Measures to Improve Diagnostic Safety in Clinical Practice
Hardeep Singh, MD, MPH,* Mark L. Graber, MD, ‡§ and Timothy P. Hofer, MD, MSc||¶

Abstract: Timely and accurate diagnosis is foundational to good clinical practice and an essential first step to achieving optimal patient outcomes. However, a recent Institute of Medicine report concluded that most of us will experience at least one diagnostic error in our lifetime. The report argues for efforts to improve the reliability of the diagnostic process through better measurement of diagnostic performance. The diagnostic process is a dynamic team-based activity that involves uncertainty, plays out over time, and requires effective communication and collaboration among multiple clinicians, diagnostic services, and the patient. Thus, it poses special challenges for measurement. In this paper, we discuss how the need to develop measures to improve diagnostic performance could move forward at a time when the scientific foundation needed to inform measurement is still evolving. We highlight challenges and opportunities for developing potential measures of “diagnostic safety” related to clinical diagnostic errors and associated preventable diagnostic harm. In doing so, we propose a starter set of measurement concepts for initial consideration that seem reasonably related to diagnostic safety and call for these to be studied and further refined. This would enable safe diagnosis to become an organizational priority and facilitate quality improvement. Health-care systems should consider measurement and evaluation of diagnostic performance as essential to timely and accurate diagnosis and to the reduction of preventable diagnostic harm.

Key Words: diagnostic errors, safety culture, quality measurement

Timely and accurate diagnosis is foundational to good clinical practice and essential to achieving optimal patient outcomes. We have learned that diagnostic errors are common, affecting approximately 1 in 20 adults each year in the United States. Yet, efforts to monitor and improve diagnostic performance are rarely, if ever, part of initiatives to improve quality and safety. Diagnosis is a complex, largely cognitive process that is more difficult to evaluate and measure than many of the other parts of the patient experience at least one diagnostic error in our lifetime and argued for efforts to improve the diagnostic process through better measurement of diagnostic performance. It reiterated that the diagnostic process is a dynamic, team-based activity that involves uncertainty, plays out over time, and requires effective communication and collaboration among multiple providers, diagnostic services, and the patient.

Measurement as a necessary first step in quality improvement is the cornerstone for many policy initiatives focused on improving quality and safety. The proliferation of health-care performance measures has been remarkable, with the National Quality Forum currently endorsing more than 600 measures in the United States. Health-care organizations (HCOs) commit substantial resources to comply with required measures from the Joint Commission and the Centers for Medicare and Medicaid Services, and many also participate in voluntary measure reporting sponsored by advocacy organizations such as the Leapfrog Group. Given the abundance of performance measures already in use, it is surprising how few are focused on diagnosis.

Multitask theory, proposed by economists Holmstrom and Milgrom, posits that when incentives put in place by an organization omit key dimensions of performance, those dimensions will receive less attention; in effect, the organization risks getting only what is measured. Thus, it would not be surprising that in the absence of specific process or outcome measures related to diagnostic safety, the HCO and its members may focus their attention elsewhere. All HCOs are resource constrained, and by necessity, they will direct their attention first to the measures specifically required by accrediting agencies and payers.

The recent IOM report creates a propitious moment to rectify this imbalance and encourages development of measures related to diagnosis. Accepting that measurement is an effective and essential component of performance improvement and that the lack of measurement is in itself deleterious, the IOM report presents a number of recommendations for improving diagnostic safety. These recommendations are informed by a wealth of evidence from the natural sciences but also are shaped by the experiences and insights of patients and families. In this paper, we highlight challenges and opportunities for developing diagnostic safety measures that can move forward at a time when the scientific foundation needed to inform measurement is still evolving. We propose a starter set of measurement concepts for initial consideration that seem reasonably related to diagnostic safety and call for these to be studied and further refined. This would enable safe diagnosis to become an organizational priority and facilitate quality improvement.

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or any other funding agency. No funding agency had any role in the design and conduct of the study; data collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.
both the opportunity and the impetus to address this dilemma. In this article, we discuss how such an initiative can move forward by balancing the need for measures and measurement with the reality that the scientific knowledge needed to inform this process is still evolving. We focus on future measures to improve diagnosis and highlight opportunities and challenges to encourage further discussion and policymaking in this area.

The Challenges of Measuring the Diagnostic Process

Despite an identified need and abundant enthusiasm to act, there is little consensus and evidence to guide selection of appropriate performance measures. Measurement begins with a definition, and the IOM defined diagnostic error as the “failure to establish an accurate and timely explanation of the patient’s health problem(s) or communicate that explanation to the patient.” This definition provides 3 key concepts that need to be operationalized: (1) accurately identifying the explanation (or diagnosis) of the patient’s problem, (2) the timely provision of this explanation, and (3) effective communication of the explanation. Although there are well-established tools for assessing communication in health care, none of these are focused primarily on discussions around diagnosis. Moreover, both the “accuracy” and the “timeliness” elements of the definition are problematic from a research perspective:

Accuracy. Inaccuracy is sometimes obvious (a patient diagnosed with indigestion who is really having a myocardial infarction), but in many other circumstances, accuracy is much harder to define. Is it acceptable to say “acute coronary syndrome” or does the label have to indicate actual infarction, or be even more specific, indicating location and transmural or not. Mental models of what is or is not an accurate diagnosis can differ even among clinicians in the same specialty.17,18 Some of these problems can be addressed by using predefined operational constructs or by using a consensus among experts, but given the uncertainties and evolving nature of diagnosis, either approach would be challenging.

Timeliness. Although we may all agree that asthma diagnosis should not require 7 visits over 3 years19 or that spinal cord compression from malignancy should probably be diagnosed within weeks rather than months,20 there are no widely accepted standards for how long diagnosis should take for any given condition. Furthermore, optimal diagnostic performance is not always about speed; sometimes, the best approach is to defer diagnosis or testing to some later time or to not make a definitive diagnosis until more information is available or if symptoms persist or evolve.

Experts have yet to define how we objectively identify clinicians or teams who excel in diagnosis and those that do not. One might argue that the best diagnosticians might be defined not only by their accuracy and timeliness but also by their efficiency (e.g., minimizing resource expenditure and limiting the patient’s exposure to risk).21 In this regard, Donabedian states, “In my opinion, the essence of quality or, in other words, ‘clinical judgment,’ is in the choice of the most appropriate strategy for the management of any given situation. The balance of expected benefits, risks, and monetary costs, as evaluated jointly by the physician and his patient, is the criterion for selecting the optimal strategy.”22 Thus, some, including authors of this paper, would argue that the measurement of the diagnostic process should really be thought of within the broader evaluation of value-based care that accounts for quality, risks, and costs, rather than using an overly simplistic focus on achieving the correct diagnosis in the shortest amount of time.23

Nevertheless, many would choose to focus on diagnostic errors as a key window into the diagnostic process, but this represents another major challenge. The instruments that organizations rely on to detect other patient safety concerns are poorly suited or fail completely in detecting diagnostic error.24 Newer approaches are needed that improve reporting by patients, physicians, and other clinicians and that take advantage of information stored in electronic medical records to detect errors or patients at risk for error.25,26 Autopsy reports, preoperative versus postoperative surgical discrepancies, escalations of care, and conducting selected chart reviews are other options for detecting missed diagnoses or preventable diagnostic delay.

Even when diagnostic errors are identified, learning from them can be challenging. Diagnosis is influenced by complex dynamics involving system-, patient-, and team-related and individual cognitive factors. While identifying these factors may be feasible in some cases,27 dissecting the root causes of these elements requires substantial inference, and there is risk of bias from looking retrospectively. Although factors can be suspected as “contributing,” it is hard to identify causal links.28 Discerning the effect of individual heuristics, biases, overconfidence, affective influences, distractions, and time constraints as well as key systems, environmental, and team factors is often not possible. For measurement to be effective and actionable, analysis needs to reflect real-world practice, in which systems, team members, and patients themselves inevitably influence the clinicians’ thought processes.28 For the many diagnoses that are made by teams, arriving at a diagnosis creates dual problems of attribution and ownership in the setting of fragmented and complex teams that exist in health care today. Thus, it might be difficult to determine who should receive the feedback that results from measurement and how to deliver useful and actionable feedback to a “team.”

Finally, there can be differences regarding whether it is more important to measure success or failure in diagnosis. Some experts29 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.”

Suggestions for Moving Forward

One of the first steps toward useful measures of diagnostic safety is to understand and use appropriate definitions of diagnostic error. In addition to the IOM definition, there are 3 other definitions of diagnostic error in active use, and each may be appropriate for research in particular circumstances. Graber et al defines it as diagnosis that was unintentionally delayed (sufficient information was available earlier), wrong (another diagnosis was made before the correct one), or missed (no diagnosis was ever made), as judged from the eventual appreciation of more definitive information.30 Schiff et al defines it as any mistake or failure in the diagnostic process leading to a misdiagnosis, missed diagnosis, or delayed diagnosis.31 Lastly, Singh defines it as missed opportunities to make a correct or timely diagnosis based on the available evidence, regardless of patient harm,32 and calls for unequivocal evidence that some critical finding or abnormality was missed or not investigated when it should have been.33 These definitions convey complimentary concepts that are useful to understand the “failure” referred to in the IOM definition and might be useful to operationalize the IOM definition as it is used in future work.

Assuming sufficient motivation exists to address and improve diagnostic safety, what measures should be considered? Recalling Donabedian’s framework, measures that focus on structures and processes can and should be considered, and where possible their downstream diagnosis-related outcomes, bearing in mind Donabedian’s admonition that none of these aspects of care are worth measuring without convincing demonstration of the causal associations between them.33 Although this framework provides
an appropriate and logical approach to begin developing measures of
diagnosis, it is critical to continue to emphasize that candidate
measures are only as good as the quality of the evidence that sup-
ports causal links between specific structures, processes, and out-
comes, underscoring the need for substantial amount of research
work that needs to be done in this area.

Table 1 describes a set of candidate measurement concepts
drawn from recent studies that focus on diagnostic error. This is
in no way a complete list but rather a conversation starter based
on emerging evidence on risks related to diagnostic safety (versus
patient safety in general). For example, many studies show lack of
timely follow-up of diagnostic test results in missed diagnosis, but
only a handful of HCOs in the United States are tracking follow-
up of abnormal test results.24,41–43 Although these proposed mea-
surement concepts are all reasonable candidates for consideration,
developing an actionable set of measures would ideally require a

<table>
<thead>
<tr>
<th>TABLE 1. Candidate Set of Measurement Concepts to Consider for Evaluation of Diagnostic Safety</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measurement Concept</strong></td>
</tr>
<tr>
<td><strong>Structure</strong></td>
</tr>
<tr>
<td>Web-based decision support tools and online reference materials are available to all providers to aid differential diagnosis.</td>
</tr>
<tr>
<td>Radiologists are available 24/7 to read stat diagnostic imaging studies in real time.</td>
</tr>
<tr>
<td>The organization has expertise to conduct a comprehensive root cause analysis in cases involving diagnostic error.34</td>
</tr>
<tr>
<td>University training programs provide specific training on diagnostic error44 that include, for example, simulated case-based learning and virtual learning platforms.</td>
</tr>
<tr>
<td>Attending staff are on site to supervise trainees 24/7.</td>
</tr>
<tr>
<td>The organization uses an interoperable and certified electronic health record with clinical decision support functionality.</td>
</tr>
<tr>
<td>The organization has an electronic health record data warehouse and informatics team to enable analytics related to diagnostic safety.</td>
</tr>
<tr>
<td>The organization has an established mechanism for providing feedback to previous clinicians when there is a significant change in diagnosis.</td>
</tr>
<tr>
<td>Health-care organizations develop processes and procedures to identify and learn from cases of diagnostic error.</td>
</tr>
<tr>
<td><strong>Process</strong></td>
</tr>
<tr>
<td>Proportion of laboratory test results or diagnostic imaging not performed within the expected turnaround time</td>
</tr>
<tr>
<td>Proportion of abnormal diagnostic test results returned but not acted upon within an appropriate time window</td>
</tr>
<tr>
<td>Proportion of clinical providers who identify a surrogate to review diagnostic test results while on vacation or when leaving employment</td>
</tr>
<tr>
<td>Proportion of patients with an unexpected hospitalization within 14 days of primary care or emergency department visit who had a differential diagnosis noted at the earlier visit.</td>
</tr>
<tr>
<td>Time from a diagnostic colonoscopy request to colonoscopy performance</td>
</tr>
<tr>
<td>Proportion of patients diagnosed with a specified target disease of interest (e.g., known diagnostic dilemmas) who received a second opinion</td>
</tr>
<tr>
<td>Proportion of patients with no-shows to cancer related diagnostic procedures</td>
</tr>
<tr>
<td>Proportion of patients who sign up for portals that actually log on to patient portals to see test results electronically</td>
</tr>
<tr>
<td>Organization monitors adenoma detection rates and provides feedback to endoscopists</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
</tr>
<tr>
<td>Proportion of patients with newly diagnosed colorectal cancer diagnosed within 60 days of first presentation of known red-flags</td>
</tr>
</tbody>
</table>
validation process that samples a broader range of informed opinion and experience in keeping with the emerging standards for the development of quality measures. Even if a particular measure is endorsed broadly, it should be considered a hypothesis to be tested. Empirical confirmation of its beneficial effect on patient outcomes should be demonstrated before it can be considered a standard to which organizations are held accountable, an essential step that is rarely considered in the development of performance measure sets.

A real challenge to implementing performance measurement in diagnosis is that harm might outweigh the benefit. Launching more measures, especially measures lacking robust evidence, tends to alienate front-line caregivers and HCOs already overburdened with other performance measures.54 Recently, experts have called for a moratorium on new measures, citing concerns that flawed measures will be used for public reporting and value-based purchasing.12,15 Turning again to the theory of performance measurement, Holmstrom and Milgrom observe that “the desirability of providing incentives for any one activity decreases with the difficulty of measuring performance in any other activities that make competing demands on the [provider]’s time and attention.” A concern that follows from this observation is that unintended consequences of performance measures will inevitably emerge and undermine efforts to improve diagnostic safety. One could easily imagine that measures of underdiagnosis might lead to higher utilization of unnecessary tests.

Summary and Recommendations

Measurement, benchmarking, and transparency of performance are playing a major role in improving health care. Current performance measures pertain almost exclusively to treatment, and a recent IOM report has strongly endorsed broadening this focus to include diagnosis. We cannot make progress toward this goal without advancing the science of measurement around diagnostic performance. Compared with most performance measures, diagnostic safety may be particularly salient to physicians and their teams, given how central diagnosis is to our professional identity and the degree of control that physicians exert over the diagnostic process.

However, the IOM also recognizes the importance of system and organizational factors in improving diagnosis. For example, improved communication and care coordination and large scale initiatives to measure and improve care delivery (such as implementation of accountable care organizations) are important targets. The United Kingdom has already embraced measurement in its large initiative focused on improving the timeliness of cancer diagnosis.55 and the United States could follow this lead as a first step in its large initiative focused on improving the timeliness of cancer care. The United Kingdom has already embraced measurement of unnecessary tests.

Innovation in diagnostic measurement, benchmarking, and transparency offers opportunities to improve diagnostic safety. In particular, measurement should be more closely tied to improving the science of diagnosis. As new measures are developed, questions of measurement should be considered: What are the appropriate time intervals to diagnose specific conditions of interest that are frequently associated with diagnostic error? How can we measure competency in clinical reasoning in real-world practice settings? What measurable physician or team behaviors characterize ideal versus suboptimal diagnostic performance? What system properties translate into safe diagnostic performance, and how can we measure those? How do we leverage information technology, including electronic health records (EHRs), to help measure and improve diagnostic safety?

How do we leverage patient experiences and reports to measure and improve diagnostic safety?

Pioneering organizations can begin by identifying “missed opportunities in diagnosis” or “diagnostic safety concerns.”52 For example, both Kaiser Permanente and the Department of Veterans Affairs are involved in initiatives to improve follow-up of abnormal test results.54,55 The case for measuring diagnostic outcomes in certain high-risk areas such as cancer diagnosis has also become clear.57 Nearly a third of patients with colorectal cancer have missed opportunities for an earlier diagnosis.48,53 Thus, outcome measures could be considered, such as ratio of early stage to late stage colorectal cancer diagnosed within the previous year and proportion of patients with newly diagnosed colorectal cancer diagnosed within 60 days of first presentation of known red flags.1,52

HCOs should also consider using their EHRs to enable diagnostic safety measurement. Although most HCOs are now using EHRs, very few are doing any analytics for patient safety improvement.12 In addition to using digital data to identify patients with potential diagnostic process failures, the EHR could be leveraged for recognizing incorrect diagnosis and internal inconsistencies suggestive of mislabeled diagnosis (patient with “coronary artery disease,” despite normal coronary angiogram; patient with “COPD” with normal lung function tests). This process would require HCOs to better capture and use structured clinical data in an electronic format for safety improvement, for which the time is now ripe.59

Additionally, in any efforts to measure underdiagnosis, it is important that attention also be paid to overdiagnosis, acknowledging that overdiagnosis has its own measurement-related conceptual challenges.61 We should learn from the mistakes of performance measurement in the treatment realm, where a single-minded focus on undertreatment in highly monitored areas of practice has led to harmful instances of overtreatment.52 We should also consider how perspectives from both patients and their care teams (physicians and other team members) can help develop novel measurement approaches that involve asking them directly about the diagnostic process and their roles. This approach is consistent with the fact that diagnosis is a “team sport” where patients play a critical role.63

Some experts caution against too much emphasis on measurement to guide decisions because of unknown and unknowable data.64 Nevertheless, evidence suggests it is now time to address measurement of diagnostic safety while balancing to avoid both underdiagnosis and overdiagnosis. We propose a starter set of measurement concepts for initial consideration that seem reasonably related to diagnostic quality and safety and call for these to be studied and further refined. This would enable safe diagnosis to become an organizational priority and facilitate quality improvement. Meanwhile, researchers should work on the evidence base needed for more rigorous measurement of structure and process elements that are connected to the real clinical outcomes of interest, more timely and accurate diagnosis, and less preventable diagnostic harm.

**REFERENCES**


