

Guidelines and Recommendations

Hardeep Singh*, Arushi Khanna, Christiane Spitzmueller and Ashley N.D. Meyer

Recommendations for using the Revised Safer Dx Instrument to help measure and improve diagnostic safety

<https://doi.org/10.1515/dx-2019-0012>

Received February 15, 2019; accepted June 7, 2019; previously published online July 9, 2019

Abstract: The medical record continues to be one of the most useful and accessible sources of information to examine the diagnostic process. However, medical record review studies of diagnostic errors have often used subjective judgments and found low inter-rater agreement among reviewers when determining the presence or absence of diagnostic error. In our previous work, we developed a structured data-collection instrument, called the Safer Dx Instrument, consisting of objective criteria to improve the accuracy of assessing diagnostic errors in primary care. This paper proposes recommendations on how clinicians and health care organizations could use the Revised Safer Dx Instrument in identifying and understanding missed opportunities to make correct and timely diagnoses. The instrument revisions addressed both methodological and implementation issues identified during initial use and included refinements to the instrument to allow broader application across all health care settings. In addition to leveraging knowledge from piloting the instrument in several health care settings, we gained insights from multiple researchers who had used the instrument in studies involving emergency care, inpatient care and intensive care unit settings. This allowed us to enhance and extend the scope of this previously

validated data collection instrument. In this paper, we describe the refinement process and provide recommendations for application and use of the Revised Safer Dx Instrument across a broad range of health care settings. The instrument can help users identify potential diagnostic errors in a standardized way for further analysis and safety improvement efforts as well as provide data for clinician feedback and reflection. With wider adoption and use by clinicians and health systems, the Revised Safer Dx Instrument could help propel the science of measuring and reducing diagnostic errors forward.

Keywords: diagnostic error; health system; measurement; patient safety; quality assurance.

Introduction

The medical record continues to be one of the most useful and accessible sources of information to examine the diagnostic process [1–3]. However, the processes involved in making a diagnosis, and those involved in a diagnostic error, are rarely black and white [4]. Judgments to identify potential diagnostic errors based on medical record reviews are correspondingly difficult and full of nuances. Clinicians in hindsight [5, 6] often do not agree on details about the clinical situation, whether decisions were made appropriately, whether certain tests or consultations should have been requested, whether specific diagnostic information was available to the treating clinician at the time of decision-making and whether an error occurred. Sometimes it is difficult to conclusively determine if a diagnosis was really missed in the context of an evolving disease condition especially when patients present with undifferentiated symptoms. Thus, most medical record reviews for diagnostic errors are associated with subjective judgments and low inter-rater agreement among reviewers [2, 3].

Better tools are needed to guide clinicians and safety professionals in evaluating cases or events comprehensively and to more objectively determine the presence or absence of diagnostic errors. In our previous work,

*Corresponding author: Hardeep Singh, MD, MPH, Center for Innovation in Quality, Effectiveness, and Safety (IQuEST) (152), Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC), 2002 Holcombe Boulevard, Houston, TX 77030, USA; and Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, TX, USA, E-mail: hardeeps@bcm.edu

Arushi Khanna and Ashley N.D. Meyer: Center for Innovation in Quality, Effectiveness, and Safety (IQuEST) (152), Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC), Houston, TX, USA; and Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, TX, USA, E-mail: ameyer@bcm.edu (A.N.D. Meyer)

Christiane Spitzmueller: University of Houston, Houston, TX, USA

we developed a structured data collection instrument called the Safer Dx Instrument, which consists of objective criteria to improve the accuracy of assessing diagnostic errors in primary care [7]. The Safer Dx Instrument is a screening tool to help determine the presence or absence of a diagnostic error for a specific episode of care. The instrument uses a previously published definition of diagnostic errors: “missed opportunities to make a correct or timely diagnosis based on the available evidence, regardless of patient harm” [8, 9]. It can be used to guide a comprehensive assessment of the diagnostic process through a detailed examination of all aspects of the patient’s medical record, including patient history, examination, diagnostic test interpretation and follow-up, ordering of additional testing or referrals, and diagnostic assessment [7]. The tool helps the reviewer think through the five main aspects of the diagnostic process described in the Safer Dx Framework (a framework describing the structure, process and outcomes involved in measuring and improving diagnosis) [10]: (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. The ultimate decision on the presence or absence of error depends on the reviewer’s overall judgment after systematically considering the individual items [10]. This basic terminology and approach helps reviewers come to a shared understanding to make this determination.

We previously conducted two studies to validate the instrument, one in primary care and the other in a pediatric intensive care unit (PICU) setting [11]. However, the use of the instrument in the PICU setting as well as preliminary feedback from other reviewers exploring the use of the instrument in additional settings suggested the need for changes in the instrument content for broader use, including the need for changes in wording of individual items and changes to the scale. Thus, in the next iteration of this work, we refined the instrument for broader application across all health care settings. We leveraged knowledge from piloting the instrument in several health care settings and gained insights from multiple researchers who had used the instrument in research studies involving emergency care, inpatient care and intensive care unit settings. This allowed us to enhance and extend the scope of this previously validated data collection instrument that had been developed primarily for research of diagnostic errors in primary care. Based on these activities and collected experiences, we then developed guidance

for health care organizations, researchers, clinicians and patient safety professionals to use this instrument to identify diagnostic errors in routine practice. In this paper, we introduce the Revised Safer Dx Instrument, summarize the refinement process and put forth recommendations for using the revised instrument with the goal of helping clinicians and health systems improve measurement of diagnostic errors more broadly.

Rationale and process of refinement

The Safer Dx Instrument requires reviewers to “decompose” their judgments, and pay systematic attention to factors that are broken up into smaller components rather than rely only on subjective judgments that may be more likely to be affected by extraneous effects [12, 13]. Instrument questions guide reviewers to assess diagnostic processes through detailed evaluation of the patient’s medical record, including history, examination, test ordering/interpretation, follow-up and diagnostic assessment.

Using multiple reviews of cases with and without diagnostic error, we iteratively modified the instrument to help determine the presence of diagnostic error in various settings (e.g. emergency, inpatient, ICU). For further refinement, we recruited 10 subject matter experts, including four researchers who independently used the instrument in their own research projects, combining knowledge from more than 400 collective records reviewed with the instrument. These researchers had an advanced understanding of the topic and personally used the instrument in a variety of settings including primary care, both medical and pediatric intensive care, medical and pediatric emergency care, and inpatient care. We solicited researchers’ feedback via telephone conferences and email. These expert users weighed in on their overall experience using the instrument, difficulties encountered in trying to utilize the tool, modifications needed for the setting in which they used it, and improvements for future use and broad dissemination. We synthesized this feedback to inform revisions to the instrument and for creating supplemental guidance for its use.

Instrument modifications

Several issues were addressed with the revision. Our previous scale was a six-point scale, yet many users reported the desire for a neutral option [14]. We considered a seven- and not a five-point scale, to increase nuance [15].

We wanted to eliminate the negatively worded items (e.g. “I am not confident in my diagnosis”) based on psychological research on measurement, which support not having items scored in two different directions when measuring an underlying latent construct [16]. We switched to the term “missed diagnostic opportunity” consistently and removed references to diagnostic error on the instrument to enhance acceptability [17].

As depicted in Figure 1, revisions to the instrument include changes to the overall structure as well as individual questions. Based on feedback from the subject matter experts, the scale of the instrument was reworked to have a neutral point, 4, on a Likert scale of 1–7 rather than 1–6. The scale was additionally reversed so higher scores are now indicative of a diagnostic error. Questions

7, 10, 11 and 12 were structurally modified to score in the same direction as the remaining questions. To ensure that the instrument addresses all potential sources of error across the Safer Dx Framework, Question 9 was added. Finally, several items were reworded; most significantly, “diagnostic error” was replaced with “diagnostic missed opportunity”.

The refined instrument (Supplementary Material, Appendix 1) consists of 12 questions to help clinician-reviewers cognitively deconstruct the diagnostic processes and one final-judgment question regarding the presence/absence of a missed opportunity to make a timely, accurate diagnosis. Each question uses a Likert scale of 1–7 to reflect the “grayness” of determining missed opportunities.

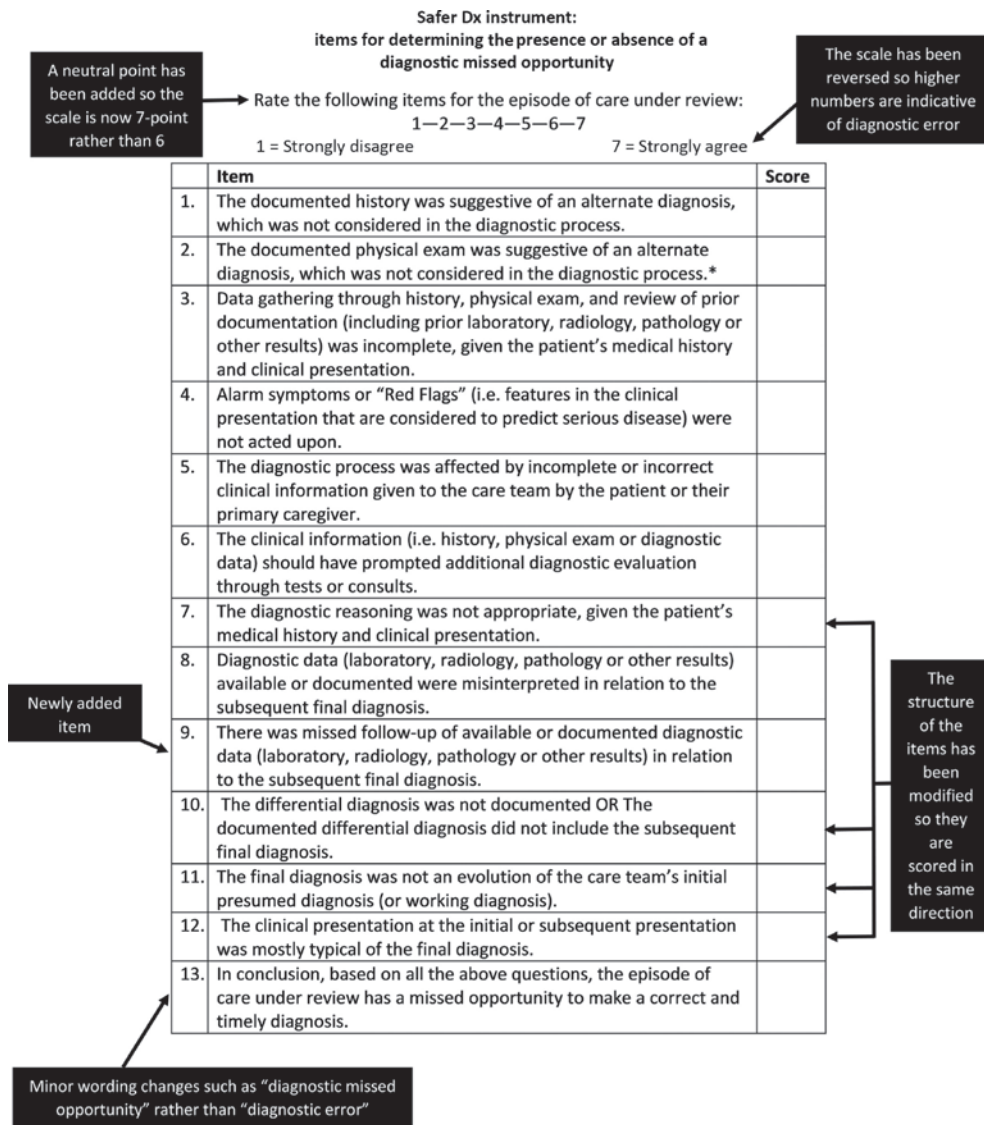


Figure 1: Summary of revisions made to the Safer Dx Instrument.

Recommendations for use

Safer Dx Instrument-facilitated measurement of diagnostic errors is intended to help users identify potential diagnostic errors in a standardized way for further review and improvement efforts. The best yield would result from review of comprehensive electronic health records (EHRs) containing longitudinal patient care data (progress notes, tests, referrals) that helps explain a patient's diagnostic journey. We developed guidance to define key terms (e.g. diagnostic errors, missed opportunities and diagnostic processes) to allow for broader use by researchers, safety professionals and clinicians, as well as to help standardize review procedures and facilitate instrument dissemination. This guidance is essential for creating a shared mental model around the complex error determination process. The following recommendations will be useful in applying the instrument across a broad range of health care settings and are each further described as follows:

- Operationalize a shared understanding of diagnostic error
- Define the episode of care being evaluated
- Consider diagnostic evolution in terms of initial presumed or working diagnosis
- Evaluate the diagnostic process rather than the ultimate patient outcome in hindsight
- Account for non-applicable and other gray zone situations
- Determine the presence/absence of missed opportunity based on an overall assessment
- Gather additional context to explain reviewer judgments
- Analyze diagnostic process breakdowns involved
- Determine preventable diagnostic harm

Recommendation 1: Operationalize a shared understanding of diagnostic error

There are several definitions of diagnostic error, including a definition recently proposed in the National Academies of Sciences, Engineering, and Medicine's (NAEM) report on improving diagnosis [1], and each contains its own nuances. The Safer Dx Instrument was developed based on extensive research using a definition where diagnostic errors are defined as missed opportunities to make correct or timely diagnoses based on the available evidence, regardless of patient harm [10]. This allows users to focus on situations where similar errors can be prevented in the future, but also minimizes the hindsight bias often seen when focusing only on harmful cases. Thus, for the

Revised Safer Dx Instrument purposes, we suggest using the following three criteria [8, 18], which we previously found useful in defining diagnostic errors [8]:

- *Case analysis reveals evidence of a missed opportunity to make a correct or timely diagnosis.* This criterion suggests the concept of a missed opportunity, i.e. could something different have been done to make the correct diagnosis earlier, a question that every reviewer must ask before they determine the presence or absence of diagnostic error. For example, if a patient with history of intravenous (IV) drug use presented to primary care with a week of neck pain and fever, a reviewer considers it a missed opportunity if magnetic resonance imaging (MRI) of the cervical spine was not considered.
- *A missed opportunity is framed within the context of an “evolving” diagnostic process.* This criterion takes into account the temporal or sequential context of events to determine missed opportunities and looks for evidence of omission (failure to do the right thing) or commission (doing something wrong) at the particular point in time at which the “error” occurred. For example, no missed opportunity would be determined when an otherwise young, healthy patient with 2 days of dry cough and no other associated signs and symptoms was treated with watchful waiting even if she returned in 5 days with fever and a subsequent chest x-ray showed community-acquired pneumonia.
- *The opportunity could be missed by the clinician, care team, system and/or patient.* This criterion suggests a system-centric vs. clinician-centric approach to diagnostic error and underscores that the missed opportunity may result from cognitive and/or system factors and usually both. For instance, a busy, fast-paced and chaotic emergency room setting with many interruptions might influence how history, examination and other forms of data gathering occurs, and diagnostic error could result from missing some critical piece of diagnostic information [19]. Very infrequently in our experience is an error attributable to more blatant factors, such as lapses in accountability or clear evidence of liability or negligence by a single individual.

Additionally, some diagnoses may be delayed or incorrect but do not necessarily involve missed opportunities because nothing different could have been done, i.e. the situation was not really preventable given currently available, best scientific knowledge and practice (a situation often seen in diagnosis of rare diseases). However, even preventable errors or delays in diagnosis sometimes occur due to factors outside the clinician's immediate control or when a clinician's performance is not contributory.

Examples include missed follow-up of an abnormal test result due to a broken interface between the institution's laboratory and EHR.

Recommendation 2: Define the episode of care being evaluated

The episode of care should include all services a patient receives for a specific health problem in a given time period. For example, if a patient presents to primary care for evaluation of new-onset shortness of breath, the episode of care includes everything that was done to evaluate and treat that health problem in a defined time period (e.g. a patient was later admitted for pneumonia, given IV antibiotics and improved). The episode of care might vary in time according to the types of patient presentations being reviewed, but reviewers should come to relative consensus on time frames suitable for review in their setting. Reviewers should look back to evaluate if the problem is new vs. old, when it started and when the patient presented as well as look forward to see what was done for the problem and how the diagnostic process unfolded. Comprehensive EHRs can facilitate this review, which often involves looking at outpatient and inpatient progress notes, tests, procedures, consultations, and other diagnostic information.

Although there is no hard and fast rule defining an episode of care, we recommend focusing on initial patient presentations and related subsequent encounters related to the situation in question [11]. For instance, for a patient whose care was escalated to the ICU from the hospital floor after a rapid response on day 2 of admission, review not only the day of rapid response but also when the patient presented to the emergency department and what happened in the first 48 h of care. In a study looking at diagnostic errors during a 7-day acute care hospitalization or a 7-day ICU stay, reviewers could consider the entire length of time of 7 days as one episode of care for review purposes. Sometimes, a single initial visit itself can guide several answers [1]. Reviewers will often need to use their clinical judgment and consider multiple visits within an episode of care.

Recommendation 3: Consider diagnostic evolution in terms of initial presumed or working diagnosis

One of the most common reasons why diagnostic errors are so complex is that diagnosis often evolves over time.

For instance, patients often present to primary care or emergency departments with subtle or undifferentiated symptoms, precluding a clear diagnosis at that point in time. A healthy-looking child with 3 days of abdominal pain and constipation but no other sign or symptom may not warrant an evaluation for appendicitis, vs. another child who presents with abdominal pain, tenderness, fever and leukocytosis who must be evaluated further. Clinicians often balance between undertesting and over-aggressive diagnostic pursuits that could be harmful and costly. Sometimes the story becomes clearer with time, or a second or third visit with the clinician when more signs and symptoms or a more succinct patient story suggests a probable diagnosis for which additional evaluation needs to be done. We thus suggest that reviewers must account for the uncertainties involved with such situations.

To prevent overcalling diagnostic errors, reviewers should use the following concepts from the NASEM report to conduct their reviews [1]: *“The working diagnosis may be either a list of potential diagnoses (a differential diagnosis) or a single potential diagnosis. Typically, clinicians will consider more than one diagnostic hypothesis or possibility as an explanation of the patient’s symptoms and will refine this list as further information is obtained in the diagnostic process.”*

An initial presumed diagnosis would be the most likely diagnosis out of several potential diagnoses for a patient’s illness that the provider considered in a differential diagnosis or one for which the clinician opted to initiate treatment or further focused diagnostic evaluation.

Recommendation 4: Evaluate the diagnostic process rather than the ultimate patient outcome in hindsight

It is easier and even tempting to look at a case in hindsight especially when outcomes were poor (the patient was harmed, died or had care escalation) and pass a judgment on the appropriateness of care. However, information available to the care team may not have allowed them to make the decision which later on seems like the correct one. Blinding the initial reviewers to the patient’s outcome is ideal but often not possible while reviewing medical records, especially when EHRs contain easily accessible and comprehensive diagnostic information about a patient’s longitudinal diagnostic journey. Thus, to avoid hindsight bias and to determine whether there was a missed opportunity, we suggest looking at the diagnostic process in depth rather than focusing on the ultimate accuracy of diagnosis or the potentially adverse

patient outcome. This also allows for the examination of processes that might have been poor, but were somehow rectified through other means (e.g. through resilience [20] or chance [21]).

The NASEM report defines the diagnostic process as follows: *“First, a patient experiences a health problem. The patient is likely the first person to consider his or her symptoms and may choose at this point to engage with the health care system. Once a patient seeks health care, there is an iterative process of information gathering, information integration and interpretation, and determining a working diagnosis. Performing a clinical history and interview, conducting a physical exam, performing diagnostic testing and referring or consulting with other clinicians are all ways of accumulating information that may be relevant to understanding a patient’s health problem. The information-gathering approaches can be employed at various times, and diagnostic information can be obtained in different orders. The continuous process of information gathering, integration and interpretation involves hypothesis generation and updating prior probabilities as more information is learned. Communication among health care professionals, the patient and the patient’s family members is critical in this cycle of information gathering, integration and interpretation”*.

The Safer Dx Instrument is heavily focused on missed opportunities in the five Safer Dx Framework’s process dimensions described before, all of which were integrated within NASEM’s conceptualization of the diagnostic process.

Recommendation 5: Account for non-applicable and other gray zone situations

For a question that does not apply for the specific case under review, the reviewer should record option 1, indicating they strongly disagree. For example, for question 4, if there are no alarm symptoms or “Red Flags” to begin with, the reviewer simply records 1 for the item and moves to the next question.

Recommendation 6: Determine the presence/absence of missed opportunity based on an overall assessment

We caution others against just using an overall score as a measure of a missed diagnosis at this point, especially because we see cases of clear missed opportunities with just one item with a high score. We recommend

the reviewers make a final judgment call on the presence/absence of missed opportunity based on an overall assessment of instrument items. Because this is only a screening instrument, if the rating by the initial reviewer is 4 or higher on question 13, we recommend the case be discussed or reviewed independently by at least one other reviewer (depending on resources and availability of clinician reviewers). If both the screening reviewer and additional reviewer(s) agree (for example, majority of ratings of >4 on question 13), the episode of care has a high likelihood of having a diagnostic error and should be reviewed in detail for learning and improvement opportunities. In cases of disagreement, i.e. a second reviewer rates question 13 as 1–3, an additional reviewer (if available) might add additional insights in the process of adjudication. However, in resource-constrained settings, it may be better to focus on cases where reviewers have indicated a much higher likelihood of the presence of a missed opportunity, such as >5. Future users could discuss and consider thresholds for adjudication through secondary and tertiary reviews depending on the availability of resources and how many records need to be reviewed.

Recommendation 7: Gather additional context to explain reviewer judgments

At times, regardless of the presence of diagnostic error, the care episode involves a management error. This could include instances of incorrect therapy or disposition after a correct diagnosis. For instance, a patient may present with a fracture that was diagnosed correctly but was treated with the wrong intervention which led to return visits and additional procedures. While occasionally there is overlap between diagnostic and management errors and at times both are simultaneously present, it is important to collect this information for context and to avoid confusing the two. We also suggest gathering specific circumstantial information, if present, about care escalation (e.g. hospitalization at subsequent visit). Care escalation could be related to worsening of an original correctly diagnosed condition, such as an abscess that progresses from a correctly diagnosed and treated skin infection, vs. related to something being missed initially, such as an infection being completely missed. Similarly, some patients refuse hospitalization or additional evaluation initially, information that should be captured to help with further analysis and adjudication. We find it most helpful when reviewers provide a few sentences of narrative text explaining their judgments that helped with the final decision, either for or against missed opportunities or management errors. This

context is useful at the time of adjudication. When possible, engaging involved providers to obtain additional context will also be useful.

Recommendation 8: Analyze diagnostic process breakdowns involved

If a diagnostic error is found to be present, consider analyzing what types of breakdowns are involved. We have found a taxonomy of five process breakdowns useful in categorizing the various types of problems in the diagnostic process (derived from the Safer Dx Framework): (1) the patient-provider encounter (problems with history, physical examination, ordering tests/referrals based on diagnostic assessment and evaluating prior patient data, such as recent hospitalization at another facility); (2) performance and interpretation of diagnostic tests (some of these problems are within the purview of diagnostic specialties such as radiology and pathology, but clinicians could also have problems with misinterpretation of test results); (3) follow-up and tracking of diagnostic information over time, such as abnormal test results (these could be from missed or delayed communication of test results to clinicians as well as notification of these results to patients); (4) subspecialty and referral-specific factors, such as communication between a hospitalist and a subspecialist; and (5) patient-related factors, such as adherence and behavioral issues [10]. More than one of these breakdowns are often involved and sometimes it is useful to seek input from involved providers if possible. We have developed another supplementary data collection tool, called the “Safer Dx Process Breakdown Supplement” that could add additional details related to the diagnostic error as well as help capture some of the details related to process breakdowns (Supplementary Material, Appendix 2).

Recommendation 9: Determine preventable diagnostic harm

To measure harm from error, we suggest using the National Coordinating Council for Medication Error Reporting and Prevention (NCC MERP) Index for Categorizing Medication Errors [22]. This index examines multiple factors that determine the level of harm such as whether the error reached the patient, if patient harm occurred and the degree of harm. We have adapted this index to reflect the potential severity of injury associated with delay or missed diagnosis. For the purposes of our data collection tool, the

harm definition has been expanded to include physical, emotional, psychological and/or financial distress. This taxonomy is included in the “Safer Dx Process Breakdown Supplement”. Harm should only be determined after an error is confirmed to have occurred, in order to minimize hindsight bias. Depending on the structure of the review team, the team may consider a later set of reviews or another set of reviewers to determine harm.

An example may be helpful to sum up recommendations. If a team of pediatric clinicians want to review potential cases of missed appendicitis, they will first develop a shared understanding of what constitutes diagnostic error for this situation, define the episode of care including time-periods before and after an appendicitis diagnosis, and account for uncertain and evolving situations including discussing when patients with presenting symptoms such as abdominal pain should get additional evaluation. We highly recommend that reviewers initially also calibrate among themselves by piloting the instrument on a set of common cases and discussing the cases to develop a shared mental model. They will then evaluate the diagnostic process in depth using the Revised Safer Dx Instrument and determine breakdowns in missed cases using the Safer Dx Process Breakdown Supplement. They should do all of this while making a distinction between appendicitis diagnosis and management issues, seeking input from involved providers when possible, and ultimately determining preventable diagnostic harm associated with missed opportunities.

Next steps: use analysis for learning

The ultimate goal of this tool is to promote learning and feedback related to diagnostic safety at multiple levels: at the micro-level, including clinician self-reflection or division/hospital peer review, to the macro-level, including health system error recognition and improvement. Thus, we envision clinicians and health systems could use this tool to identify specific learning opportunities that could be communicated to the right audiences for feedback, system changes and improvement strategies [17]. To identify a set of records to review, health systems could leverage their existing safety and quality improvement infrastructure and personnel, and apply diagnostic safety e-trigger tools that select high-risk records to review [23]. Use of e-triggers followed by review using the Revised Safer Dx Instrument can help identify diagnostic events more efficiently, allowing health systems to

monitor event rates, study contributory factors, learn from these events and prevent similar events in the future. This basic approach and associated institutional safety infrastructure can stimulate review exercises. This tool could also be used to inform and improve peer-review processes at hospitals and help use some of those data for learning and improvement [24]. Clinicians could also use this tool for self-reflection on their own records. Currently, no such procedures exist but based on anecdotal data from training reviewers on this tool, they suggested it provided them with useful insights about how to improve their own documentation and processes of care.

In conclusion, the Revised Safer Dx Instrument-facilitated measurement of diagnostic errors could help propel the science of measuring and reducing diagnostic errors forward. By helping reviewers identify potential diagnostic errors in a standardized way for further analysis, it can lead to feedback and reflection for clinicians as well as help inform strategies to reduce diagnostic missed opportunities at the health system level [25].

Acknowledgments: We would like to acknowledge the following subject matter experts for their valuable input to revise the Safer Dx Instrument: Drs. Rita Fernholm, Christina Cifra, Joe Grubenhoff, Prashant Mahajan, Andrew Olson, Paul Bergl, Divvy Upadhyay, Viralkumar Vaghani, Grant Shafer and Suresh Gautham.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

Research funding: Dr. Singh is funded by the Veterans Affairs (VA) Health Services Research and Development Service (HSR&D) (VA IIR-17-127 and the Presidential Early Career Award for Scientists and Engineers USA 14-274), the Agency for Healthcare Research and Quality (R01HS022087 and R18HS017820), the VA National Center for Patient Safety, the Houston VA HSR&D Center for Innovations in Quality, Effectiveness, and Safety (CIN 13-413), Gordon and Betty Moore Foundation and a CanTest Research Collaborative funded by a Cancer Research UK Population Research Catalyst award (C8640/A23385). Dr. Meyer is additionally funded by the Veterans Affairs (VA) Health Services Research and Development Service (HSR&D) (VA HSR&D 1 IK2 HX002586-01A1). The authors' views do not represent those of any of the funders.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Improving Diagnosis in Health Care. National Academies of Sciences Engineering and Medicine 2015. <http://iom.nationalacademies.org/Reports/2015/Improving-Diagnosis-in-Healthcare.aspx>. Accessed: 14 June 2016.
2. 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;173:418–25.
3. Zwaan L, de Bruijne M, Wagner C, Thijs A, Smits M, van der Wal G, et al. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med* 2010;170:1015–21.
4. Zwaan L, Singh H. The challenges in defining and measuring diagnostic error. *Diagnosis* 2015;2:97–103.
5. Singh H, Giardina T, Forjuoh S, Reis MD, Kosmach S, Khan MM, et al. Electronic health record-based surveillance of diagnostic errors in primary care. *BMJ Qual Saf* 2012;21:93–100.
6. Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Lambert BL, et al. Diagnostic error in medicine: analysis of 583 physician-reported errors. *Arch Intern Med* 2009;169:1881–7.
7. Al-Mutairi A, Meyer AN, Thomas EJ, Etchegaray JM, Roy KM, Davalos MC, et al. Accuracy of the Safer Dx Instrument to identify diagnostic errors in primary care. *J Gen Intern Med* 2016;31:602–8.
8. Singh H. Editorial: Helping health care organizations to define diagnostic errors as missed opportunities in diagnosis. *Jt Comm J Qual Patient Saf* 2014;40:99–101.
9. Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. *BMJ Qual Saf* 2016;26:484–94.
10. Singh H, Sittig DF. Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework. *BMJ Qual Saf* 2015;24:103–10.
11. Davalos MC, Samuels K, Meyer AN, Thammasitboon S, Sur M, Roy K, et al. Finding diagnostic errors in children admitted to the PICU. *Pediatr Crit Care Med* 2017;18:265–71.
12. Arkes HR, González-Vallejo C, Bonham AJ, Kung YH, Bailey N. Assessing the merits and faults of holistic and disaggregated judgments. *J Behav Decis Mak* 2010;23:250–70.
13. Ruscio J. Holistic judgment in clinical practice. *Sci Rev Ment Health Pract* 2003;2:38–48.
14. Adelson JL, McCoach DB. Measuring the mathematical attitudes of elementary students: the effects of a 4-point or 5-point Likert-type scale. *Educ Psychol Meas* 2010;70:796–807.
15. Finstad K. Response interpolation and scale sensitivity: evidence against 5-point scales. *J Usability Stud* 2010;5:104–10.
16. Dalal DK, Carter NT. Negatively worded items negatively impact survey research. In: Lance CE, Vandenberg RJ, editors. *More Statistical and Methodological Myths and Urban Legends*. Abingdon-on-Thames: Routledge, 2015:112–32.
17. Meyer AN, Singh H. The path to diagnostic excellence includes feedback to calibrate how clinicians think. *J Am Med Assoc* 2019;321:737–8.

18. Singh H, Sittig DF. Setting the record straight on measuring diagnostic errors. Reply to: 'bad assumptions on primary care diagnostic errors' by Dr Richard Young. *BMJ Qual Saf* 2015;24:345–8.
19. Upadhyay DK, Sittig DF, Singh H. Ebola US patient zero: lessons on misdiagnosis and effective use of electronic health record. *Diagnosis* 2014;1:283–8.
20. Smith MW, Davis GT, Murphy DR, Laxmisan A, Singh H. Resilient actions in the diagnostic process and system performance. *BMJ Qual Saf* 2013;22:1006–13.
21. Reason J. Human error: models and management. *Br Med J* 2000;320:768–70.
22. Hartwig SC, Denger SD, Schneider PJ. Severity-indexed, incident report-based medication error-reporting program. *Am J Hosp Pharm* 1991;48:2611–6.
23. 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;28:151–9.
24. 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;23:1023–30.
25. McGlynn EA, McDonald KM, Cassel CK. Measurement Is Essential for Improving Diagnosis and reducing diagnostic error: a report from the institute of medicine. *J Am Med Assoc* 2015;314:2501–2.

Supplementary Material: The online version of this article offers supplementary material (<https://doi.org/10.1515/dx-2019-0012>).