Issue Brief 1

Operational Measurement of Diagnostic Safety:

State of the Science

Agency for Healthcare Research and Quality

PATIENT SAFETY

Operational Measurement of Diagnostic Safety:

State of the Science

Prepared for:
Agency for Healthcare Research and Quality
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract Number HHSP233201500022I/75P00119F37006

Prepared by:
Center for Innovations in Quality, Effectiveness, and Safety (IQuESt), Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX
Hardeep Singh, M.D., M.P.H.
Andrea Bradford, Ph.D.

MedStar Health Institute for Quality and Safety, AHRQ Patient Safety Learning Lab Grantee, and AHRQ National Advisory Council
Christine Goeschel, Sc.D., RN

AHRQ Publication No. 20-0040-1-EF
April 2020

This project was funded under contract number HHSP233201500022I/75P00119F37006 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of AHRQ. Readers should not interpret any statement in this product as an official position of AHRQ or of the U.S. Department of Health and Human Services. None of the authors has any affiliation or financial involvement that conflicts with the material presented in this product.

Suggested citation: Singh H, Bradford A, Goeschel C. Operational Measurement of Diagnostic Safety: State of the Science. Rockville, MD: Agency for Healthcare Research and Quality; April 2020. AHRQ Publication No. 20-0040-1-EF.

Introduction

Diagnostic errors pose a significant threat to patient safety, resulting in substantial preventable morbidity and mortality and excess healthcare costs.[1] Diagnostic errors are also among the most frequent reasons for medical malpractice claims.[2] In 2015, the National Academies of Sciences, Engineering, and Medicine (NASEM) highlighted the scope and significance of diagnostic safety in its report *Improving Diagnosis in Health Care*.[3] Among NASEM's recommendations is a call for accrediting bodies to require healthcare organizations (HCOs) to "monitor the diagnostic process and identify, learn from, and reduce diagnostic errors and near misses in a timely fashion."[3]

Measurement of diagnostic performance is necessary for any systematic effort to improve diagnostic quality and safety. While numerous healthcare performance measures exist (the National Quality Forum [NQF] currently endorses hundreds of measures),[4,5] none is being used routinely to assess and address diagnostic errors. Only a few U.S. healthcare organizations have explored measurement of diagnostic safety, and the development of diagnostic safety measures remains in its infancy.[6] However, diagnostic errors are increasingly prominent in the national conversation on patient safety.

Although a coordinated national strategy to measure diagnostic safety remains an aspirational goal, recent research has yielded practical guidance for HCOs to start using measurement to enhance diagnostic safety.

Several stakeholders have recently launched initiatives and research projects to advance development and implementation of diagnostic safety measurement.[7-11] These stakeholders include the Agency for Healthcare Research and Quality, the NQF, the Centers for Medicare & Medicaid Services, the Society to Improve Diagnosis in Medicine, and philanthropic foundations such as the Gordon and Betty Moore Foundation. It is thus reasonable to expect that HCOs will face increasing expectations to measure and improve diagnostic safety as part of their quality and safety programs.

In parallel with calls for improving diagnostic safety is a growing emphasis on the concept of a learning health system (LHS). LHSs can be conceptualized at various levels, including the care team level, the HCO level, the external level (e.g., State or national), or within a specific improvement initiative or collaborative. In LHSs, leaders are committed to improvement, outcomes are systematically gathered, variation in care within the system is systematically analyzed to inform and improve care, and a continuous feedback cycle facilitates quality improvement based on evidence.[12]

Rigorous measurement of quality and safety should be an essential component of an LHS.[13,14] Further, measurement is only valuable to the extent that it is actionable and leads to improved care delivery and outcomes.[15] New and emerging measures of diagnostic safety should therefore be evaluated not only in terms of their validity but also their potential to inform pragmatic strategies to improve diagnosis.[16] At present, most of the tools and strategies that organizations use to detect patient safety concerns cannot always specifically detect diagnostic error, and even when these errors are uncovered, analysis and learning are limited.[17]

Although a coordinated national strategy to measure diagnostic safety remains an aspirational goal, recent research has yielded practical guidance for HCOs to start using measurement to enhance diagnostic safety. Measurement strategies validated in research settings are approaching readiness for implementation in an operational context. Equipped with this knowledge, HCOs have an opportunity to develop and implement strategies to learn from diagnostic safety events within their own walls.

In this Issue Brief, we discuss the state of the science of operational measurement of diagnostic safety, informed by recent peer-reviewed scientific publications, innovations in real-world healthcare settings, and initiatives to spur further development of diagnostic safety measures. Our aim is to provide knowledge and recommendations to encourage HCOs to begin to identify and learn from diagnostic errors.

Special Considerations for Measurement of Diagnostic Safety

Defining Diagnostic Errors and Diagnostic Performance

Measurement begins with a definition. The recent NASEM report defines 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."[3] This definition provides three key concepts that need to be operationalized, namely:

1. Accurately identifying the explanation (or diagnosis) of the patient's problem,
2. Providing this explanation in a timely manner, and
3. Effectively communicating the explanation.

Whereas these criteria are clear-cut in some cases (e.g., a patient diagnosed with bronchitis who is having a pulmonary embolus), they are more ambiguous in others, lacking consensus even among experienced clinicians.[18,19] Uncertainty in the diagnostic process is ubiquitous and accuracy is often poorly defined, with no widely accepted standards for how long a diagnosis should take for most conditions. Furthermore, patient presentations often evolve over time, and sometimes the best approach is to defer diagnosis or testing to a later time, or to not make a definitive diagnosis until more information is available or if symptoms persist or evolve.[20]

Diagnostic performance can be defined not only by accuracy and timeliness but also by efficiency (e.g., minimizing resource expenditure and limiting the patient's exposure to risk).[21] Measurement of diagnostic performance should hence consider the broader context of value-based care, including quality, risks, and costs, rather than focus simply on achieving the correct diagnosis in the shortest time.[16,22]

Understanding the Multifactorial Context of Diagnostic Safety

Diagnostic errors are best conceptualized as events that may be distributed over time and place and shaped by multiple contextual factors and complex dynamics involving system-related, patient-related, team-related, and individual cognitive factors. Availability of clinical data that provide a longitudinal picture across care settings is essential to understanding a patient's diagnostic journey.

A multidisciplinary framework, the Safer Dx framework[14] (Figure 1), has been proposed to help advance the measurement of diagnostic errors. The framework follows the Donabedian Structure-Process-Outcome model,[23] which approaches quality improvement in three domains:

1. Structure (characteristics of care providers, their tools and resources, and the physical/organizational setting);
2. Process (both interpersonal and technical aspects of activities that constitute healthcare); and
3. Outcome (change in the patient's health status or behavior).

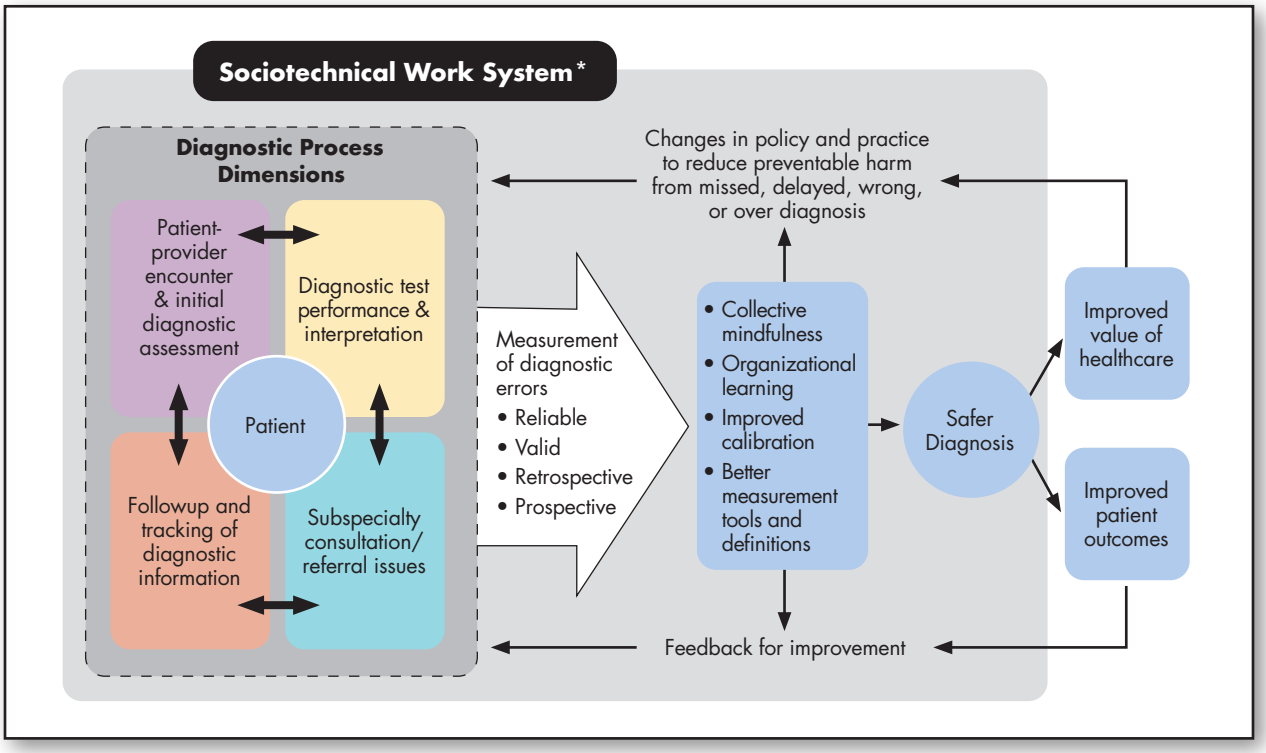
The recent NASEM report adapted concepts from the Safer Dx framework and other sources to generate a similar framework that emphasizes system-level learning and improvement.[3]

Measurement must account for all aspects of the diagnostic process as well as the distributed nature of diagnosis, i.e., evolving over time and not limited to what happens during a single clinician-patient visit. The five interactive process dimensions of diagnosis include:

1. Patient-clinician encounter (history, physical examination, ordering of tests/referrals based on assessment);
2. Performance and interpretation of diagnostic tests;
3. Followup and tracking of diagnostic information over time;
4. Subspecialty and referral-specific factors; and
5. Patient-related factors.[24]

Diagnostic performance is the outcome of these processes within a complex, adaptive sociotechnical system.[25,26] Safe diagnosis (as opposed to missed, delayed, or wrong) is an intermediate outcome compared with more distal patient and healthcare delivery outcomes.

Figure 1. Safer Dx Framework



Source: Singh H, Sittig DF. Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework. BMJ Qual Saf 2015;24(2):103-10. doi: 10.1136/bmjqs-2014-003675.

* The sociotechnical work system includes eight technological and nontechnological dimensions, as well as external factors affecting diagnostic performance and measurement, such as payment systems, legal factors, national quality measurement initiatives, accreditation, and other policy and regulatory requirements.

The Safer Dx framework is interactive and acts as a continuous feedback loop. It includes the complex adaptive sociotechnical work system in which diagnosis takes place (structure), the process dimensions in which diagnoses evolve beyond the doctor's visit (process), and resulting outcomes of "safe diagnosis" (i.e., correct and timely), as well as patient and healthcare outcomes (outcomes). Valid and reliable measurement of diagnostic errors results in collective mindfulness, organizational learning, improved collaboration, and better measurement tools and definitions. In turn, these proximal outcomes enable overall safer diagnosis, which contributes to both improved patient outcomes and value of healthcare. The knowledge created by measurement will also lead to changes in policy and practice to reduce diagnostic errors as well as feedback for improvement.

Understanding these special considerations for measurement is essential to avoid confusion with other care processes (e.g., screening, prevention, management, or treatment). It also keeps diagnostic safety as a concept distinct from other common concerns such as communication breakdowns, readmissions, and care transitions, which may be either contributors or outcomes related to diagnostic errors. The Safer Dx framework underscores that diagnostic errors can emerge across multiple episodes of care and that corresponding clinical data across the care continuum are essential to inform measurement. However, in reality these data are not always available.

Providers and HCOs, often unaware of their patients' ultimate diagnosis-related outcomes, could benefit from measurement of diagnostic performance, which can then contribute to better system-level understanding of the problem and enable improvement through feedback and learning.[27-29] Developing and implementing measurement methods that focus on what matters is consistent with LHS approaches.[30] In the absence of robust and accurate measures for diagnostic safety or any current external accountability metric, HCOs should focus on implementing measurement strategies that can lead to actionable data for improvement efforts.[31]

Choosing Data Sources for Measurement

Measurement must account for real-world clinical practice, taking into consideration not only the individual clinician's decision making and reasoning but also the influence of systems, team members, and patients on the diagnostic process. As a first step, safety professionals need access to both reliable and valid data sources, ideally across the longitudinal continuum of patient care, as well as pragmatic tools to help measure and address diagnostic error.

Measurement must account for real-world clinical practice, taking into consideration not only the individual clinician's decision making and reasoning but also the influence of systems, team members, and patients on the diagnostic process.

Many methods have been suggested to study diagnostic errors.[32] While most of them have been evaluated in research settings, few HCOs take a systematic approach to measure or monitor diagnostic error in routine clinical care.[33] However, with appropriate tools and guidance, all HCOs should be able to adopt systematic approaches to measure and learn about diagnostic safety.

Not all methods to measure diagnostic performance are feasible to implement given limited time and resources. For instance, direct video observations of encounters between clinicians and patients would be excellent, but these are expensive and difficult to sustain on the scale needed for systemwide change.[34] In contrast, a more viable alternative for most HCOs is to leverage existing data sources, such as large electronic health record (EHR) data repositories.

EHRs can be useful for measuring both discrete (and perhaps relatively uncommon) events and risks that are systematic, recurrent, or outside clinicians' awareness. For example, HCOs with the resources to query or mine electronic data repositories have options for both retrospective and prospective measurements of diagnostic safety.[35]

Even in the absence of standardized methods for measuring diagnostic error, HCOs should measure, learn from, and intervene to address threats to diagnostic safety by following basic measurement principles. For instance, the recently released Salzburg Statement on Moving Measurement Into Action was published through a collaboration between the Salzburg Global Seminar and the Institute for Healthcare Improvement. The statement recognizes the absence of standard and effective safety measures and provides eight broad principles for patient safety measurement. These principles act as "a call to action for all stakeholders in reducing harm, including policymakers, managers and senior leaders, researchers, health care professionals, and patients, families, and communities."[36]

In the same spirit, we propose several general assumptions that should guide and facilitate institutional practices for measurement of diagnostic safety and lead to actionable intelligence that identifies opportunities for improvement:

- The underlying motivation for measurement is to learn from errors and improve clinical operations, as opposed to merely responding to external incentives or disincentives.
- Efforts should focus on conditions that have been shown to be relatively common in being missed, which leads to patient harm.
- Measurement should not be used punitively to identify provider failures but rather should be used to uncover system-level problems.[37] Information should be used to inform system fixes and related organizational processes and procedures that constitute the sociotechnical context of diagnostic safety.
- Use of a single method to identify diagnostic error, such as manual chart reviews or voluntary reporting, will have limited effectiveness.[38,39] For ongoing, widespread monitoring at an organizational level, a combination of methods will have higher yield.
- Hindsight bias, i.e., when knowledge of the outcome significantly affects perception of past events, is inherent in most analysis and will need to be minimized by focusing on how missed opportunities can become learning opportunities.

With this general guidance in mind, we recommend that HCOs begin to monitor diagnostic safety using the most robust data sources currently available. Table 1 describes various data sources and strategies that could enable measurement of diagnostic error, along with their respective strengths and limitations. The following sections describe these data sources and strategies and the evidence to support their use for measurement of diagnostic safety.

Table 1. Data Sources and Related Strategies for Measurement of Diagnostic Safety

Data Source	What Can Be Learned	Approaches to Operationalizing Data Collection	Approaches to Operationalizing Data Synthesis	Strengths	Limitations	Examples
Routinely recorded quality and safety events	Awareness of the impact and harm of diagnostic safety events Emerging patterns that may suggest high-risk situations	Peer review Morbidity and mortality conferences Adverse event or incident reports Risk management sources Malpractice claims Autopsy	Content analysis of related sources of clinical data Record reviews	High level of detail Focus on events with clear potential for harm	Captured events that are not representative Unknown scope and frequency of errors	Meeks, et al., 2014[42] Cifra, et al., 2015[113] Gupta, et al., 2017[114] Shojania, et al., 2003[115]
Solicited clinician reports	Understanding of system-related and cognitive factors that affect diagnostic safety	Web-based, telephone-based, written, and in-person mechanisms to report suspected safety breakdowns Followup interviews with clinicians	Content analysis Record reviews Real-time assessment or intervention*	May capture safety breakdowns undetected by algorithmic methods, including "near-miss" and low-harm events Can capture events in real time Needs engaged clinicians and institutional support for maintenance	Captured events that are not representative Underestimates of frequency of safety events	Okafor, et al., 2015[52]
Solicited patient reports	Understanding of system-related, patient-related, and communication factors that affect diagnostic safety	Surveys Interviews Followup of patient complaints to institutions, licensing boards, and accrediting organizations Real-time reporting (e.g., inpatient hotline)	Content analysis Record review Real-time assessment or intervention*	May capture safety breakdowns undetected by algorithmic methods or manual chart review Can capture events in real time	Lack of validated measurement strategies Captured events that may not be representative Underestimates of frequency of safety events	Walton, et al., 2017[116] Fowler, et al., 2008[117] Giardina, et al., 2018[61]

Data Source	What Can Be Learned	Approaches to Operationalizing Data Collection	Approaches to Operationalizing Data Synthesis	Strengths	Limitations	Examples
Administrative billing data	Detection of diagnostic safety-related patterns and events within a cohort defined by high-risk indicators	ICD-based indicators applied to administrative datasets to select cohorts for further review (ICD = International Classification of Diseases)	Statistical analysis	Data that are routinely collected and widely accessible	Validity for case finding not established Variable data quality Lack of detailed and contextual clinical data in administrative datasets	Liberman, Newman-Toker, 2018[64] Mahajan, et al., 2020[66]
Medical records (random, selective, and e-trigger enhanced)	Detection of diagnostic safety events within a sample of screened or randomly selected records Insight into vulnerable structures and processes	Manual review by one or more trained raters Selection of high-risk cohorts (such as patients with cancer or epidural abscess) Use of electronic triggers to flag charts for manual review	Record review (e.g., using structured review tool such as the Safer Dx instrument) Real-time assessment or intervention*	High level of detail regarding the course of clinical care Acceptable agreement for case finding by trained reviewers Intervention and investigation with minimal recall bias possible due to detection of safety concerns in real time	Inefficient (more efficient with selected high-risk cohorts or use of e-triggers) Variable quality of clinical documentation Limited to data systems with large repositories of data that can be queried Dedicated monitoring resources needed for real-time application	Bhise, et al., 2017[70] Murphy, et al., 2014[78] Singh, et al., 2012[19]
Advanced data science methods, including natural language processing of EHR data*	Detection of possible diagnostic safety events (retrospectively or in real time) based on machine-learning algorithms	Application of algorithms to clinical documentation or other unstructured EHR data to search for language indicating possible safety concerns	Record review Real-time assessment or intervention*	Can capture data that may be missed by administrative data-based screening May enhance efficiency of record review	Validity for case finding unknown Limited to data systems with large repositories of data that can be queried Dedicated monitoring resources needed for real-time application	None known

Note: Approaches marked with an asterisk (*) have potential for future application but are not currently developed for real-time surveillance.

Learning From Known Incidents and Reports

No single data source will capture the full range of diagnostic safety concerns. Valuable information can be gleaned from even limited data sources so long as those who use the data remain mindful of its limitations for a given purpose.[40] For instance, many routinely recorded discrete events lend themselves to retrospective analysis of diagnostic safety. Most healthcare organizations have incident reporting systems, although reporting has included few diagnostic events.[41]

There is also an opportunity to leverage peer review programs to improve diagnostic self-assessment, feedback, and improvement.[42] Similarly, autopsy reports,[43] diagnostic discrepancies at admission versus discharge,[44,45] escalations of care,[46,47] and malpractice claims[48-51] may be reviewed with special attention to opportunities to improve diagnosis. These data sources may not shed light on the frequency or scope of a problem, but they can help raise awareness of the impact and harm of diagnostic errors and, in some cases, specific opportunities for improvement.

Voluntary reports solicited specifically from clinicians who make diagnoses are another potentially useful source of data on diagnostic safety. For example, reports from clinicians who have witnessed diagnostic error have the advantage of rich detail that, at least in some cases, may offer insight into ways to prevent or mitigate future errors. However, no standardized mechanisms exist to report diagnostic errors. Despite widespread efforts to enable providers to report errors,[17,52] clinicians find reporting tools onerous and are often unaware of errors they make.[53]

It has also become clear that a local champion and quality improvement team support are needed to sustain reporting behavior. At present, few facilities or organizations allocate "protected time" that is essential for clinicians to report, analyze, and learn from safety events. Some of these challenges could be overcome by having frontline clinicians report events briefly and allowing organizational safety teams (which include other clinicians) to analyze them. Still, voluntary reporting alone cannot address the multitude of complex diagnostic safety concerns, and reporting can only be one aspect of a comprehensive measurement strategy.[54]

Patients are often underexplored sources of information, and many organizations already conduct patient surveys, interviews, and reviews of patient complaints to learn about safety risks.[55-58] Prior work in other areas of patient safety (e.g., medication errors, infection control) has examined the potential of engaging patients proactively to monitor safety risks and problems.[59,60] With subsequent development, similar mechanisms could be used to monitor diagnosis-related issues.

One limitation of current patient reporting systems is the lack of validated patient-reported questions or methods to detect diagnostic safety concerns. Barriers to patient engagement in safety initiatives, including low health literacy, lack of wider acceptance of safety monitoring as part of the patient role, provider expectations and attitudes, and communication differences, will also need to be addressed to make the most of such efforts.[59,61,62] Real-time adverse event and "near-miss" reporting systems, akin to those intended for use by clinicians, are another potential mechanism to collect patient-reported data on diagnostic safety.[63]

Learning From Existing Large Datasets

Whereas direct reports from patients and clinicians may offer unique insights, HCOs may also be ready to use large datasets to uncover trends in diagnostic performance and events that are not otherwise easily identified. Administrative and billing data are widely available in most modern HCOs and have been proposed as one source of data for detecting missed opportunities for accurate and timely diagnosis.[64-66]

For example, diagnosis codes assigned at successive clinical encounters may be used as a proxy for the evolution of a clinical diagnosis; if significant discrepancies are found, it may lead to a search for reasons.[44] Using symptom-disease-based dyads, such as abdominal pain followed by appendicitis a few days later or dizziness followed by stroke, are examples of this approach.[66,67]

This strategy may be intuitive to patient safety leaders because several other safety measurement methods are also based on diagnosis codes extracted from administrative datasets (e.g., patient safety indicators).[68,69] However, unlike specific safety events that can be coded with good sensitivity (e.g., retention of foreign objects during procedures), administrative data are not sufficiently sensitive to detect diagnostic errors. Moreover, administrative data lack relevant clinical details about diagnostic processes that can be improved. Administrative data sources are mainly useful insofar as they can be used to identify patterns or specific cohorts of patients to further review for presence of diagnostic error, based on diagnosis or other relevant characteristics that may be considered risk factors or of special interest for diagnostic safety improvement.

Medical records can be a rich source of data, as they contain clinical details and reflect the patient's longitudinal care journey. Although medical record reviews are considered valuable and sometimes even gold standard for detecting diagnostic errors, it is often not clear which records to review. Reviewing records at random can be burdensome and resource intensive. However, more selective methods can identify a high-yield subset (e.g., reviewing records of patients diagnosed with specific clinical conditions at high risk of being missed, such as colorectal cancer or spinal epidural abscess).[70,71] Such selective methods can be more efficient compared with voluntary reporting or nonselective or random manual review.

Another way to select records is through the "trigger" approach, which aims to "alert patient safety personnel to possible adverse events so they can review the medical record to determine if an actual or potential adverse event has occurred."[72-75] EHRs and clinical data warehouses make it possible to identify signals suggestive of missed diagnosis prior to detailed reviews. HCOs can search EHRs on a scale that would be untenable using manual or random search methods using electronic algorithms, or "e-triggers," which mine vast amounts of clinical and administrative data to identify these signals.[74,76,77]

For example, algorithms could identify patients with a certain type of abnormal test result (denominator) and identify which results have still not been acted upon after a certain length of time (numerator).[78] This type of algorithm is possible because the data about an abnormal result (e.g., abnormal hemoglobin value and microcytic anemia) and the followup action needed (e.g., colonoscopy for the 65-year-old) are (or should be) coded in the EHR. Similarly, certain patterns, such as unexpected hospitalizations after a primary care visit, can be identified more accurately if corresponding clinical data are available.[19]

To enhance the yield of record reviews, e-triggers can be developed to alert personnel to potential patient safety events and enable targeted review of high-risk patients.[35] The Institute for Healthcare Improvement's Global Trigger Tools,[79] which include both manual and electronic trigger tools to detect inpatient events,[80-82] are widely used but were not specifically designed to detect diagnostic errors. More targeted e-triggers for diagnostic safety measurement and monitoring can be developed and integrated within existing patient safety surveillance systems in the future and may enhance the yield of record reviews.[46,83-86] Other examples of possible e-triggers that are either in development or approaching wider testing and implementation are described in Table 2.

Table 2. Examples of Potential Safer Dx e-Triggers Mapped to Diagnostic Process Dimensions of the Safer Dx Framework[14] (adapted from Murphy DR, et al.[35])

Safer Dx Diagnostic Process	Safer Dx Trigger Example	Potential Diagnostic Error
Patient-provider encounter	Emergency department or primary care visit followed by unplanned hospitalization	Missed red flag findings or incorrect diagnosis during initial office visit
	ED/PC visit within 72 hours after ED or hospital discharge	Missed red flag findings during initial ED/PC or hospital visit
	Unexpected transfer from hospital general floor to ICU within 24 hours of ED admission	Missed red flag findings during admission
Performance and interpretation of diagnostic tests	Amended imaging report	Missed findings on initial read or lack of communication of amended findings
Followup and tracking of diagnostic information	Abnormal test result with no timely followup action	Abnormal test result missed
Referral-related factors	Urgent specialty referral followed by discontinued referral within 7 days	Delay in diagnosis from lack of specialty expertise
Patient-related factors	Poor rating on patient experience scores post ED/PC visit	Patient report of communication barriers related to missed diagnosis

Both selective and e-trigger enhanced reviews are advantageous because they can allow development of potential measures for wider testing and adoption. For instance, measures of diagnostic test result followup may focus specifically on abnormal test results that are suggestive of serious or time-sensitive diagnoses, such as cancer.[87] Examples of diagnostic safety measures could include the proportion of documented "red-flag" symptoms or test results that receive timely followup, or the proportion of patients with cancer newly diagnosed within 60 days of first presentation of known red flags.[16] Depending on an HCO's priorities, safety leaders could consider additional development, testing, and potential implementation of measure concepts proposed by NQF and other researchers.[8,16,88]

Other mechanisms for mining clinical data repositories, such as natural language processing (NLP) algorithms,[89] are too early in development but may be a useful complement to future e-triggers for detection of diagnostic safety events. Whereas e-triggers leverage structured data to identify possible safety concerns, a substantial proportion of data in EHRs is unstructured and therefore inaccessible to e-triggers. NLP and machine-learning techniques could help analyze and interpret large volumes of unstructured textual data, such as those in clinical narratives and free-text fields, and reduce the burden of medical record reviews by selecting records with potentially highest yield. While research has examined the predictive validity of NLP algorithms for detection of safety incidents and adverse events,[90-92] to date this methodology has not been applied to or validated for measurement of diagnostic safety.

Synthesizing Data and Enhancing Confidence in Measurement

Determining the presence of diagnostic error is complex and requires additional evaluation for missed opportunities.[93] A binary classification (e.g., error or no error) may be insufficient for cases involving greater uncertainty, which call for more graded assessment approaches reflecting varying degrees of confidence in the determination of error.[20,94]

Depending on the measurement method, a thorough content review may be sufficient to identify missed opportunities for diagnosis, contributing factors, and harm or impact. However, systematic approaches to surveillance and measurement of diagnostic safety often warrant the use of structured data collection instruments[95,96] that assess diagnostic errors using objective criteria, as well as taxonomies to classify process breakdowns. For example, the Revised Safer Dx Instrument[97] is a validated tool that can be used to do an initial screen for presence or absence of diagnostic error in a case. It helps users identify potential diagnostic errors in a standardized way for further analysis and safety improvement efforts.

Structured assessment instruments can also be used to provide clinicians with data for feedback and reflection that they otherwise may not receive.[98] Furthermore, process breakdowns can be analyzed using approaches such as the Diagnostic Error Evaluation and Research taxonomy[99] or the Safer Dx Process breakdown taxonomy,[97] both of which help to identify where in the diagnostic process a problem occurred. Other factors related to diagnostic error, such as the presence of patient harm (e.g., clear evidence of harm versus "near-misses"), preventability, and actionability, may also be important to define in advance so that the selected measurement strategy aligns with the learning and improvement goals. Nevertheless, the science of understanding the complex mix of cognitive and sociotechnical contributory factors and implementing effective solutions based on these data still needs additional development.[100-102]

Getting Ready for Measurement: Overcoming Barriers and Taking Next Steps

HCOs are resource constrained, and efforts to measure quality and safety of healthcare are anchored foremost to the measures specifically required by accrediting agencies and payers. Currently, measures of diagnostic performance and safety are not among these required measures, and the burden of additional data collection is a daunting prospect for HCOs that already struggle to meet existing requirements. Furthermore, we need national consensus regarding what aspects of diagnostic safety can be measured pragmatically in real-world care settings.

The current lack of incentives and models to measure diagnostic safety can make it difficult to get started even in the best case scenario when leadership support and resources are available to support improvement efforts focused on diagnosis. Nevertheless, as the burden of diagnostic errors is increasingly recognized and as measurement strategies become better defined, diagnostic safety will no longer remain sidelined in the landscape of healthcare quality and safety.

Despite the additional burden that measurement for improvement would entail, the benefit of these activities for improving patient safety and stimulating learning and feedback could outweigh concerns.

Despite the additional burden that measurement for improvement would entail, the benefit of these activities for improving patient safety and stimulating learning and feedback could outweigh concerns. In fact, some HCOs have already started on their journeys, including one that has pursued measurement on multiple fronts[103] and another that

aims to pursue measurement as part of becoming a "Learning and Exploration of Diagnostic Excellence (LEDE)" organization.[33] Now is the time for other HCOs to use the data already available to them to begin to detect, understand, and learn from diagnostic errors.

While diagnostic errors occur across the spectrum of medical practice, measurement should be strategic and focused on areas with strong potential for learning and impact.

While diagnostic errors occur across the spectrum of medical practice, measurement should be strategic and focused on areas with strong potential for learning and impact. As with other aspects of patient safety measurement, goals should include:

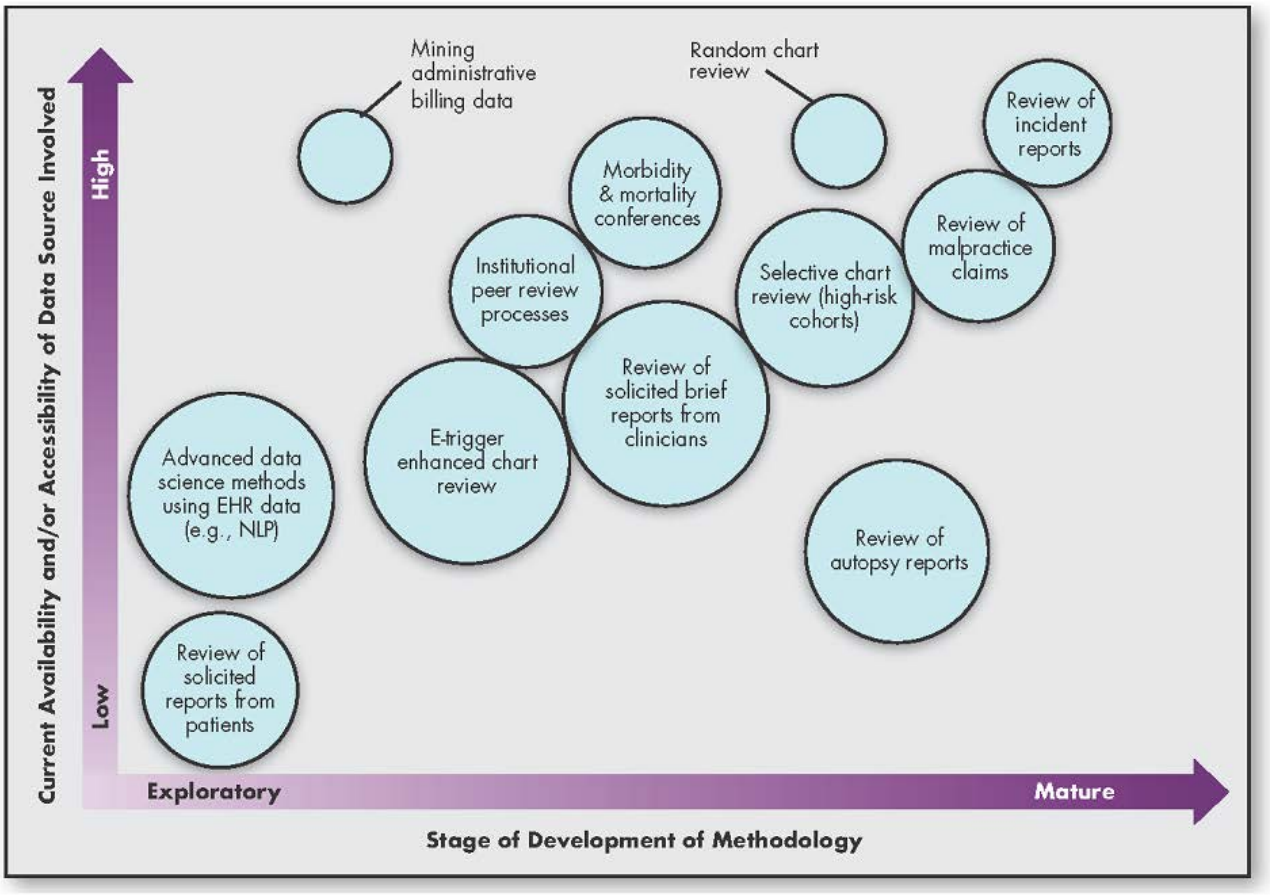
- Creating learning opportunities from past events with both potential and real harm,
- Ensuring reliability in diagnostic safety,
- Anticipating and preparing for problems related to diagnosis, and
- Integrating and learning from the knowledge generated.[13]

Absent a specific local or institutional need or mandate calling for a specific measurement target, we recommend that HCOs new to measurement of diagnostic safety focus initial measurement efforts on a limited set of realistic measures that map to one or a few specific diagnoses or care processes. Measurement targets may be in line with:

- Published literature (e.g., research showing a high rate of missed test results),
- Priorities identified in national initiatives (e.g., timely diagnosis of cancer), or
- Local needs identified by quality and safety committees (e.g., focusing on specific symptoms such as abdominal pain[104] or specific diseases such as spinal epidural abscess).[70]

Once a target is identified, HCOs should use measurement strategies that balance validity (i.e., for case finding) with yield. Figure 2 visualizes the implementation readiness of several diagnostic safety measurement strategies discussed in this brief. For example, e-trigger enhanced structured chart review appears suitable for operational measurement with additional development. Systems without EHR capabilities can use other data sources (e.g., selective reviews, event reports) to begin measurement as soon as possible, rather than wait for the information technology infrastructure needed to implement EHR-based algorithms. The strategies depicted in Figure 2 represent the current state of the science and are subject to change as new innovations are validated.

Figure 2. Implementation Readiness of Diagnostic Safety Measurement Strategies



Note: Larger circles denote higher potential yield for cases that can inform systemwide learning and improvement. Measurement strategies that are ready for implementation balance validity and yield (i.e., an estimate of the proportion of cases with diagnostic errors that could lead to learning and improvement relative to measurement effort). The relative position of these methods will vary according to local context; Figure 2 illustrates data sources for a typical healthcare organization equipped with an EHR and moderate resources for quality and safety improvement.

Table 3 presents a summary of the strategies and the factors affecting their use.

Table 3. Implementation Readiness of Diagnostic Safety Measurement Strategies and Estimated Yield Relative to Effort

Measurement Strategy	Stage of Development	Current Potential Availability and/or Accessibility of Data Source	Estimated Yield Relative to Effort
Review of solicited reports from patients	Exploratory	Low	Medium
Advanced data science methods using EHR data (e.g., NLP)	Exploratory	Low	Very large
Mining administrative billing data	Exploratory	High	Very small
E-trigger enhanced chart review	Moderate	Moderate	Very large
Institutional peer review processes	Moderate	High	Medium
Morbidity and mortality conferences	Moderate	High	Medium
Review of solicited brief reports from clinicians	Moderate	Moderate	Very large
Selective chart review of high-risk cohorts	Mature	High	Large
Random chart review	Mature	High	Very small
Review of autopsy reports	Mature	Low	Large
Review of malpractice claims	Mature	High	Medium
Review of incident reports	Mature	High	Small

Based on what HCOs are learning, they could conduct additional activities related to quality improvement, such as:

- Better managing test results to make sure they are acted on,[105]
- Conducting a self-assessment of reliability of communication and reporting of test results,[106]
- Closing referral loops,[107] and
- Enhancing teamwork and communication with patients.[108]

Meanwhile, research should stimulate the development of more rigorous measures that are better correlated with not just more timely and accurate diagnosis but also less preventable diagnostic harm.

For additional implementation and adoption, especially beyond measurement to inform institutional quality improvement activities that we discuss in this brief, additional balance measures will need to be developed. For example, measures that focus on underdiagnosis of a certain condition (stroke) may lead to overtesting for that condition (magnetic resonance imaging on all patients with dizziness in primary or emergent care). Thus, research efforts must continue to inform and strengthen the rigor of measurement.

Other unintended consequences, including gaming and measure fatigue, may emerge from implementing suboptimal and inadequately validated performance measures for public reporting, performance incentives, or penalties. For measures to be used for accountability, they will first need to be tested and validated and linked to improving suboptimal care or outcomes.[109] This document and its recommendations thus focus largely on measurement for improvement at the HCO level and not measurement for accountability purposes.

In the long term, measures will need to be developed to address diagnostic excellence, defined in terms of the ability to make a diagnosis using the fewest resources while maximizing patient experiences, managing and communicating uncertainty to patients, and tolerating watchful waiting when unfocused treatment may be harmful.[110] As measurement methods evolve, they should address concepts such as uncertainty and clinician calibration (the degree to which clinicians' confidence in the accuracy of their diagnostic decision making aligns with their actual accuracy[111]). Understanding and measuring these concepts in the diagnostic safety equation is essential to optimize the balance between reducing overuse and addressing underuse of diagnostic tests and other resources.[112]

Conclusion

While measurement of diagnostic error has been challenging, recent research and early implementation in operational settings have provided new evidence that can now stimulate progress. Whether by leveraging current resources or building capacity for additional data gathering, HCOs have a variety of options to begin measurement to reduce preventable diagnostic harm. The measurements can be implemented progressively and, eventually, can be supported by payers and accrediting organizations. This overview is a call to action for leaders within both large and small HCOs in any setting to assess how they could begin their journey to measure and reduce preventable diagnostic harm and then to act on that assessment.

References

1. Singh H, Graber ML. Improving diagnosis in health care - the next imperative for patient safety. N Engl J Med. 2015;373(26):2493-5. https://www.ncbi.nlm.nih.gov/pubmed/26559457. Accessed April 1, 2020.
2. Saber Tehrani AS, Lee H, Mathews SC, Shore A, Makary MA, Pronovost PJ, et al. 25-year summary of U.S. malpractice claims for diagnostic errors 1986-2010: an analysis from the National Practitioner Data Bank. BMJ Qual Saf. 2013;22(8):672-80. https://www.ncbi.nlm.nih.gov/pubmed/23610443. Accessed April 1, 2020.
3. National Academies of Sciences, Engineering, and Medicine. Improving Diagnosis in Health Care. http://iom.nationalacademies.org/Reports/2015/Improving-Diagnosis-in-Healthcare.aspx. Accessed March 31, 2020.
4. National Quality Forum. National Quality Forum home page. http://www.qualityforum.org/Home.aspx. Accessed March 31, 2020.
5. National Quality Forum. Measures, Reports & Tools. http://www.qualityforum.org/Measures_Reports_Tools.aspx. Accessed March 31, 2020.
6. Graber ML, Wachter RM, Cassel CK. Bringing diagnosis into the quality and safety equations. JAMA. 2012;308(12):1211-2. https://www.ncbi.nlm.nih.gov/pubmed/23011708. Accessed April 1, 2020.
7. Gordon and Betty Moore Foundation. New projects aim to develop clinical quality measures to improve diagnosis. https://www.moore.org/article-detail?newsUrlName=new-projects-aim-to-develop-clinical-quality-measures-to-improve-diagnosis. Accessed March 31, 2020.
8. National Quality Forum. Reducing Diagnostic Error: Measurement Considerations. October 2019. https://www.qualityforum.org/ProjectDescription.aspx?projectID=90704. Accessed March 31, 2020.
9. Agency for Healthcare Research and Quality. Diagnostic Safety and Quality. https://www.ahrq.gov/topics/diagnostic-safety-and-quality.html. Accessed March 31, 2020.
10. Centers for Medicare & Medicaid Services. TEP Current Panels. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/MMS/TEP-Current-Panel#a0214-1. Accessed March 31, 2020.
11. Health Research & Educational Trust. Improving Diagnosis in Medicine. Diagnostic Error Change Package. https://www.improvediagnosis.org/wp-content/uploads/2018/11/improving-diagnosis-in-medicine-change-package-11-8.pdf. Accessed April 1, 2020.
12. Olsen LA, Aisner D, McGinnis JM, eds. Institute of Medicine (US) Roundtable on Evidence-Based Medicine. The Learning Healthcare System: Workshop Summary. The National Academies Collection: Reports funded by National Institutes of Health. Washington, DC: National Academies Press; 2007. https://www.ncbi.nlm.nih.gov/pubmed/21452449. Accessed April 1, 2020.
13. Vincent C, Burnett S, Carthey J. Safety measurement and monitoring in healthcare: a framework to guide clinical teams and healthcare organisations in maintaining safety. BMJ Qual Saf. 2014;23(8):670-7. https://www.ncbi.nlm.nih.gov/pubmed/24764136. Accessed April 1, 2020.

14. Singh H, Sittig DF. Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework. BMJ Qual Saf. 2015;24(2):103-10. https://www.ncbi.nlm.nih.gov/pubmed/25589094. Accessed April 1, 2020.

15. McGlynn EA, McDonald KM, Cassel CK. Measurement is essential for improving diagnosis and reducing diagnostic error: a report from the Institute of Medicine. JAMA. 2015;314(23):2501-2. https://www.ncbi.nlm.nih.gov/pubmed/26571126. Accessed April 1, 2020.

16. Singh H, Graber ML, Hofer TP. Measures to improve diagnostic safety in clinical practice. J Patient Saf. 2019;15(4):311-6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5398940/. Accessed April 1, 2020.

17. Graber ML, Trowbridge RL, Myers JS, Umscheid CA, Strull W, Kanter MH. The next organizational challenge: finding and addressing diagnostic error. Jt Comm J Qual Patient Saf. 2014;40(3):102-10. https://www.ncbi.nlm.nih.gov/pubmed/24730205. Accessed April 1, 2020.

18. Zwaan L, de Bruijne M, Wagner C, Thijs A, Smits M, van der Wal G, Timmermans DR. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. Arch Intern Med. 2010;170(12):1015-21. https://www.ncbi.nlm.nih.gov/pubmed/20585065. Accessed April 1, 2020.

19. Singh H, Giardina T, Forjuoh S, Reis M, Kosmach S, Khan M, Thomas EJ. Electronic health record-based surveillance of diagnostic errors in primary care. BMJ Qual Saf. 2012;21(2):93-100. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3680372/. Accessed April 1, 2020.

20. Zwaan L, Singh H. The challenges in defining and measuring diagnostic error. Diagnosis. 2015;2(2):97-103. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4779119/. Accessed April 1, 2020.

21. Singh H. Diagnostic errors: moving beyond 'no respect' and getting ready for prime time. BMJ Qual Saf. 2013;22(10):789-92. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3786612/. Accessed April 1, 2020.

22. Hofer TP, Kerr EA, Hayward RA. What is an error? Eff Clin Pract 2000;3:261-9. PMID:11151522.

23. Donabedian A. The quality of care. How can it be assessed? JAMA. 1988;260(12):1743-8. https://www.ncbi.nlm.nih.gov/pubmed/3045356. Accessed April 1, 2020.

24. Singh H, Giardina TD, Meyer AN, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings. JAMA Intern Med. 2013;173(6):418-25. https://www.ncbi.nlm.nih.gov/pubmed/23440149. Accessed April 1, 2020.

25. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care. 2010;19(Suppl 3):i68-i74. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3120130/. Accessed April 1, 2020.

26. Henriksen K, Brady J. The pursuit of better diagnostic performance: a human factors perspective. BMJ Qual Saf. 2013;22(Suppl 2):ii1-ii5. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3786636/. Accessed April 1, 2020.

27. Dhaliwal G. Web Exclusives. Annals for hospitalists inpatient notes - diagnostic excellence starts with an incessant watch. Ann Intern Med. 2017;167(8):HO2-HO3. https://www.ncbi.nlm.nih.gov/pubmed/29049776. Accessed April 1, 2020.

28. Lane KP, Chia C, Lessing JN, Limes J, Mathews B, Schaefer J, Seltz LB, Turner G, Wheeler B, Wooldridge D, Olson AP. Improving resident feedback on diagnostic reasoning after handovers: the LOOP Project. J Hosp Med. 2019;14(10):622-5. https://www.ncbi.nlm.nih.gov/pubmed/31433779. Accessed April 1, 2020.

29. Shenvi EC, Feupe SF, Yang H, El-Kareh R. "Closing the loop": a mixed-methods study about resident learning from outcome feedback after patient handoffs. Diagnosis (Berl). 2018;5(4):235-42. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6556111/. Accessed April 1, 2020.

30. Thomas EJ, Classen DC. Patient safety: let's measure what matters. Ann Intern Med. 2014;160(9):642-3. https://www.ncbi.nlm.nih.gov/pubmed/24798525. Accessed April 1, 2020.

31. Solberg LI, Mosser G, McDonald S. The three faces of performance measurement: improvement, accountability, and research. Jt Comm J Qual Improv. 1997;23(3):135-47. https://www.ncbi.nlm.nih.gov/pubmed/9103968. Accessed April 1, 2020.

32. Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf. 2013;22 Suppl2:ii21-ii7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3786666/. Accessed April 1, 2020.

33. Singh H, Upadhyay DK, Torretti D. Developing health care organizations that pursue learning and exploration of diagnostic excellence: an action plan. Acad Med. 2019 Oct 29. https://www.ncbi.nlm.nih.gov/pubmed/31688035. Accessed April 1, 2020.

34. Amelung D, Whitaker KL, Lennard D, Ogden M, Sheringham J, Zhou Y, Walter FM, Singh H, Vincent C, Black G. Influence of doctor-patient conversations on behaviours of patients presenting to primary care with new or persistent symptoms: a video observation study. BMJ Qual Saf. 2019. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7057803/. Accessed April 1, 2020.

35. Murphy DR, Meyer AN, Sittig DF, Meeks DW, Thomas EJ, Singh H. Application of electronic trigger tools to identify targets for improving diagnostic safety. BMJ Qual Saf. 2019;28(2):151-9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6365920/. Accessed April 1, 2020.

36. Institute for Healthcare Improvement and Salzburg Global Seminar. The Salzburg Statement on Moving Measurement Into Action: Global Principles for Measuring Patient Safety. https://www.salzburgglobal.org/fileadmin/user_upload/Documents/2010-2019/2019/Session_622/SalzburgGlobal_Statement_622_Patient_Safety_01.pdf. Accessed March 31, 2020.

37. Smith MW, Davis GT, Murphy DR, Laxmisan A, Singh H. Resilient actions in the diagnostic process and system performance. BMJ Qual Saf. 2013;22(12):1006-13. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3833904/. Accessed April 1, 2020.

38. Murff HJ, Patel VL, Hripcsak G, Bates DW. Detecting adverse events for patient safety research: a review of current methodologies. J Biomed Inform. 2003;36(1-2):131-43. https://www.ncbi.nlm.nih.gov/pubmed/14552854. Accessed April 1, 2020.

39. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med. 2003;18(1):61-7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1494808/. Accessed April 1, 2020.

40. Rosen AK. Are we getting better at measuring patient safety? Perspectives on Safety 2010 Nov 1. https://psnet.ahrq.gov/perspective/are-we-getting-better-measuring-patient-safety. Accessed April 1, 2020.

41. Levtzion-Korach O, Frankel A, Alcalai H, Keohane C, Orav J, Graydon-Baker E, Barnes J, Gordon K, Puopulo AL, Tomov EI, Sato L, Bates DW. 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):402-10. https://www.ncbi.nlm.nih.gov/pubmed/20873673. Accessed April 1, 2020.

42. Meeks DW, Meyer AN, Rose B, Walker YN, Singh H. Exploring new avenues to assess the sharp end of patient safety: an analysis of nationally aggregated peer review data. BMJ Qual Saf. 2014;23(12):1023-30. https://www.ncbi.nlm.nih.gov/pubmed/25260781. Accessed April 1, 2020.

43. Tejerina EE, Padilla R, Abril E, Frutos-Vivar F, Ballen A, Rodriguez-Barbero JM, Lorente JÁ, Esteban A. Autopsy-detected diagnostic errors over time in the intensive care unit. Hum Pathol. 2018;76:85-90. https://www.ncbi.nlm.nih.gov/pubmed/29530753. Accessed April 1, 2020.

44. Hautz WE, Kammer JE, Hautz SC, Sauter TC, Zwaan L, Exadaktylos AK, Birrenbach T, Maier V, Müller M, Schauber SK. Diagnostic error increases mortality and length of hospital stay in patients presenting through the emergency room. Scand J Trauma Resusc Emerg Med. 2019;27(1):54. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6505221/. Accessed April 1, 2020.

45. Gov-Ari E, Leann Hopewell B. Correlation between pre-operative diagnosis and post-operative pathology reading in pediatric neck masses--a review of 281 cases. Int J Pediatr Otorhinolaryngol. 2015;79(1):2-7. https://www.ncbi.nlm.nih.gov/pubmed/25479698. Accessed April 1, 2020.

46. Bhise V, Sittig DF, Vaghani V, Wei L, Baldwin J, Singh H. An electronic trigger based on care escalation to identify preventable adverse events in hospitalised patients. BMJ Qual Saf. 2018;27(3):241-6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5867429/. Accessed April 1, 2020.

47. Davalos MC, Samuels K, Meyer AN, Thammasitboon S, Sur M, Roy K, Al-Mutairi A, Singh H. Finding diagnostic errors in children admitted to the PICU. Pediatr Crit Care Med. 2017;18 (3):265-71. https://www.ncbi.nlm.nih.gov/pubmed/28125548. Accessed April 1, 2020.

48. Schiff GD, Puopolo AL, Huben-Kearney A, Yu W, Keohane C, McDonough P, et al. Primary care closed claims experience of Massachusetts malpractice insurers. JAMA Intern Med. 2013;173(22):2063-8. https://www.ncbi.nlm.nih.gov/pubmed/24081145. Accessed April 1, 2020.

49. Gandhi TK, Kachalia A, Thomas EJ, Puopolo AL, Yoon C, Brennan TA, Studdert DM. Missed and delayed diagnoses in the ambulatory setting: a study of closed malpractice claims. Ann Intern Med. 2006;145(7):488-96. https://www.ncbi.nlm.nih.gov/pubmed/17015866. Accessed April 2, 2020.

50. Kachalia A, Gandhi TK, Puopolo AL, Yoon C, Thomas EJ, Griffey R, Brennan TA, Studdert DM. Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. Ann Emerg Med. 2007;49(2):196-205. https://www.ncbi.nlm.nih.gov/pubmed/16997424. Accessed April 2, 2020.

51. Singh H, Thomas EJ, Petersen LA, Studdert DM. Medical errors involving trainees: a study of closed malpractice claims from 5 insurers. Arch Intern Med. 2007;167(19):2030-6. https://www.ncbi.nlm.nih.gov/pubmed/17954795. Accessed April 2, 2020.

52. Okafor N, Payne VL, Chathampally Y, Miller S, Doshi P, Singh H. Using voluntary reports from physicians to learn from diagnostic errors in emergency medicine. Emerg Med J. 2015. https://www.ncbi.nlm.nih.gov/pubmed/26531860. Accessed April 2, 2020.

53. Schiff GD. Minimizing diagnostic error: the importance of follow-up and feedback. Am J Med. 2008;121(5, Supplement 1):S38-S42.

54. Shojania K. The elephant of patient safety: what you see depends on how you look. Jt Comm J Qual Patient Saf. 2010;36(9):399-401. doi: 10.1016/s1553-7250(10)36058-2.

55. Ward JK, Armitage G. Can patients report patient safety incidents in a hospital setting? A systematic review. BMJ Qual Saf. 2012;21(8):685-99. https://www.ncbi.nlm.nih.gov/pubmed/22562875. Accessed April 2, 2020.

56. Scott J, Heavey E, Waring J, De Brun A, Dawson P. Implementing a survey for patients to provide safety experience feedback following a care transition: a feasibility study. BMC Health Serv Res. 2019;19(1):613. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6716906/. Accessed April 2, 2020.

57. Haroutunian P, Alsabri M, Kerdiles FJ, Adel Ahmed Abdullah H, Bellou A. Analysis of factors and medical errors involved in patient complaints in a European emergency department. Adv J Emerg Med. 2018;2(1):e4. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6548105/. Accessed April 2, 2020.

58. Gillespie A, Reader TW. Patient-centered insights: using health care complaints to reveal hot spots and blind spots in quality and safety. Milbank Q. 2018;96(3):530-67. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6131356/. Accessed April 2, 2020.

59. Longtin Y, Sax H, Leape LL, Sheridan SE, Donaldson L, Pittet D. Patient participation: current knowledge and applicability to patient safety. Mayo Clin Proc. 2010;85(1):53-62. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2800278/. Accessed April 2, 2020.

60. Weingart SN, Toth M, Eneman J, Aronson MD, Sands DZ, Ship AN, Davis RB, Phillips RS. Lessons from a patient partnership intervention to prevent adverse drug events. Int J Qual Health Care. 2004;16(6):499-507. https://www.ncbi.nlm.nih.gov/pubmed/15557360. Accessed April 2, 2020.

61. Giardina TD, Haskell H, Menon S, Hallisy J, Southwick FS, Sarkar U, Royse KE, Singh H. Learning from patients' experiences related to diagnostic errors is essential for progress in patient safety. Health Aff (Millwood). 2018;37(11):1821-7. https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2018.0698?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed. Accessed April 2, 2020.

62. Smith K, Baker K, Wesley D, Zipperer L, Clark MD. Guide to Improving Patient Safety in Primary Care Settings by Engaging Patients and Families: Environmental Scan Report. (Prepared by: MedStar Health Research Institute under Contract No. HHSP233201500022I/HHSP23337002T.). Rockville, MD: Agency for Healthcare Research and Quality. AHRQ Publication No. 17-0021-2-EF. https://www.ahrq.gov/patient-safety/reports/engage/envscan.html. Accessed April 2, 2020.

63. Huerta TR, Walker C, Murray KR, Hefner JL, McAlearney AS, Moffatt-Bruce S. Patient safety errors: leveraging health information technology to facilitate patient reporting. J Healthc Qual. 2016;38(1):17-23. https://www.ncbi.nlm.nih.gov/pubmed/26730805. Accessed April 2, 2020.

64. Liberman AL, Newman-Toker DE. Symptom-Disease Pair Analysis of Diagnostic Error (SPADE): a conceptual framework and methodological approach for unearthing misdiagnosis-related harms using big data. BMJ Qual Saf. 2018;27(7):557-66. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6049698/. Accessed April 2, 2020.

65. Michelson KA, Buchhalter LC, Bachur RG, Mahajan P, Monuteaux MC, Finkelstein JA. Accuracy of automated identification of delayed diagnosis of pediatric appendicitis and sepsis in the ED. Emerg Med J. 2019;36(12):736-40. https://www.ncbi.nlm.nih.gov/pubmed/31597671. Accessed April 2, 2020.

66. Mahajan P, Basu T, Pai CW, Singh H, Petersen N, Bellolio MF, Gadepalli SK, Kamdar NS. Factors associated with potentially missed diagnosis of appendicitis in the emergency department. JAMA Netw Open. 2020;3(3):e200612. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7063499/. Accessed April 2, 2020.

67. Newman-Toker DE, Moy E, Valente E, Coffey R, Hines AL. Missed diagnosis of stroke in the emergency department: a cross-sectional analysis of a large population-based sample. Diagnosis. 2014;1(2):155-66. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5361750/. Accessed April 2, 2020.

68. Southern DA, Burnand B, Droesler SE, Flemons W, Forster AJ, Gurevich Y, Harrison J, Quan H, Pincus HA, Romano PS, Sundararajan V, Kostanjsek N, Ghali WA. Deriving ICD-10 codes for Patient Safety Indicators for large-scale surveillance using administrative hospital data. Med Care. 2017;55(3):252-60. https://www.ncbi.nlm.nih.gov/pubmed/27635599. Accessed April 2, 2020.

69. Miller MR, Elixhauser A, Zhan C, Meyer GS. Patient Safety Indicators: using administrative data to identify potential patient safety concerns. Health Serv Res. 2001;36(6 Pt 2):110-32. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1383610/. Accessed April 2, 2020.

70. Bhise V, Meyer AND, Singh H, Wei L, Russo E, Al-Mutairi A, Murphy DR. Errors in diagnosis of spinal epidural abscesses in the era of electronic health records. Am J Med. 2017;130(8):975-81. https://www.ncbi.nlm.nih.gov/pubmed/28366427. Accessed April 2, 2020.

71. Singh H, Daci K, Petersen L, Collins C, Petersen N, Shethia A, El-Serag HB. Missed opportunities to initiate endoscopic evaluation for colorectal cancer diagnosis. Am J Gastroenterol. 2009;104(10):2543-54. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2758321/. Accessed April 2, 2020.

72. Agency for Healthcare Research and Quality. Triggers and Targeted Injury Detection Systems (TIDS) Expert Panel Meeting: Conference Summary Report. Rockville, MD. AHRQ Pub. No. 09¬0003. February 2009. https://psnet.ahrq.gov/issue/triggers-and-targeted-injury-detection-systems-tids-expert-panel-meeting-conference-summary. Accessed April 2, 2020.

73. Mull HJ, Nebeker JR, Shimada SL, Kaafarani HM, Rivard PE, Rosen AK. Consensus building for development of outpatient adverse drug event triggers. J Patient Saf. 2011;7(2):66-71. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558884/. Accessed April 2, 2020.

74. Szekendi MK, Sullivan C, Bobb A, Feinglass J, Rooney D, Barnard C, Noskin GA. Active surveillance using electronic triggers to detect adverse events in hospitalized patients. Qual Saf Health Care. 2006;15(3):184-90. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2464849/. Accessed April 2, 2020.

75. Agency for Healthcare Research and Quality. Patient Safety Primer. Triggers and Trigger Tools. Last updated September 2019. https://psnet.ahrq.gov/primer/triggers-and-trigger-tools. Accessed April 2, 2020.

76. Shenvi EC, El-Kareh R. Clinical criteria to screen for inpatient diagnostic errors: a scoping review. Diagnosis. 2015 Feb;2(1):3-19. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4474234/. Accessed April 2, 2020.

77. Singh H, Thomas EJ, Khan MM, Petersen LA. Identifying diagnostic errors in primary care using an electronic screening algorithm. Arch Intern Med. 2007;167(3):302-8. https://www.ncbi.nlm.nih.gov/pubmed/17296888. Accessed April 2, 2020.

78. Murphy DR, Laxmisan A, Reis BA, Thomas EJ, Esquivel A, Forjuoh SN, Parikh R, Khan MM, Singh H. Electronic health record-based triggers to detect potential delays in cancer diagnosis. BMJ Qual Saf. 2014;23(1):8-16. https://www.ncbi.nlm.nih.gov/pubmed/23873756. Accessed April 2, 2020.

79. Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. JAMA. 1991;266(20):2847-51. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1744018/. Accessed April 2, 2020.

80. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, Whittington JC, Frankel A, Seger A, James BC. "Global trigger tool" shows that adverse events in hospitals may be ten times greater than previously measured. Health Aff (Millwood). 2011;30(4):581-9. https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2011.0190?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed. Accessed April 2, 2020. Accessed April 2, 2020.

81. Doupi P SH, Bjørn B, Deilkås E, Nylén U, Rutberg H. Use of the Global Trigger Tool in patient safety improvement efforts: Nordic experiences. Cogn Technol Work. 2015;17(1):45-54. https://link.springer.com/article/10.1007/s10111-014-0302-2.

82. Classen D, Li M, Miller S, Ladner D. An electronic health record-based real-time analytics program for patient safety surveillance and improvement. Health Aff (Millwood). 2018;37(11):1805-12. https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2018.0728?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed. Accessed April 2, 2020.

83. Danforth KN, Smith AE, Loo RK, Jacobsen SJ, Mittman BS, Kanter MH. Electronic clinical surveillance to improve outpatient care: diverse applications within an integrated delivery system. eGEMs 2014;2(!):1056. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4371433/. Accessed April 2, 2020.

84. Murphy DR, Meyer AN, Vaghani V, Russo E, Sittig DF, Richards KA, Wei L, Wu L, Singh H. Application of electronic algorithms to improve diagnostic evaluation for bladder cancer. Appl Clin Inform. 2017;8(1):279-90. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5373770/. Accessed April 2, 2020.

85. Murphy DR, Wu L, Thomas EJ, Forjuoh SN, Meyer AN, Singh H. Electronic trigger-based intervention to reduce delays in diagnostic evaluation for cancer: a cluster randomized controlled trial. J Clin Oncol. 2015;33(31):3560-7. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4622097/. Accessed April 2, 2020.

86. Murphy DR, Thomas EJ, Meyer AN, Singh H. Development and validation of electronic health record-based triggers to detect delays in follow-up of abnormal lung imaging findings. Radiology. 2015:142530. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4613876/. Accessed April 2, 2020.

87. Singh H, Hirani K, Kadiyala H, Rudomiotov O, Davis T, Khan MM, Wahls TL. Characteristics and predictors of missed opportunities in lung cancer diagnosis: an electronic health record-based study. J Clin Oncol. 2010;28(20):3307-15. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2903328/. Accessed April 2, 2020.

88. National Quality Forum. Improving Diagnostic Quality and Safety Final Report. (Developed under Department of Health and Human Services Contract HHSM-500-2012-00009I, Task Order HHSM-500-T0026.) Washington, DC: NQF; September 19, 2017. http://www.qualityforum.org/Publications/2017/09/Improving_Diagnostic_Quality_and_Safety_Final_Report.aspx. Accessed April 2, 2020.

89. Young IJB, Luz S, Lone N. A systematic review of natural language processing for classification tasks in the field of incident reporting and adverse event analysis. Int J Med Inform. 2019;132:103971. https://www.ncbi.nlm.nih.gov/pubmed/31630063. Accessed April 2, 2020.

90. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. J Am Med Inform Assoc. 2005;12(4):448-57. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1174890/. Accessed April 2, 2020.

91. Fong A, Harriott N, Walters DM, Foley H, Morrissey R, Ratwani RR. Integrating natural language processing expertise with patient safety event review committees to improve the analysis of medication events. Int J Med Inform. 2017;104:120-5. https://www.ncbi.nlm.nih.gov/pubmed/28529113. Accessed April 2, 2020.

92. Jagannatha A, Liu F, Liu W, Yu H. Overview of the First Natural Language Processing Challenge for Extracting Medication, Indication, and Adverse Drug Events From Electronic Health Record Notes (MADE 1.0). Drug Saf. 2019;42(1):99-111. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6860017/. Accessed April 2, 2020.

93. Singh H, Graber M, Onakpoya I, Schiff GD, Thompson MJ. The global burden of diagnostic errors in primary care. BMJ Qual Saf. 2016;26(6):484-94. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5502242/. Accessed April 2, 2020.

94. Al-Mutairi A, Meyer AN, Thomas EJ, Etchegaray JM, Roy KM, Davalos MC, Sheikh S, Singh H. Accuracy of the Safer Dx Instrument to identify diagnostic errors in primary care. J Gen Intern Med. 2016;31(6):602-8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4870415/. Accessed April 2, 2020.

95. Cifra CL, Ten Eyck P, Dawson JD, Reisinger HS, Singh H, Herwaldt LA. Factors associated with diagnostic error on admission to a PICU: a pilot study. Pediatr Crit Care Med. 2020. https://www.ncbi.nlm.nih.gov/pubmed/32097247. Accessed April 2, 2020.

96. Bergl PA, Taneja A, El-Kareh R, Singh H, Nanchal RS. Frequency, risk factors, causes, and consequences of diagnostic errors in critically ill medical patients: a retrospective cohort study. Crit Care Med. 2019;47(11):e902-e10. https://www.ncbi.nlm.nih.gov/pubmed/31524644. Accessed April 2, 2020.

97. Singh H, Khanna A, Spitzmueller C, Meyer A. Recommendations for using the Revised Safer Dx Instrument to help measure and improve diagnostic safety. Diagnosis 2019 Nov 26;6(4):315-23. https://www.ncbi.nlm.nih.gov/pubmed/31287795. Accessed April 2, 2020.

98. Mathews BK, Fredrickson M, Sebasky M, Seymann G, Ramamoorthy S, Vilke G, Sloane C, Thorson E, El-Kareh R. Structured case reviews for organizational learning about diagnostic vulnerabilities: initial experiences from two medical centers. Diagnosis (Berl). 2020;7(1):27-35. https://www.ncbi.nlm.nih.gov/pubmed/31444963. Accessed April 2, 2020.

99. Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Lambert BL, Elstein AS, Hasler S, Kabongo ML, Krosnjar N, Odwazny R, Wisniewski MF, McNutt RA. Diagnostic error in medicine: analysis of 583 physician-reported errors. Arch Intern Med. 2009;169(20):1881-7. https://www.ncbi.nlm.nih.gov/pubmed/19901140. Accessed April 2, 2020.

100. Reilly JB, Myers JS, Salvador D, Trowbridge RL. Use of a novel, modified fishbone diagram to analyze diagnostic errors. Diagnosis. 2014;1(2):167-71. https://www.ncbi.nlm.nih.gov/pubmed/29539996. Accessed April 2, 2020.

101. Graber ML, Rencic J, Rusz D, Papa F, Croskerry P, Zierler B, Harkless G, Giuliano M, Schoenbaum S, Colford C, Cahill M, Olson APJ. Improving diagnosis by improving education: a policy brief on education in healthcare professions. Diagnosis (Berl). 2018;5(3):107-18. doi: 10.1515/dx-2018-0033.

102. Henriksen K, Dymek C, Harrison MI, Brady PJ, Arnold SB. Challenges and opportunities from the Agency for Healthcare Research and Quality (AHRQ) research summit on improving diagnosis: a proceedings review. Diagnosis (Berl). 2017;4(2):57-66. https://www.ncbi.nlm.nih.gov/pubmed/29536924. Accessed April 2, 2020.

103. Ohio Hospital Association. OPSI Offers Free Diagnostic Errors Webinar on Sept. 19. https://ohiohospitals.org/News-Publications/Subscriptions/Member-Newsletters/HIINformation/HIINformation/OPSI-Offers-Free-Diagnostic-Errors-Webinar-on-Sept. Accessed April 2, 2020.

104. Medford-Davis L, Park E, Shlamovitz G, Suliburk J, Meyer AN, Singh H. Diagnostic errors related to acute abdominal pain in the emergency department. Emerg Med J. 2016;33(4):253-9. https://www.ncbi.nlm.nih.gov/pubmed/26531859. Accessed April 2, 2020.

105. Singh H, Graber M. Reducing diagnostic error through medical home-based primary care reform. JAMA. 2010;304(4):463-4. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3120138/. Accessed April 2, 2020.

106. Sittig DF, Singh H. Toward more proactive approaches to safety in the electronic health record era. Jt Comm J Qual Patient Saf. 2017;43(10):540-7. doi: 10.1016/j.jcjq.2017.06.005.

107. Institute for Healthcare Improvement/National Patient Safety Foundation. Closing the Loop: A Guide to Safer Ambulatory Referrals in the EHR Era. Cambridge, MA: IHI; 2017. http://www.ihi.org/resources/Pages/Publications/Closing-the-Loop-A-Guide-to-Safer-Ambulatory-Referrals.aspx. Accessed April 2, 2020.

108. Agency for Healthcare Research and Quality. Team STEPPS. https://www.ahrq.gov/teamstepps/index.html. Accessed April 2, 2020.

109. Bates DW, Singh H. Two decades since *To Err Is Human*: an assessment of progress and emerging priorities in patient safety. Health Aff (Millwood). 2018;37(11). https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2018.0738?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dpubmed. Accessed April 2, 2020.

110. Meyer AND, Singh H. The path to diagnostic excellence includes feedback to calibrate how clinicians think. JAMA. 2019;321(8):737-8. doi: 10.1001/jama.2019.0113.

111. Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study. JAMA Intern Med. 2013;173(21):1952-8. https://www.ncbi.nlm.nih.gov/pubmed/23979070. Accessed April 2, 2020.

112. Schiff GD, Martin SA, Eidelman DH, Volk LA, Ruan E, Cassel C, Galanter W, Johnson M, Jutel A, Kroenke K, Lambert BL, Lexchin J, Myers S, Miller A, Mushlin S, Sanders L, Sheikh A. Ten principles for more conservative, care-full diagnosis. Ann Intern Med. 2018;169(9):643-5. https://annals.org/aim/fullarticle/2705208/ten-principles-more-conservative-care-full-diagnosis. Accessed April 2, 2020.

113. Cifra CL, Jones KL, Ascenzi JA, Bhalala US, Bembea MM, Newman-Toker DE, Fackler JC, Miller MR. Diagnostic errors in a PICU: insights from the morbidity and mortality conference. Pediatr Crit Care Med. 2015;16(5):468-76. https://www.ncbi.nlm.nih.gov/pubmed/25838150. Accessed April 2, 2020.

114. Gupta A, Snyder A, Kachalia A, Flanders S, Saint S, Chopra V. Malpractice claims related to diagnostic errors in the hospital. BMJ Qual Saf. 2017;27(1). https://www.ncbi.nlm.nih.gov/pubmed/28794243. Accessed April 2, 2020.

115. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. JAMA. 2003;289(21):2849-56. https://www.ncbi.nlm.nih.gov/pubmed/12783916. Accessed April 2, 2020.

116. Walton MM, Harrison R, Kelly P, Smith-Merry J, Manias E, Jorm C, Iedema R. Patients' reports of adverse events: a data linkage study of Australian adults aged 45 years and over. BMJ Qual Saf. 2017;26(9):743-50. https://www.ncbi.nlm.nih.gov/pubmed/28356333. Accessed April 2, 2020.

117. Fowler FJ, Jr, Epstein A, Weingart SN, Annas CL, Bolcic-Jankovic D, Clarridge B, Schneider EC, Weissman JS. Adverse events during hospitalization: results of a patient survey. Jt Comm J Qual Patient Saf. 2008;34(10):583-90. https://www.ncbi.nlm.nih.gov/pubmed/18947118. Accessed April 2, 2020.

AHRQ Pub. No. 20-0040-1-EF
April 2020