlshahtrust@gmail.com](mailto:dlshahtrust@gmail.com)**

**You can also access these journals on our website: [www.dlshahtrust.org](http://www.dlshahtrust.org)**

<b>Sponsored by: D. L. Shah Trust For Applied Science, Technology, Arts &amp; Philosophy Mumbai. email: <u><a href="mailto:dlshahtrust@yahoo.co.in">dlshahtrust@yahoo.co.in</a></u> Ph: 022-22838890</b>	<b>Edited by Hari Taneja, Mumbai email: <u><a href="mailto:dlshahtrust@yahoo.co.in">dlshahtrust@yahoo.co.in</a></u></b>
--	---