Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework

Hardeep Singh,1 Dean F Sittig2

ABSTRACT
Diagnostic errors are major contributors to harmful patient outcomes, yet they remain a relatively understudied and unmeasured area of patient safety. Although they are estimated to affect about 12 million Americans each year in ambulatory care settings alone, the conceptual and pragmatic scientific foundation for their measurement is under-developed. Health care organizations do not have the tools and strategies to measure diagnostic safety and most have not integrated diagnostic error into their existing patient safety programs. Further progress toward reducing diagnostic errors will hinge on our ability to overcome measurement-related challenges. In order to lay a robust groundwork for measurement and monitoring techniques to ensure diagnostic safety, we recently developed a multifaceted framework to advance the science of measuring diagnostic errors (The Safer Dx framework). In this paper, we describe how the framework serves as a conceptual foundation for system-wide safety measurement, monitoring and improvement of diagnostic error. The framework accounts for the complex adaptive sociotechnical system in which diagnosis takes place (the structure), the distributed process dimensions in which diagnoses evolve beyond the doctor’s visit (the process) and the outcomes of a correct and timely “safe diagnosis” as well as patient and health care outcomes (the outcomes). We posit that the Safer Dx framework can be used by a variety of stakeholders including researchers, clinicians, health care organizations and policymakers, to stimulate both retrospective and more proactive measurement of diagnostic errors. The feedback and learning that would result will help develop subsequent interventions that lead to safer diagnosis, improved value of health care delivery and improved patient outcomes.

INTRODUCTION
Diagnostic errors are major contributors to harmful patient outcomes,1–7 yet they remain a relatively understudied and unmeasured area of patient safety.8 Diagnostic errors are estimated to affect about 12 million Americans each year in ambulatory care settings alone.9 Many studies, including those involving record reviews and malpractice claims, have shown that common conditions (ie, not just rare or difficult cases) are often missed, leading to patient harm.1 4 10 11 However, both the conceptual and pragmatic scientific foundation for measurement of diagnostic errors is underdeveloped. Healthcare organisations do not have the tools and strategies to measure diagnostic safety and most have not integrated diagnostic error into their existing patient safety programmes.12 Thus, diagnostic errors are generally underemphasised in the systems-based conversation on patient safety.

Many methods have been suggested to study diagnostic errors, including autopsies, case reviews, surveys, incident reporting, standardised patients, second reviews and malpractice claims.13 While most of these have been tried in research settings, most organisations do not use any of these systematically to measure and monitor diagnostic error in routine practice. Most methods look at a select sample of diagnostic errors and do not individually provide a comprehensive picture of the problem. Compared with other safety concerns, there are also fewer sources of valid and reliable data that could enable measurement. Furthermore, unique challenges to reliable measurement such as difficult conceptualisation of the diagnostic process and often blurred boundaries between errors related to diagnosis versus those in other healthcare processes (eg, screening, prevention, management or treatment) make measurement more difficult. Because improving measurement of patient safety is essential to reducing adverse events,14 further progress toward reducing
diagnostic errors will hinge on our ability to overcome these measurement-related challenges.

In 2015, the Institute of Medicine (IOM) will release a comprehensive report on diagnostic errors in continuation of the IOM Health Care Quality Initiative that previously published the landmark IOM reports on quality and safety.\textsuperscript{15} The complexity of measuring diagnostic errors is thus bound to gather widespread attention. In order to lay a robust groundwork for measurement and monitoring techniques to ensure diagnostic safety, we recently developed a multifaceted framework to advance the science of measuring diagnostic errors (Safer Dx framework: figure 1). This framework provides a conceptual foundation for system-wide safety measurement, monitoring and improvement and is further described in this paper. We cannot measure what we cannot define and thus for purposes of this paper, we used an operational definition of diagnostic error that has been extensively used in our work.\textsuperscript{16} In brief, we define diagnostic errors as missed opportunities to make a correct or timely diagnosis based on the available evidence, regardless of patient harm.\textsuperscript{16}

**FRAMEWORK RATIONALE**

We envision the need for the Safer Dx framework because current organisational processes, procedures and policies do not necessarily facilitate or focus on measurement of diagnostic safety, which is essential for creating future feedback and learning to improve patient care. The Safer Dx framework is expected to facilitate feedback and learning to help accomplish two short-term goals: (1) refine the science of measuring diagnostic error and (2) make diagnostic error an organisational priority by securing commitment from organisational leadership and refocusing the organisation’s clinical governance structure. High quality diagnostic performance requires both a well-functioning healthcare system along with high-performing individual providers and care teams within that system.\textsuperscript{17} Therefore, the framework posits that ideal measurement and monitoring techniques should facilitate feedback and learning at all of these levels.

**OPERATIONALISING THE SAFER DX FRAMEWORK**

Our framework follows Donabedian’s Structure–Process–Outcome model,\textsuperscript{18} which approaches quality improvement in three domains: (1) structure (characteristics of care providers, their tools and resources, and the internal physical/organisational setting along with the external regulatory environment); (2) process (both interpersonal and technical aspects of activities that constitute healthcare); and (3) outcome (change in the patient’s health status or behaviour). We made several modifications to adapt these three domains to diagnostic error-related measurement work.

**Structure**

Given the widespread adoption of health information technologies and its significant implications for both measuring and improving the distributed diagnostic process that evolves over multiple care settings, we need to understand the ‘sociotechnical’ context of diagnosis. Our work illustrates the many technical and non-technical (workflow, organisational, people and internal and external policy related) variables that are likely to affect diagnostic safety, supporting the need for a multifaceted sociotechnical approach to improve diagnosis-related performance and measurements.\textsuperscript{19} Examples of underlying vulnerabilities that require sociotechnical approaches include failure to track patients and their abnormal findings over time, failure to recognise ‘red flags’ in patient presentation even within health...
information technology-enabled care settings, and lack of reliable quality measurement and feedback systems to analyse diagnostic performance comprehensively. Within the sociotechnical context, many system-related factors could shape cognitive performance and critical thinking and make clinicians vulnerable to losing situational awareness.20 External structural factors that can affect diagnostic performance and measurement include clinical productivity and legal pressures, reimbursement issues, administrative demands and other confounding factors related to ongoing mandatory quality measurements. Thus, the framework includes factors such as payment systems, legal factors, national quality measurement initiatives, accreditation and other policy and regulatory requirements.

In our framework, structure reflects the entire complex adaptive sociotechnical system and has eight technical and non-technical dimensions21 as described in table 1.

### Prerequisites of safer diagnosis

In the absence of specific diagnosis-related measurements, organisations cannot maintain situational awareness of their diagnostic safety and thus our framework proposes that improved measurement of diagnostic error should lead to better collective mindfulness and organisational learning.24 Collective mindfulness is a concept from high-reliability organisations that emphasises awareness and action on the part of all involved. It describes the ability of everyone within an organisation to consistently focus on issues that have the potential to cause harm without losing sight of the ‘forest among the trees’25 and to act effectively upon data (even when signals are subtle) to positively transform the organisation.26

The Safer Dx framework is expected to lead to learning and feedback to improve diagnostic safety. While organisational safety and risk management personnel often do not have the tools to measure and address diagnostic error,27 achieving collective mindfulness acknowledges that organisations and practices will need to become ‘preoccupied’ with diagnostic errors to push measurement ahead. Health care organisations (HCOs) will need to gather ‘intelligence’ related to diagnostic safety through multiple retrospective and prospective surveillance methods to inform good measures and solutions. Organisational learning about diagnostic safety will lead to better calibration related to diagnostic performance at both provider and system level and better methods for assessment and feedback about what matters in improving safety.28 Provider calibration would lead to better alignment between actual and perceived diagnostic accuracy and system calibration would entail better alignment between what needs to be measured and what the organisation is actually measuring. Providers receive very little, if any, feedback on the accuracy of their diagnoses and are often unaware of their patients’ ultimate outcomes.29 This also applies to both small clinics and larger institutions where such feedback could be useful to change structures and

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware and software</td>
<td>Computing infrastructure used to support and operate clinical applications and devices</td>
</tr>
<tr>
<td>Clinical content</td>
<td>The text, numeric data and images that constitute the ‘language’ of clinical applications</td>
</tr>
<tr>
<td>Human–computer interface</td>
<td>All aspects of technology that users can see, touch or hear as they interact with it</td>
</tr>
<tr>
<td>People</td>
<td>Everyone who is involved with patient care and/or interacts in some way with healthcare delivery (including technology). This would include patients, clinicians and other healthcare personnel, information technology (IT) developers and other IT personnel and informaticians</td>
</tr>
<tr>
<td>Workflow and communication</td>
<td>Processes to ensure that patient care is carried out effectively</td>
</tr>
<tr>
<td>Internal organisational features</td>
<td>Policies, procedures, work environment and culture</td>
</tr>
<tr>
<td>External rules and regulations</td>
<td>Federal or state rules (eg, CMS’s Physician Quality Reporting Initiative22 or requirements for associating a diagnosis with a request for a diagnostic test23) that facilitate or constrain preceding dimensions</td>
</tr>
<tr>
<td>Measurement and monitoring</td>
<td>Processes to evaluate both intended and unintended consequences</td>
</tr>
</tbody>
</table>

CMS, Centers for Medicare & Medicaid Services.

Table 1 Sociotechnical dimensions21 comprising the ‘structure’ of the Safer Dx framework

Process

Processes in the Safer Dx framework account for the fact that diagnosis evolves over time and is not limited to events during a single provider visit. Accordingly, the framework includes five interactive process dimensions: (1) the patient–provider encounter (history, physical examination, ordering tests/referrals based on assessment); (2) performance and interpretation of diagnostic tests; (3) follow-up and tracking of diagnostic information over time; (4) subspecialty and referral-specific factors; and (5) patient-related factors.10

Outcome

Outcomes in our Safer Dx framework include intermediate outcomes, such as safe diagnosis (correct and timely as opposed to missed, delayed, wrong or over-diagnosis), and account for patient and healthcare delivery outcomes. In much of the measurement-related work, we envision that the intermediate outcome of a safe diagnosis would be the default primary outcome.
processes involved in diagnosis and lead to changes in policy and practice to reduce diagnostic error.

The application of the Safer Dx framework emphasises learning about diagnostic errors, which should lead to development of better measurement tools and rigorous definitions of what missed opportunities the organisation should focus on. Thus, it is expected to provide more evidence to support measurement including the iterative development of rigorous ‘sharp-end’ outcome measures to determine errors as they occur as well as ‘blunt-end’ measures of system-level diagnostic performance.

CHALLENGES OF REAL-WORLD DIAGNOSTIC ERROR MEASUREMENT

While the Safer Dx framework could provide a useful impetus for using both retrospective and prospective methods to improve the ‘basic science’ of measuring diagnostic safety, several challenges remain. Most measurement methods are still evolving and have limitations. In this section, we illustrate challenges of certain current methods to routinely measure diagnostic errors.

As Graber et al have emphasised, there is ‘not a single HCO in the United States that is systematically measuring the rate of diagnostic error in its clinics, hospitals, or emergency departments.’ One of the reasons is that HCOs do not know how to find these errors. Nevertheless, organisations will need to start somewhere before we arrive at better measurement and monitoring techniques.

While there have been some early successes with voluntary reporting by providers, there are no standardised mechanisms or incentives for providers to report diagnostic errors. Further, diagnostic error incident reports from providers are not included in the standardised error reporting protocols in the common formats developed by the Agency for Healthcare Research and Quality. It has also become obvious that quality improvement team support and local champions are needed to sustain reporting behaviour. At present, few HCOs allocate the ‘protected time’ that is essential for clinicians to report, analyse and learn from safety events. A recently implemented programme at Maine Medical Center was found to be resource intensive and faced challenges in sustainability as well as in identifying the underlying contributory factors and developing and implementing strategies to reduce error. Even if these challenges could be overcome, voluntary reporting alone cannot address the multitude of complex diagnostic safety concerns. Many times, physicians are unaware of errors they make. Although accounts from providers necessarily play an important role in understanding breakdowns in the diagnostic process, provider reporting can only be one aspect of a comprehensive measurement strategy. Patient reports could be a source of measurements about events not discovered through other means and recent studies have collected data about diagnostic errors from patients. For example, of 1697 patients surveyed in seven primary care practices, 265 reported physician mistakes and 227 of these involved a diagnostic error. Although patient reporting has significant future potential, it has not been validated as a method to measure diagnostic error yet.

While medical record reviews are considered valuable for detecting diagnostic errors, non-selective record review is burdensome and resource intensive. To improve the yield of record reviews, we have developed ‘triggers’ to alert personnel to potential patient safety events, which enables selective targeted review of high-risk patients. The robust electronic health record (EHR) and large clinical data warehouses within the Department of Veterans Affairs (VA) healthcare system have allowed us to pursue research on electronic trigger development, enabling us to identify preventable conditions likely to result in harm. Triggers can be used both retrospectively and prospectively, but they are yet to move beyond the research environment into real-world practice.

The use of administrative data (eg, use of billing data as proxy for evolution of clinical diagnosis) to identify possible diagnostic errors is not sufficiently sensitive to detect events of interest. Recent work has tried to refine approaches to use administrative datasets for measurement of diagnostic errors related to stroke and myocardial infarction. However, this approach alone is unlikely to become a viable basis for rigorous measurement in and of itself in part because of the absence of clinical details related to diagnostic evolution and corresponding care processes.

Prospective methods such as use of unannounced standardised patients help to overcome limitations of other methods, such as chart reviews, which often lack certain details including those related to patient context. However, direct observations and simulations, though relatively rigorous, have distinct drawbacks; direct observations are expensive, and simulations may not generalise well to real-world practice. An alternative strategy is to develop means of gaining more valid and actionable insights from existing data sources in a proactive fashion. For example, many organisations are conducting safety huddles and reviews of patient presentations, though we are unaware of any efforts to monitor these sources for diagnosis-related issues. Similarly, we found that peer review may be a useful tool for healthcare organisations when assessing their sharp-end clinical performance, particularly for diagnosis-related safety events. However, current peer review systems are inadequate and need to include root cause analysis approaches and performance improvement thinking to detect and address system issues. This enhancement of existing peer review programmes could greatly improve processes for diagnostic self-assessment, feedback and improvement.
Once HCOs start to measure diagnostic safety, ideally using multiple methods (both retrospective and prospective), the findings would lead to further refinement of the science of measuring diagnostic error. This could lead to creation of standardised methods for measuring and monitoring diagnostic error and determining the effects of intervention or improvement. To make this into a reality, external rules, regulations and policies would be necessary to make diagnostic error an organisational priority and to secure commitment from the organisational leadership. Once we have the focus of the organisation’s clinical governance structure, the Safer Dx framework suggests that two additional challenges must be overcome for conducting and using measurements at the institutional level.

First, information must be directed to the right people in the appropriate organisational roles. Currently, most institutions do not have dedicated personnel who can analyse and act upon this information. Risk and safety managers routinely collect quality and safety data from multiple sources and are often privy to data unavailable to others, but they have little guidance on how to approach diagnostic errors.\(^4^9\) Second, measurements must generate feedback and learning to improve calibration. However, currently the mechanisms of generating and delivering feedback at the organisational level and provider level remain unknown.

To facilitate measurement and learning, policy makers working on monitoring and quality improvement and researchers would need to address the following questions about giving feedback: (1) what type of content is appropriate for feedback (eg, a return visit based on administrative data which may or may not be related to error vs a clear preventable diagnostic adverse event determined only after a comprehensive review of an incident), (2) what should be the method of delivery (eg, verbal vs written; in person vs by phone; individual vs group; anonymous vs identified), (3) what should be the timing and frequency (eg, delayed until an event is clear vs real-time as details become available) and (4) how do we create feedback that is unambiguous, non-threatening and non-punitive. Use of the Safer Dx framework should lead to shared accountability beyond the clinicians directly involved, including improving the processes and systems they practice in. Because the framework also focuses on value as an outcome, it should help address the challenge of ensuring we mitigate unintended consequences due to feedback-related hypervigilance, for example, due to increased testing/treatment.

**SUGGESTED APPLICATION OF THE SAFER DX FRAMEWORK AND NEXT STEPS**

Use of the Safer Dx framework accounts for the complexity of diagnostic error and will lead to addressing the sociotechnical context while understanding and preventing error. We envision that it will help overcome some of the challenges in exploration, measurement and ultimately improvement of the diagnostic care process within a technology-enabled healthcare system. We have used the approach suggested by this framework in our current work on measuring and reducing delays from missed follow-up of abnormal diagnostic test results in EHR-enabled healthcare. The Safer Dx framework has enhanced our conceptual understanding of the problem and informed intervention development. For example, we often suspected specific aspects of the EHR led to missed follow-up (such as the software or the user interface)\(^5^0\) but we learnt that the user interface is only one, albeit an important, part of the problem. Conversely, we identified and fixed a software configuration error that prevented providers at our site from receiving notifications of abnormal faecal occult blood test results,\(^5^1\) but in doing so we also found problems related to organisational factors, workflow and human–computer interaction, all of which needed to be addressed to improve follow-up.\(^5^2\) In a different study, we found that the same technical intervention to improve test result follow-up in two different VA facilities led to improved outcomes in only one facility.\(^5^3\) The findings remind us about the importance of accounting for the local sociotechnical context as well as the external financial and regulatory environment in which safety innovations are embedded.\(^5^4\)

The Safer Dx framework also encourages its users to leverage the transformational power of state-of-the-art health information technology and the EHR in developing interventions and feedback for improvement. Based on the Safer Dx framework, organisations should be harnessing a wealth of electronic data and these data need to be put to use to provide actionable information on diagnostic error. As an example, with funding from the VA,\(^5^5\) we are now developing an automated measurement and surveillance intervention to improve timely diagnosis and follow-up of cancer-related abnormal findings and applying the Safer Dx framework. We are using a large, national data warehouse to trigger medical records with evidence of potential delays in follow-up of cancer-related abnormal test results\(^5^0\) and have continually worked over past 2 years to make measurements more valid and reliable. Additionally, to integrate near real-time surveillance and communication of information about at-risk patients into the point of care, we are adhering to many of the organisational learning elements of the Safer Dx framework. Safer Dx encourages the involvement of organisation’s existing safety/quality improvement (QI) infrastructure, including risk managers and safety managers, and to more effectively leverage these personnel to take action and improve collective mindfulness. Through qualitative work, we are working with our organisational partners to ensure that the data we provide them will be used meaningfully, will lead to improvement in collective mindfulness and learning,
and be fed back to the providers on the frontlines to improve diagnostic processes. The framework reminds us that various internal and external policies and measurement procedures which, if not designed and implemented with the entire sociotechnical context in mind, can affect diagnosis performance and measurements.

The Safer Dx framework could also help jump-start the improvement of diagnostic safety measurement by informing the design of proactive risk assessment strategies at the organisational or practice level. As used in other high-risk settings and industries, proactive risk assessment involves self-assessment to identify specific areas of vulnerability and ensure readiness to apply preventive measures to promote safety and reduce adverse events. Stakeholders, including policymakers, QI/safety personnel or researchers, could help develop tools for risk assessment based on the framework. For example, with support from the Office of the National Coordinator for Health Information Technology, we previously used a rigorous, iterative process to develop a set of nine self-assessment tools to optimise the safety and safe use of EHRs using our eight-dimension sociotechnical model.66–68 These tools, referred to as the Safety Assurance Factors for EHR Resilience guides, are each organised around a set of ‘recommended practices’69. The goal is also to change local culture and foster the implementation of risk-reducing practices. Similar tools could be developed based on Safer Dx framework and use a multifaceted approach to assess risk proactively and drive improvements in the quality and safety of diagnosis.58

Last, policymakers working on diagnostic error-related measurement issues would need more evidence on what structural, process and outcomes-related issues need to be measured for improving diagnostic safety. This groundwork could inform the development of future performance measures at a national level with partnership with the National Quality Forum.

SUMMARY
Diagnostic errors are considered harder to tackle8 and remain elusive to improvement efforts in part because they are difficult to define and measure.60 A new conceptual framework (the Safer Dx framework) can be used by a variety of stakeholders including researchers, clinicians, policymakers and healthcare organisations to stimulate the development of local, regional and national strategies and tools for measuring and monitoring diagnostic error. The resulting knowledge would be instrumental to lay the foundation for both retrospective and more proactive measurement of diagnostic errors. The feedback and learning that would result will help develop subsequent interventions that lead to safer diagnosis, improved value of healthcare delivery and improved patient outcomes.

Contributors HS and DFS have significantly contributed to the planning, conducting and reporting of the work reported in this article. Concept and design; analysis and interpretation; drafting and revision of the manuscript; administrative and technical support: HS, DFS.

Funding HS is supported by the VA Health Services Research and Development Service (CRE 12-033; Presidential Early Career Award for Scientists and Engineers USA 14-274), the VA National Center for Patient Safety and the Agency for Healthcare Research and Quality (RO1HS022807). This work is supported in part by the Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety (CIN 13-413). HS presented the Safer Dx framework to the Institute of Medicine Diagnostic Error in Health Care Consensus Study Committee on 6 August 2014.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/