ImproveDx

The Newsletter of the
Society to Improve Diagnosis in Medicine
Volume 3 · Number 6
December 2016

Measuring Diagnostic Performance

By Susan Carr
Newsletter Editor

Measurement is central to improving the safety, quality, and value of healthcare services. There is debate, however, about what and how to measure and the potential for measurement to result in unintended consequences. Does desire for data focus our attention only on things that are easy to count, concealing problems that don't lend themselves to measurement? Do documentation requirements drain physicians' time and attention, to the detriment of patient care? Do patients find the information provided by quality measures helpful? Do they trust it? How often do they use the information in decision-making?

The current fascination with measurement (counting errors and quantifying the results of treatment and care) and measures (gauging quality and driving change with data) has the potential to skew understanding. Information and evidence, even when accurate and objective, may tell only part of the story.

Robert Berenson, fellow at the Urban Institute, points out a pivotal and common misunderstanding about measurement. He reports that W. Edwards Deming, who is often credited with having said, "You can't manage what you can't measure," actually said, "It is wrong to suppose that if you can't measure it, you can't manage it—a costly myth."[1] Being hard to define and measure are among the reasons why diagnostic error has been largely left out of efforts to improve patient safety and quality.[2] According to Berenson, diagnostic error has been largely invisible:

Because diagnostic error is hard to define and measure, it has been largely left out of efforts to improve patient safety and quality.

> *Another major problem… is the assumption that if a quality problem isn't being measured, it basically doesn't exist. A prime example is diagnosis errors.*[1]

In a report issued in 2015,[2] the Institute of Medicine (IOM; now called the National Academy of Medicine) lists reasons why diagnostic processes and errors are difficult to measure: the complexity of the process, number of people involved, subjective nature of clinical reasoning, and lack of feedback from patients and clinicians, as well as multiple definitions and incomplete documentation. The IOM also recommends methods for measuring the incidence of diagnostic error and assessing contributing factors, potential failures, and improvement interventions. The IOM report authors caution, however, that measurement is by its nature imperfect and refer to research[3] and commentary[4] that show different approaches to data collection and measurement result in different stories. The most reliable approach is to use more than one method.

Quality Measures for Diagnosis

In November 2016, the Centers for Medicare and Medicaid Services (CMS) published its annual list of measures intended to inform quality improvement and value-based purchasing programs.[5] Only seven of the 97 measures under consideration—information transfer at transitions and screening for malnutrition, frailty,

Also in This Issue…

SOCIETY to IMPROVE DIAGNOSIS in MEDICINE

ImproveDx is a bimonthly publication of the not-for-profit Society to Improve Diagnosis in Medicine. The opinions expressed in this publication are not necessarily those of the Society to Improve Diagnosis in Medicine or its Board of Directors.

malnutrition, alcohol, tobacco, and opioid use—relate to diagnosis. While these measures are an important start, they are small steps and may not drive significant improvement, especially in regard to the errors that cause the most harm, for example those involving misdiagnosis of cancer, cardiovascular conditions, and infections.

Assessing Diagnostic Safety

Amid this kaleidoscope of challenges, researchers continue to study ways to identify diagnostic errors, analyze processes of care, improve outcomes, and increase accountability. Some are also exploring new ways to approach safety, which involve a mindset and interventions different from traditional methods of measurement and improvement.

In "Measures to improve diagnostic safety in clinical practice,"[6] Singh, Graber, and Hofer discuss ways to assess performance consistent with traditional methods that have not yet been applied to diagnosis. Their starter set of "measurement concepts"[6(p3)] includes concrete suggestions categorized in Avedis Donabedian's framework of structure, process, and outcomes.[7] Drawn from recent research in diagnostic error, the suggestions include measures known to be relevant and successful. For example, because diagnostic errors are often shown to have occurred in cases that have no documented differential diagnosis, Singh and colleagues suggest using the availability of "web-based decision support tools and online reference material…to all providers to aid differential diagnosis"[6(p3)] to evaluate the safety of the diagnostic process.

The authors acknowledge their efforts represent the beginning of a lengthy process, which should include further study of the incidence of diagnostic error and direct input from patients and families. They also point out that the framework they use is but one approach to measurement, about which there is some debate:

> *Finally, there can be differences regarding whether it is more important to measure success or failure in diagnosis. Some experts have argued that "safety is better measured by how everyday work goes well than by how it fails." This represents a paradigm change from the current dominant focus on errors that would substantially change how we would design a measurement system of "diagnostic safety."*[6(pe2)]

Safety II and Resilience Engineering

The study of success—how systems manage to operate safely in the face of hazard—is referred to as Safety II.[8,9] Safety I includes traditional patient safety activities, such as error reporting, root cause analysis, and risk management. It focuses on decreasing the incidence of adverse events by identifying errors, learning from them, and preventing their recurrence. Safety II, on the other hand, looks proactively at everyday performance to understand why it usually goes well, even in the face of danger. While the patient safety movement usually defines safety as the absence of error and harm,[2] Safety II defines it …"as the ability to make things go right and not merely the absence of failures or adverse outcomes."[8(p418)]

Current mainstream efforts to improve patient safety use Safety I methods. Although leaders of the Safety II movement sometimes complain that narrowly applying Safety I to eliminate medical errors has constrained improvement, they acknowledge a role for Safety I. Ideally, the two approaches would coexist, and complement each other.[8,9]

Diagnosis is a complex cognitive activity, rooted as much in the human psyche and the dynamic setting in which it occurs as it is in clinical knowledge. That makes diagnosis a provocative subject for Safety II, which embraces the unpredictable, highly variable nature of clinical practice and looks for ways to support physicians in that environment.

Safety II prizes resilience, which has been studied since the 1970s as an attribute of systems and individuals.[10] Eric Hollnagel, professor at the University of Southern Denmark, explains that resilience in the sense used by Safety II has

Save the Date

May 24–25, 2017

1st Australasian
Diagnostic Error in Medicine Conference
Toward Safer Diagnosis—A Team Effort
Melbourne, Australia
www.improvediagnosis.org/page/AusDEM17

October 8–10, 2017

Diagnostic Error in Medicine
10th International Conference
Boston, Massachusetts
Stay tuned for details in 2017
www.improvediagnosis.org

With thanks for their assistance with this article:

Jeannine Cyr Gluck, MLS

Mark L. Graber, MD

Michael Grossman, MD

Lorri Zipperer, MA

evolved from indicating a system is safe because it can adapt to unexpected circumstances to one that learns and adapts continually, in good times as well as bad.[10] Resilient performance responds to opportunities for learning and improvement across the spectrum of everyday activities. The dynamic nature of diagnosis—performed under production pressure in a complex environment, with varied sources and amounts of information—demands resilience from all concerned, including patients.

Safety II for Patients and Families

Organizations and advocates encourage patients and family members to help improve patient safety by catching errors: ask questions, remind clinicians to wash their hands, examine records for mistakes, and so on. All of these activities represent new and sometimes challenging roles for patients, and they all fit into Safety I.

The IOM's 2015 report on improving diagnosis[2] casts patients and family members in a role that can be characterized as Safety II. To the best of their abilities, patients and family members are asked to report useful information about symptoms and experience related to their medical problems. Clinicians count on them to participate in making a diagnosis. The quality of the information matters, and most patients and family members receive no training in how to be effective partners in diagnosis. That training should include ways to stimulate success as well as protect against errors and harm.

If follow-up visits or conversations are involved, clinicians need to know how the patient's condition has evolved and to hear about the effect of medications and other details or questions that may seem relevant. With training on both sides, that exchange could be seen as a Safety II exploration. What do and don't we understand at this point? What else do we need to know? What is contributing to or impeding our progress?

Measurement and Safety II

In the context of Safety II and resilience, what can measurement contribute to diagnostic safety? The activities that Singh, Graber, and Hofer recommend as indicators of diagnostic safety— performing effective root cause analyses, tracking test results, alerting physicians to a changed or incorrect diagnosis, to name a few—are Safety I activities.[6] Safety II activities would change the focus of these discussions from what went wrong to what went right. Incorporating both of these perspectives in discussing diagnostic performance would provide novel opportunities to learn and improve.

Safety II activities would change the focus of these discussions from what went wrong to what went right.

Cognitive psychologist Gary Klein suggests that storytelling is a good way to record and share examples of what he describes as moments of insight—"aha moments"—in which we see a problem or situation in a new light.[11] Often those are moments of success in which an individual solves a problem, large or small, or recognizes a creative solution, similar to moments in clinical reasoning and diagnosis.

Whether through storytelling or another device, discovering and sharing factors that contribute to resilient performance can improve diagnosis and the experience of all stakeholders. Adding the Safety II mindset would align with efforts to improve the safety culture of institutions, joy in work for clinicians, and patient experience. To incorporate Safety II's focus on resilience, support, and creativity will take commitment from leaders to learn from clinicians about their daily work and from patients on the frontline and then spread those lessons.

References

1 Berenson R. JAMA Forum: if you can't measure performance, can you improve it? news@JAMA Web site. https://newsatjama.jama.com/2016/01/13/jama-forum-if-you-cant-measure-performance-can-you-improve-it/. Published January 13, 2016. Accessed November 13, 2016.

2 Balogh EP, Miller BT, Ball JR, eds. *Improving Diagnosis in Health Care.* Washington, DC: National Academies Press; 2015.

3 Levtzion-Korach O, Frankel A, Alcalai H, et al. Integrating incident data from five reporting systems to assess patient safety: making sense of the elephant. *Jt Comm J Qual Patient Saf.* 2010;36(9):402–410.

4 Shojania KG. The elephant of patient safety: what you see depends on how you look. *Jt Comm J Qual Patient Saf.* 2010;36(9):399–401.

5 Goodrich K. CMS finalizes its measures under consideration list for pre-rulemaking. The CMS blog. https://blog.cms.gov/2016/11/22/cms-finalizes-its-measures-under-consideration-list-for-pre-rulemaking/. Published November 22, 2016. Accessed November 23, 2016.

6 Singh H, Graber ML, Hofer TP. Measures to improve diagnostic safety in clinical practice. *J Patient Safety.* 2016;00:00:e1–e6

7 Donabedian A. Evaluating the quality of medical care. *Milbank Memorial Fund Q.* 1966;44(3)(suppl):166–206.

8 Braithwaite J, Wears RL, Hollnagel E. Resilient health care: turning patient safety on its head. *Int J Qual Health Care.* 2015;27(5):418–420.

9 Hollnagel E, Wears RL, Braithwaite J. Middelfart. *From Safety-I to Safety-II: A White Paper.* Denmark: Resilient Health Care Net; 2015.

10 Hollnagel E. Resilience engineering. Web site of Erik Hollnagel. http://erikhollnagel.com/ideas/resilience-engineering.html. Accessed November 20, 2016.

11 Klein G. *Seeing What Others Don't. The Remarkable Ways We Gain Insights.* New York, NY: Public Affairs; 2013.

Multiple Misdiagnoses Caused by Inappropriate Use of Genetic Testing

A 13-year-old boy's sudden, unexplained death led to a series of misdiagnoses throughout his extended family, as reported by CNN in October[1] and in an article published in *Mayo Clinic Proceedings* in November.[2]

An autopsy performed on the 13-year-old did not determine a clear cause of death, and although a sample was available, no post-mortem genetic testing was done. The boy's parents and brother underwent cardiac testing. The surviving brother was diagnosed with familial long QT syndrome (LQTS) and received an implanted defibrillator. He then received genetic testing, which confirmed the diagnosis. It was, therefore, assumed that his brother's death had been caused by LQTS. Genetic testing was performed among the extended family; more than 24 individuals were given a diagnosis of LQTS.

The brother and parents, plus other family members—none of whom displayed clinical symptoms of LQTS—sought a second opinion at Mayo Clinic. Genetic testing of the deceased brother's sample and his parents failed to show evidence of LQT1. The boy's death was then determined to have been caused by a different, non-familial form of heart disease.

Physicians involved with this case at Mayo Clinic found that a timely molecular autopsy would have supplied crucial information and likely have prevented the epidemic of misdiagnoses in this family. The original physicians treating the boy's family members relied on the interpretation supplied by the genetic testing company, to the exclusion of other test results or physical examination. The authors observe that "Genetic testing is a powerful tool, but it can also be a dangerous weapon. … More than ever, we must also strive to be wise clinicians who recognize that phenotyping still matters most."[1(p1615)]

References

1 Howard J. Misdiagnoses: a hidden risk of genetic testing. CNN. http://www.cnn.com/2016/10/31/health/genetic-testing-misdiagnosis-risk/index.html. Published October 31, 2016. Accessed November 30, 2016.

2 Ackerman JP, Bartos DC, Kapplinger JD, Tester DJ, Delisle BP, Ackerman MJ. The promise and peril of precision medicine: phenotyping still matters most. *Mayo Clin Proc.* 2016;91(11):1606-1616.

Diagnostic Error in Medicine: From IOM to Action

The Society to Improve Diagnosis in Medicine (SIDM) hosted the ninth annual Diagnostic Error in Medicine (DEM) conference in Hollywood, California, in early November. Many attendees—roughly 75% by show of hands—were experiencing DEM for the first time. They represented 40 states, the Distric of Columbia, New Zealand, Australia, the United Kingdom, Canada, and The Netherlands. Consistent with earlier DEM conferences, attendees represented a wide range of professions and organizations interested in diagnosis.

Using goals described in the Institute of Medicine's (IOM's) 2015 report[1] as a guide, the program for DEM 2016 focused on advancements in education, research, technology, patient engagement, and risk management.

DEM 2016 offered two days of preconference programing. Saturday, November 5, featured separate summits for patients and researchers. Half-day programs on Sunday included an introductory course, separate workshops for medical educators and clinicians, and a program on cognitive psychology.

Kicking off the main conference at 1 PM on Sunday, SIDM's president, Mark L. Graber, highlighted evidence that efforts to improve diagnosis have accelerated since the IOM report was released immediately before DEM 2015:

- With help from Hardeep Singh and Sue Sheridan on behalf of SIDM, the Centers for Disease Control and Prevention and the Quinnipiac Group have developed recommendations for measures to improve diagnostic performance, which are being forwarded to the secretary of the U.S. Department of Health & Human Services. The measures promote the use of autopsies and better management of test results.
- The number of Centers of Excellence for diagnosis is growing, including the newly announced Center for Dioagnostic Excellence at the Armstrong Institute at Johns Hopkins.
- SIDM's new Practice Improvement Committee will collect stories about actions hospitals, health systems, insurers, and patient safety organizations are taking to improve diagnosis. These do not have to be completed studies or fully validated results. SIDM is interested in all active initiatives in diagnosis. Groups that want to learn more and share their stories should contact Mark Graber (mark.graber@improvediangosis.org).
- Three physicians have completed their work as the first class in SIDM's Diagnostic Medicine Fellowship.
- SIDM has received a grant from the Macy Foundation to develop a curriculum on diagnosis and diagnostic error to be used in medical schools nationally.
- Internationally, the first DEM held outside the United States convened in Rotterdam in June and July 2016. John Ely's checklists have been translated into three languages, and there are SIDM interest groups in Romania, Japan, and China. The World Health Organization is developing a technical series on primary care that will include a focus on diagnostic error and be distributed to every country in the world. In Australia, Amanda Walker is writing performance measures for diagnosis that will become part of the national standards for hospitals.

Reference

1 Balogh EP, Miller BT, Ball JR, eds. *Improving Diagnosis in Health Care.* Washington, DC: National Academies Press; 2015.