



OPEN ACCESS

Application of electronic trigger tools to identify targets for improving diagnostic safety

Daniel R Murphy,^{1,2} Ashley ND Meyer,^{1,2} Dean F Sittig,^{1,3,4}
Derek W Meeks,^{1,2} Eric J Thomas,⁴ Hardeep Singh^{1,2}

¹Center for Innovations in Quality, Effectiveness and Safety, Michael E DeBakey Veterans Affairs Medical Center, Houston, Texas, USA

²Department of Medicine, Baylor College of Medicine, Houston, Texas, USA

³School of Biomedical Informatics, University of Texas Health Science Center, Houston, Texas, USA

⁴Department of Medicine, University of Texas-Memorial Hermann Center for Healthcare Quality and Safety, Houston, Texas, USA

Correspondence to

Dr Daniel R Murphy, Center for Innovations in Quality, Effectiveness and Safety, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX 77030, USA; drmurphy@bcm.edu

Received 14 March 2018

Revised 20 June 2018

Accepted 14 August 2018

Published Online First

5 October 2018

ABSTRACT

Progress in reducing diagnostic errors remains slow partly due to poorly defined methods to identify errors, high-risk situations, and adverse events. Electronic trigger (e-trigger) tools, which mine vast amounts of patient data to identify signals indicative of a likely error or adverse event, offer a promising method to efficiently identify errors. The increasing amounts of longitudinal electronic data and maturing data warehousing techniques and infrastructure offer an unprecedented opportunity to implement new types of e-trigger tools that use algorithms to identify risks and events related to the diagnostic process. We present a knowledge discovery framework, the Safer Dx Trigger Tools Framework, that enables health systems to develop and implement e-trigger tools to identify and measure diagnostic errors using comprehensive electronic health record (EHR) data. Safer Dx e-trigger tools detect potential diagnostic events, allowing health systems to monitor event rates, study contributory factors and identify targets for improving diagnostic safety. In addition to promoting organisational learning, some e-triggers can monitor data prospectively and help identify patients at high-risk for a future adverse event, enabling clinicians, patients or safety personnel to take preventive actions proactively. Successful application of electronic algorithms requires health systems to invest in clinical informaticists, information technology professionals, patient safety professionals and clinicians, all of who work closely together to overcome development and implementation challenges. We outline key future research, including advances in natural language processing and machine learning, needed to improve effectiveness of e-triggers. Integrating diagnostic safety e-triggers in institutional patient safety strategies can accelerate progress in reducing preventable harm from diagnostic errors.

process.⁷ Use of health information technology (HIT) is essential to monitor patient safety⁸ but has received limited application in diagnostic error detection. Widespread recent adoption of comprehensive electronic health records (EHR) and clinical data warehouses have advanced our ability to collect, store, use and analyse vast amounts of electronic clinical data that helps map the diagnostic process.

Triggers have helped measure safety in hospitals; for example, use of inpatient naloxone administration outside of the postanesthesia recovery room could suggest oversedation due to opioid administration. Trigger development and use have steadily increased over the past decade in prehospital,⁹ emergency room,¹⁰ inpatient,¹¹ ambulatory care¹² and home health settings,¹³ and helped identify adverse drug reactions,¹⁰ surgical complications^{14 15} and other potentially preventable harm.¹⁴ Electronic trigger (e-trigger) tools,¹⁶ which mine vast amounts of clinical and administrative data to identify signals for likely adverse events,^{17–19} offer a promising method to detect patient safety events. Such tools are more efficient and effective in detecting adverse events as compared with voluntary reporting or use of patient safety indicators^{20 21} and offer the ability to quickly mine large data sets, reducing the number of records requiring human review to those at highest risk of harm. While most e-triggers rely on structured (non-free text) data, some can detect specific words within progress notes or reports.²²

The most widely used trigger tools (the Institute for Healthcare Improvement's Global Trigger Tools)²³ include both manual and e-trigger tools to detect inpatient events.^{20 24–28} However, they were not designed to detect diagnostic errors.

Nearly two decades after the Institute of Medicine's report '*To Err is Human*',¹ medical errors remain frequent.^{2–4} Methods are needed to efficiently and effectively identify high-risk situations to prevent harm as well as identify patient safety events to enable organisational learning for error prevention.^{5 6} Measurement needed for improving diagnosis is particularly challenging due to the complexity of an evolving diagnostic



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Murphy DR, Meyer AND, Sittig DF, et al. *BMJ Qual Saf* 2019;**28**:151–159.

Meanwhile, other types of trigger tools have been developed for diagnostic errors and are getting ready for integration within existing patient safety surveillance systems.^{4 29 30} To stimulate progress in this area, we present a knowledge discovery framework, the Safer Dx Trigger Tools Framework, that could enable health systems to develop and implement e-trigger tools that measure diagnostic errors using comprehensive EHR data. Health systems would also need to leverage and/or develop their existing safety and quality improvement infrastructure and personnel (such as clinical leadership, HIT professionals, safety managers and risk management) to operationalise this framework. In addition to showcasing the application of diagnostic safety e-trigger tools, we highlight several strategies to bolster their development and implementation. Triggers can identify diagnostic events, allowing health systems to monitor event rates and study contributory factors, and thus potentially learn from these events and prevent similar events in the future. Some e-triggers additionally allow monitoring of data more prospectively and help identify patients at high risk for future adverse events, enabling clinicians, patients or safety personnel to take preventive actions proactively.

CONCEPTUALISING DIAGNOSTIC SAFETY E-TRIGGERS

Triggers are not new to patient safety measurement. Several existing triggers focus on identifying errors related to medications, such as administering incorrect dosages, or procedure complications, such as returning to the operating room. Only recently has this concept been adapted to detect potential problems with diagnostic processes, such as patterns of care suggestive of missed or delayed diagnoses.¹⁹ For instance, a clinic visit followed several days later by an unplanned hospitalisation or subsequent visit to the emergency department could be indicative of something missed at the first visit.³¹ Similarly, misdiagnosis could be suggested by an unusually prolonged hospital stay for a given diagnosis¹⁹ or an unexpected inpatient transfer to a higher level of care,^{19 32} particularly when

considering younger patients with minimal comorbidity.³³ Event identification can promote organisational learning with the goal of addressing underlying factors that led to the error, similar to what was proposed earlier in the 2015 Safer Dx framework for measuring diagnostic errors.³⁴ Additionally, e-triggers enable tracking of events over time to allow an assessment of the impact of efforts to reduce adverse events.

Certain e-trigger tools can additionally monitor for high-risk situations prospectively, such as when risk of harm is high, even if no harm has yet occurred. Several studies have shown that e-trigger tools offer promise in detecting errors of omission, such as detecting delays in care after an abnormal test result suspicious for cancer,^{26–30 35} kidney failure,^{29 36} infection²⁹ and thyroid conditions,³⁷ as well as patients at risk of delayed action on pregnancy complications.^{38 39} Such triggers can identify situations where earlier intervention can potentially improve patient outcomes. Future e-triggers could explore other process breakdowns associated with diagnostic errors, such as when insufficient history has been collected or diagnostic testing is not completed for a high-risk symptom (eg, no documented fever assessment or temperature recording in patients with back pain, where an undiagnosed spinal epidural abscess might be missed).^{40 41} In table 1, we provide several examples of ‘Safer Dx’ e-trigger tools that align with the process dimensions of the Safer Dx framework. To promote the uptake of Safer Dx trigger tools by health systems, we now discuss essential steps for their development and implementation.

SAFER DX TRIGGER TOOLS FRAMEWORK

Overview

e-Trigger development may be viewed as a form of data mining or pattern matching to discover knowledge about clinical processes. Several knowledge discovery frameworks have evolved from fields such as statistics, machine learning and database research. Hripcsak *et al* proposed a framework for mining complex clinical data for patient safety research, which is composed of seven iterative steps: define the target events, access the

Table 1 Examples of potential Safer Dx e-triggers mapped to diagnostic process dimensions of the Safer Dx framework³⁴

Safer Dx diagnostic process	Safer Dx trigger example	Potential diagnostic error
Patient-provider encounter	Primary care office visit followed by unplanned hospitalisation	Missed red flag findings or incorrect diagnosis during initial office visit
	ER visit within 72 hours after ER or hospital discharge	Missed red flag findings during initial ER or hospital visit
	Unexpected transfer from hospital general floor to ICU	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
Follow-up and tracking of diagnostic information over time	Abnormal test result with no timely follow-up action	Abnormal test result missed
Referral and/or patient-specific factors	Urgent specialty referral followed by discontinued referral or patient no-show within 7 days	Delay in diagnosis from lack of specialty expertise

ER, emergency room; ICU, intensive care unit.

clinical data repository, use natural language processing (NLP) for interpreting narrative data, generate queries to detect and classify events, verify target detection, characterise errors using system or cognitive taxonomies, and provide feedback.⁴² We build on essential components of Hripcsak's framework to demonstrate the steps of Safer Dx e-trigger tools design and development (figure 1), with an emphasis on operationalising them using a multidisciplinary approach.

Development methods

These development methods (table 2) have now been validated to identify several diagnostic events of interest.^{25–29 31 35 36 43}

Step 1: Identify and prioritise diagnostic error of interest

The choice of which diagnostic error to focus on could be guided by high-risk areas identified in prior research and/or local priorities.⁴⁴ Because of challenges to define error, we recommend risk areas where clear evidence exists of a missed opportunity to make a correct or timely diagnosis^{1 45–48} since this emphasises preventability (focusing efforts where improvement is more feasible) and accounts for the evolution of diagnosis over time.

Take, for example, a potentially missed diagnosis of lung cancer related to delayed follow-up after an abnormal chest radiograph.^{26 35} A robust body of literature suggests that poor outcomes and malpractice suits can result from delays in follow-up of abnormal imaging when potential lung malignancies are missed.^{49–51}

Step 2: Operationally define criteria to detect diagnostic error

Developing operational definitions involves creating unambiguous language to objectively describe all inclusion and exclusion characteristics to identify the event. For example, an operational definition of 'unexpected readmission' might be 'unplanned readmission to the same hospital for the same patient within 14 days of discharge'. In many cases, standard definitions will not exist and will need to be developed by patient safety and clinical stakeholders. Published literature, clinical guidelines from academic societies, and input from clinicians, staff and other stakeholders with expertise or involvement in related care processes will allow development of rigorous criteria matched to local processes and site characteristics. Consensus may be achieved by having a designated team review and approve all final criteria or Delphi-like methods⁵² with iterative revisions based on individual feedback and re-review by the group.

In the above example, defining what is an 'abnormal' radiograph, a follow-up action and length of time that should be considered a 'delay' is seemingly straightforward, but in absence of any standards, a key step. 'Abnormal radiographs' could include any plain chest radiograph where the radiologist documents findings

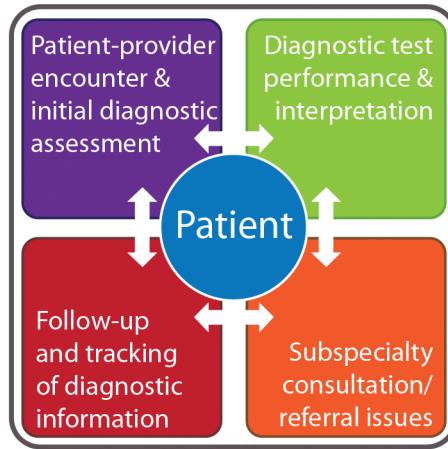
suspicious for new lung malignancy and 'timely' follow-up could include repeat imaging or a lung biopsy performed within 30 days of the initial radiograph. While the 30-day time frame is longer than what is required to act on an abnormal radiograph, it is short enough to 'catch' an abnormality before clinically significant disease progression, allowing an opportunity to intervene. Consensus on this time frame might involve primary care physicians, pulmonologists, oncologists and patient safety experts, and definitions may vary from site to site. The criteria should also exclude patients where additional diagnostic evaluation is unnecessary, such as in those with known lung cancer or terminal illness.

Step 3: Determine potential electronic data sources

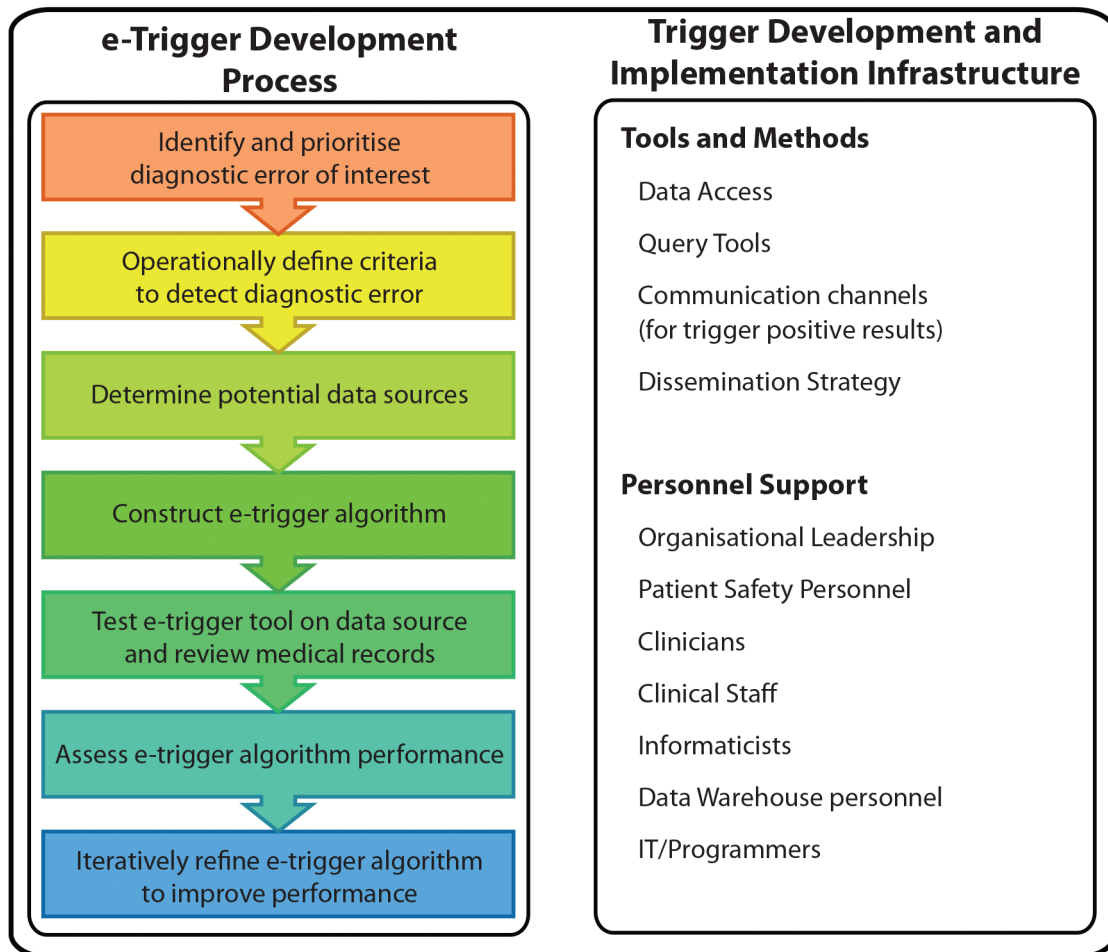
The nature and quality of the available data will determine to what extent the trigger can reliably capture the desired cohort, and operational definitions will often require refinements based on available data. All safety triggers ultimately involve manual medical record reviews to both validate (during trigger development) and act on (during trigger implementation) trigger output. EHR built-in functionality may provide sufficient data access and querying capabilities for e-trigger development where only a few simple criteria are required, but a data warehouse may be required when numerous inclusion and exclusion criteria are present. In addition to access to clinical and administrative data, e-trigger development relies on query software to develop, refine and test algorithms, as well as temporary storage for holding data from identified records.

Both structured and/or unstructured data can be used. Structured, or 'coded', data, such as International Classification of Diseases (ICD) codes and lab results can be used to objectively identify data items. More advanced text mining algorithms, like NLP, can be optionally added to an e-trigger to allow use of the vast amounts of unstructured (ie, free-text) data, particularly when a structured data field for a key criterion does not exist, but the relevant data are contained in progress notes or reports. For example, a structured Breast Imaging Reporting and Data System (BIRADS) code may be helpful in detecting possible cancers on mammography results; however, no analogous coding system is widely used for detecting liver masses on abdominal imaging tests. Instead, an NLP algorithm could scan abdominal imaging result text for radiologist interpretations describing the presence of liver masses.⁵³ While NLP methods are being actively explored, barriers to further deployment include limited access to shared data for comparisons, lack of annotated data sets for training and lack of user-centred development and concerns regarding reproducibility of results in different settings.⁵⁴ Deployment of NLP systems usually requires an expert developer to build algorithms specific to the concepts being queried

Diagnostic Process Dimensions



Process breakdowns lead to diagnostic errors



Organisational Learning for Improving Diagnostic Safety

Prevention of Diagnostic Harm

Figure 1 The Safer Dx e-trigger tools framework. Diagnostic process dimensions adapted from Safer Dx framework.³⁴

Table 2 Safer Dx e-trigger tool development process

e-trigger tool development steps	Stakeholders involved	Example
Identify and prioritise diagnostic error of interest.	Organisational leadership and patient safety personnel	Delays in follow-up of lung nodules identified as a patient safety concern
Operationally define criteria to detect diagnostic error.	Clinicians and staff involved in diagnostic process and patient safety personnel	Trigger development team defines delay as a patient with a lung nodule on a chest imaging, but no repeat imaging or specialty visit within 30 days.
Determine potential data sources.	Informaticists, IT/programmers and data warehouse personnel	Team identifies necessary structured data elements for imaging results and specialty visits within local clinical data warehouse.
Construct e-trigger algorithm.	Clinicians and staff involved in diagnostic process, informaticists, IT/programmers and data warehouse personnel	Programmer develops electronic query based on operational definition of delayed lung nodule follow-up.
Test e-trigger on data source and review medical record.	Clinicians and staff involved in diagnostic process, informaticists, IT/programmers and data warehouse personnel	Triggers are applied to data warehouse and clinicians perform chart reviews on 50 randomly selected records from those identified by the trigger.
Assess e-trigger algorithm performance.	Clinicians and staff involved in diagnostic process, informaticists and patient safety personnel	Positive and negative predictive values, sensitivity and specificity of the trigger are calculated to understand the trigger's performance.
Iteratively refine e-trigger algorithm to improve performance.	Clinicians and staff involved in diagnostic process, informaticists, IT/programmers and patient safety personnel	Trigger development team determines terminal illness to be a major cause of false positive results and adds this to the exclusion criteria.

IT, information technology.

in the free-text data, and are often not easily reused in subsequent projects.⁵⁴ This may put NLP-based triggers beyond current user capabilities, requiring more developer support and limiting wider use. Similarly, unsupervised machine learning, where computers act without being explicitly programmed, can help develop and improve triggers.⁵⁵ Such algorithms could potentially 'learn' to identify patterns in clinical data and make predictions on potential diagnostic errors. However, application of machine learning to make triggers 'smarter' requires more research and development and not ready for widespread implementation.

Step 4: Construct an e-trigger algorithm to obtain cohort of interest

The clinical logic for selecting a cohort of interest can be converted into the necessary query language to extract electronic data. This requires individuals with database and query programming expertise, such as Structured Query Language programming knowledge. Detailed understanding of the clinical event of interest and available data sources are needed to generate patient cohorts for subsequent validation, which requires clinical experts to work closely with the query programmer.

While inclusion criteria will initially identify at-risk patients, a robust set of exclusions is needed to narrow down the population of interest. These exclusions could remove patients in hospice care or those unlikely to have a diagnostic error, such as patients where timely follow-up actions were already performed (eg, imaging test or biopsy done within 30 days) or patients hospitalised electively for procedures rather than unexpected admissions after primary care visits.

The remaining cohort will include an enriched sample of patients with the highest risk for error.

Step 5: Test e-trigger on data source and review medical records

Depending on algorithm complexity, individual inclusion and exclusion criterion should be validated via reviewing a small sample of records. This may isolate potential algorithm or data-related issues (eg, additional ICD codes that need to be considered) not immediately apparent when initially testing the full algorithm. For instance, a small sample of records will reveal if exclusions such as terminal illness, known lung cancer, imaging testing within 30 days and biopsy testing within 30 days indeed were made accurately.

Application of the full e-trigger algorithm yields a list of patients at high risk of diagnostic error ('e-trigger positive' patients). Medical records of e-trigger-positive patients should be reviewed by a clinician to assess for presence or absence of diagnostic error. For instance, when timely follow-up was performed at an outside institution or when the return visit was planned and mentioned only in a free-text portion of a progress note, the record will be false positive. Reviews also help determine whether initial criteria require refinement to increase future predictive value. A review of patients excluded from the cohort ('e-trigger negative' patients) may identify information to help refine the e-trigger (eg, ensuring appropriate data are used for selection and whether additional patient information should be incorporated into the trigger). Directed interviews of involved clinicians (eg, physicians, nurses) and subject matter experts may also yield information to modify criteria.⁴²

Step 6: Assess e-trigger algorithm performance

Several assessment measures can be used to evaluate e-triggers, including positive predictive values (PPV) based on the number of records flagged by the e-trigger tool confirmed as diagnostic error on review (numerator) divided by the total number of records flagged (denominator).⁵⁶ If 'negative' records (ie, those not flagged by the trigger) are reviewed, negative predictive values (NPV; number of patients without the diagnostic error divided by all patients not flagged by the e-trigger), sensitivity and specificity can additionally be calculated. Use of criteria to select higher risk populations will often yield higher PPVs (eg, including lipase orders to identify patients presenting with acute abdominal pain to the emergency department).⁵⁷ Trade-offs will often be needed to achieve the best discrimination of patients of interest from patients without the target or event of interest. e-Triggers with higher PPVs will reduce resources spent on manual confirmatory reviews, while those with higher NPVs will miss fewer records that contain the event of interest. With uncommon events, such as seen in patient safety research, it may only be possible to provide an estimated NPV by reviewing a modest sample of records (eg, 100) because the number of 'e-trigger negative' records that need to be reviewed to find a single event is vast and cost prohibitive. Higher sensitivity may be desirable for certain e-triggers where the importance of all events being captured outweighs the additional review burden introduced by false positives.

The PPV helps plan for human resources to review records and to act on e-trigger output. Clinical personnel would intervene in high-risk situations to prevent harm, whereas patient safety personnel would investigate events and factors that contributed to errors. Process improvement and organisational learning activities would follow. Reviews and actions for missed opportunities to close the loop on abnormal test results will require just a few minutes per patient, allowing a single individual to handle many records per week. However, others related to whether a cognitive error occurred and subsequent investigation and debriefs about what transpired will take much longer.

Step 7: Iteratively refine e-triggers to improve trigger performance

Using the knowledge gained from the previous steps, the e-trigger may be iteratively refined to improve capture of the defined cohort. This may involve simply changing the value of a structured field or potentially redesigning the entire algorithm to better capture the clinical event. Similar to initial development, revision should be informed by content experts and iterative review of the available data. Clinicians can also suggest revisions based on clinical circumstances.

Trigger implementation and use

The ultimate goal of Safer Dx e-trigger development is to improve patient safety through better measurement

and discovery of diagnostic errors by leveraging electronic data. After e-trigger tools are developed and validated to capture the desired cohort of patients with acceptable performance, safety analysis activities and potential solutions can result based on what is learnt.^{58,59} Use of e-triggers as diagnostic safety indicators is promising for identifying historic trends, generating feedback and learning, facilitating understanding of underlying contributory processes and informing improvement strategies. Additionally, certain e-triggers can help health systems intervene to prevent patient harm if applied prospectively.

In addition to having leadership support, health systems will need to either leverage existing or build additional infrastructure necessary to develop and implement diagnostic safety triggers. In organisations with advanced safety programmes, development and implementation will require only modest additional investment of resources; but for others in initial stages, trigger tools could provide a useful starting point. We envision many validated algorithms could be freely shared across institutions to reduce development workload.^{60,61} All health systems will need to convene a multidisciplinary team to harvest knowledge generated by the e-trigger tool. This group should address implementation factors related to how best to use e-trigger results, including who should receive them and how. Prospective application warrants these findings to be communicated to clinicians to take action. Traditional methods of communicating such findings have posed challenges⁶²; thus, additional work to reliably deliver such information is needed. Certain detected events may require further investigation and dedicated patient safety and process improvement teams will need time and resources to collect and analyse data and recommend improvement strategies. Such a group should be composed of clinicians involved in the care processes, informaticists, patient safety professionals, and patients and garner multidisciplinary expertise for understanding data, safety events, and creating and implementing effective solutions. While there is need to invest in additional resources and infrastructure, building such an institutional programme could make significant advances in the quality, accuracy and timeliness of diagnoses.

DISCUSSION

We demonstrate the application of a knowledge discovery framework to guide development and implementation of e-triggers to identify targets for improving diagnostic safety. This approach has shown early promise to identify and describe diagnostic safety concerns within health systems using comprehensive EHRs.^{25–28,35} This discovery approach could ensure progress towards the goal of using the EHR to monitor and improve patient safety, the most advanced and challenging aspect of EHR use.^{8,63}

Application of the Safer Dx e-trigger tool framework is not without limitations. First, a sizeable proportion of

healthcare information is contained in free-text notes or documents. This may necessitate use of NLP-based methods if e-trigger performance is inadequate to detect the cohort of interest, but NLP requires additional expertise and methods to improve portability, and accessibility of NLP tools is still being explored.^{64 65} Use of statistical models to estimate the likelihood of a diagnostic error or machine learning^{55 66} to program a computer to ‘learn’ from data patterns and make subsequent predictions may allow subsequent improvements in performance. Maturation of these techniques will stimulate the development and use of more sophisticated and effective e-trigger tools. Second, data availability and quality remain important issues that impact trigger feasibility and performance. Even at organisations that provide comprehensive and longitudinal care, we have found data sharing across institutions to be incomplete, requiring deliberate processes to actively collect and record external findings.⁶⁰ This highlights the need for more meaningful sharing of data across institutions in a manner computers can use. Efforts to improve data sharing are already under way, but in early states (eg, view-only versions of data from external organisations). As data sharing improves, e-trigger tools will have better opportunities to impact patient safety.⁶⁷ Furthermore, even when all care is delivered within a single organisation, absent, incomplete, outdated or incorrect data can affect trigger tool performance. Similarly, certain elements of patients’ histories, exams or assessments may not be recorded in the medical record, limiting both e-trigger performance and subsequent chart reviews used to verify trigger results.^{68 69} However, this is a limitation of most current safety measurement methods.

CONCLUSION

Use of HIT and readily available electronic clinical data can enable better patient safety measurement. The Safer Dx Trigger Tools Framework discussed here has potential to advance both real-time and retrospective identification of opportunities to improve diagnostic safety. Development and implementation of diagnostic safety e-trigger tools along with institutional investments to do so can improve our knowledge on reducing harm from diagnostic errors and accelerate progress in patient safety improvement.

Contributors All authors contributed to the development, review and revision of this manuscript.

Funding Work described is heavily drawn from research funded by the Veteran Affairs Health Services Research and Development Service CREATE grant (CRE-12-033), the Agency for Healthcare Research and Quality (R18HS017820) and the Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety (CIN 13-413). Dr Murphy is additionally funded by an Agency for Healthcare Research & Quality Mentored Career Development Award (K08-HS022901), and Dr Singh is additionally supported by the VA Health Services Research and Development Service (Presidential Early Career Award for Scientists and Engineers USA 14-274), the VA National Center for Patient Safety, the

Agency for Health Care Research and Quality (R01HS022087) and the Gordon and Betty Moore Foundation. Drs Sittig and Thomas are supported in part by the Agency for Health Care Research and Quality (P30HS023526). These funding sources had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; and preparation, review or approval of the manuscript.

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

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

REFERENCES

- 1 Kohn LT, Corrigan JM, Donaldson MS, eds. Committee on quality of health care in America, Institute of Medicine. *To Err is human: building a safer health system*. Washington, DC: The National Academies Press, 2000.
- 2 Nepple KG, Joudi FN, Hillis SL, *et al*. Prevalence of delayed clinician response to elevated prostate-specific antigen values. *Mayo Clin Proc* 2008;83:439–45.
- 3 Singh H, Meyer AN, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014;23:727–31.
- 4 *Improving diagnostic quality and safety*. Washington, DC: National Quality Forum, 2017.
- 5 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;27:557–66.
- 6 Dhaliwal G, Shojania KG. The data of diagnostic error: big, large and small. *BMJ Qual Saf* 2018;27:499–501.
- 7 Lorincz C, Drazen E, Sokol P. *Research in ambulatory patient safety 2000–2010: a 10-year review*. Chicago IL: American Medical Association, 2011.
- 8 Sittig DF, Singh H. Electronic health records and national patient-safety goals. *N Engl J Med* 2012;367:1854–60.
- 9 Howard IL, Bowen JM, Al Shaikh LAH, *et al*. Development of a trigger tool to identify adverse events and harm in Emergency Medical Services. *Emerg Med J* 2017;34:391–7.
- 10 De Almeida SM, Romualdo A, De Abreu Ferraresi A, *et al*. Use of a trigger tool to detect adverse drug reactions in an emergency department. *BMC Pharmacol Toxicol* 2017;18:71.
- 11 Unbeck M, Lindemalm S, Nydert P, *et al*. Validation of triggers and development of a pediatric trigger tool to identify adverse events. *BMC Health Serv Res* 2014;14:655.
- 12 Lipitz-Snyderman A, Classen D, Pfister D, *et al*. Performance of a trigger tool for identifying adverse events in oncology. *J Oncol Pract* 2017;13:e223–e230.
- 13 Lindblad M, Schildmeijer K, Nilsson L, *et al*. Development of a trigger tool to identify adverse events and no-harm incidents that affect patients admitted to home healthcare. *BMJ Qual Saf* 2018;27:502–511.
- 14 Sammer C, Miller S, Jones C, *et al*. Developing and evaluating an automated all-cause harm trigger system. *Jt Comm J Qual Patient Saf* 2017;43:155–65.

- 15 Menendez ME, Janssen SJ, Ring D. Electronic health record-based triggers to detect adverse events after outpatient orthopaedic surgery. *BMJ Qual Saf* 2016;25:25–30.
- 16 *Triggers and Targeted Injury Detection Systems (TIDS) Expert Panel Meeting: conference summary report*. Rockville, MD: Agency for Healthcare Research and Quality, 2009.
- 17 Glossaries. *AHRQ Patient safety network*, 2005.
- 18 Clancy CM. Common formats allow uniform collection and reporting of patient safety data by patient safety organizations. *Am J Med Qual* 2010;25:73–5.
- 19 Shenvi EC, El-Kareh R. Clinical criteria to screen for inpatient diagnostic errors: a scoping review. *Diagnosis* 2015;2:3–19.
- 20 Classen DC, Resar R, Griffin F, *et al*. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff* 2011;30:581–9.
- 21 Asgari H, Esfahani SS, Yaghoubi M, *et al*. Investigating selected patient safety indicators using medical records data. *J Educ Health Promot* 2015;4:54.
- 22 Kane-Gill SL, MacLasco AM, Saul MI, *et al*. Use of text searching for trigger words in medical records to identify adverse drug reactions within an intensive care unit discharge summary. *Appl Clin Inform* 2016;7:660–71.
- 23 Classen DC, Pestotnik SL, Evans RS. Computerized surveillance of adverse drug events in hospital patients. *JAMA* 1991;266:2847–51.
- 24 Doupi R, Svaar H, Bjørn B, *et al*. Use of the global trigger tool in patient safety improvement efforts: nordic experiences. *Cogn Technol Work* 2015;17:45–54.
- 25 Murphy DR, Meyer AN, Vaghani V, *et al*. Application of electronic algorithms to improve diagnostic evaluation for bladder cancer. *Appl Clin Inform* 2017;8:279–90.
- 26 Murphy DR, Wu L, Thomas EJ, *et al*. Electronic trigger-based intervention to reduce delays in diagnostic evaluation for cancer: a cluster randomized controlled trial. *J Clin Oncol* 2015;33:3560–7.
- 27 Murphy DR, Laxmisan A, Reis BA, *et al*. Electronic health record-based triggers to detect potential delays in cancer diagnosis. *BMJ Qual Saf* 2014;23:8–16.
- 28 Murphy DR, Thomas EJ, Meyer AN, *et al*. Development and validation of electronic health record-based triggers to detect delays in follow-up of abnormal lung imaging findings. *Radiology* 2015;277:81–7.
- 29 Danforth KN, Smith AE, Loo RK, *et al*. Electronic clinical surveillance to improve outpatient care: diverse applications within an integrated delivery system. *EGEMS* 2014;2:1056.
- 30 Wandtke B, Gallagher S. Reducing delay in diagnosis: multistage recommendation tracking. *AJR Am J Roentgenol* 2017;209:970–5.
- 31 Singh H, Thomas EJ, Khan MM, *et al*. Identifying diagnostic errors in primary care using an electronic screening algorithm. *Arch Intern Med* 2007;167:302–8.
- 32 Resar RK, Rozich JD, Simmonds T, *et al*. A trigger tool to identify adverse events in the intensive care unit. *Jt Comm J Qual Patient Saf* 2006;32:585–90.
- 33 Bhise V, Sittig DF, Vaghani V, *et al*. An electronic trigger based on care escalation to identify preventable adverse events in hospitalised patients. *BMJ Qual Saf* 2018;27:241–246.
- 34 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.
- 35 Murphy DR, Meyer AN, Bhise V, *et al*. Computerized triggers of big data to detect delays in follow-up of chest imaging results. *Chest* 2016;150:613–20.
- 36 Sim JJ, Rutkowski MP, Selevan DC, *et al*. Kaiser permanente creatinine safety program: a mechanism to ensure widespread detection and care for chronic kidney disease. *Am J Med* 2015;128:1204–11.
- 37 Meyer AND, Murphy DR, Al-Mutairi A, *et al*. Electronic detection of delayed test result follow-up in patients with hypothyroidism. *J Gen Intern Med* 2017;32:753–9.
- 38 Hedriana HL, Wiesner S, Downs BG, *et al*. Baseline assessment of a hospital-specific early warning trigger system for reducing maternal morbidity. *Int J Gynaecol Obstet* 2016;132:337–41.
- 39 Shields LE, Wiesner S, Klein C, *et al*. Use of maternal early warning trigger tool reduces maternal morbidity. *Am J Obstet Gynecol* 2016;214:527.e1–527.e6.
- 40 Singh H. Improving diagnostic safety in primary care by unlocking digital data. *Jt Comm J Qual Patient Saf* 2017;43:29–31.
- 41 Bhise V, Meyer AND, Singh H, *et al*. Errors in diagnosis of spinal epidural abscesses in the era of electronic health records. *Am J Med* 2017;130:975–81.
- 42 Hripcsak G, Bakken S, Stetson PD, *et al*. Mining complex clinical data for patient safety research: a framework for event discovery. *J Biomed Inform* 2003;36(1-2):120–30.
- 43 Singh H, Giardina TD, Forjuoh SN, *et al*. Electronic health record-based surveillance of diagnostic errors in primary care. *BMJ Qual Saf* 2012;21:93–100.
- 44 Olson APJ, Graber ML, Singh H. Tracking progress in improving diagnosis: a framework for defining undesirable diagnostic events. *J Gen Intern Med* 2018;33:1187–91.
- 45 Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005;165:1493–9.
- 46 Singh H, Hirani K, Kadiyala H, *et al*. Characteristics and predictors of missed opportunities in lung cancer diagnosis: an electronic health record-based study. *J Clin Oncol* 2010;28:3307–15.
- 47 Zwaan L, Thijs A, Wagner C, *et al*. Relating faults in diagnostic reasoning with diagnostic errors and patient harm. *Acad Med* 2012;87:149–56.
- 48 Zwaan L, Schiff GD, Singh H. Advancing the research agenda for diagnostic error reduction. *BMJ Qual Saf* 2013;22(Suppl 2):ii52–ii57.
- 49 Berlin L. Malpractice issues in radiology admitting mistakes. *AJR Am J Roentgenol* 1999;172:879–84.
- 50 Gale BD, Bissett-Siegel DR, Davidson SJ, *et al*. Failure to notify reportable test results: significance in medical malpractice. *J Am Coll Radiol* 2011;8:776–9.
- 51 Gandhi TK, Kachalia A, Thomas EJ, *et al*. Missed and delayed diagnoses in the ambulatory setting: a study of closed malpractice claims. *Ann Intern Med* 2006;145:488–96.
- 52 Brender J, Ammenwerth E, Nykänen P, *et al*. Factors influencing success and failure of health informatics systems—a pilot Delphi study. *Methods Inf Med* 2006;45:125–36.
- 53 Danforth KN, Early MI, Ngan S, *et al*. Automated identification of patients with pulmonary nodules in an integrated health system using administrative health plan data, radiology reports, and natural language processing. *J Thorac Oncol* 2012;7:1257–62.
- 54 Chapman WW, Nadkarni PM, Hirschman L, *et al*. Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions. *J Am Med Inform Assoc* 2011;18:540–3.
- 55 Deo RC. Machine learning in medicine. *Circulation* 2015;132:1920–30.

- 56 Bramer M. *Principles of data mining*. London: Springer London, 2013.
- 57 Medford-Davis L, Park E, Shlamovitz G, *et al*. Diagnostic errors related to acute abdominal pain in the emergency department. *Emerg Med J* 2016;33:253–9.
- 58 Rosen AK, Mull HJ, Kaafarani H, *et al*. Applying trigger tools to detect adverse events associated with outpatient surgery. *J Patient Saf* 2011;7:45–59.
- 59 Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;12:39–45.
- 60 Murphy DR, Meyer AND, Vaghani V, *et al*. Electronic triggers to identify delays in follow-up of mammography: harnessing the power of big data in health care. *J Am Coll Radiol* 2018;15:287–95.
- 61 Murphy DR, Meyer AND, Vaghani V, *et al*. Development and validation of trigger algorithms to identify delays in diagnostic evaluation of gastroenterological cancer. *Clin Gastroenterol Hepatol* 2018;16:90–8.
- 62 Meyer AN, Murphy DR, Singh H. Communicating findings of delayed diagnostic evaluation to primary care providers. *J Am Board Fam Med* 2016;29:469–73.
- 63 Meeks DW, Takian A, Sittig DF, *et al*. Exploring the sociotechnical intersection of patient safety and electronic health record implementation. *J Am Med Inform Assoc* 2014;21(e1):e28–e34.
- 64 Soundrarajan BR, Ginter T, DuVall SL. An interface for rapid natural language processing development in UIMA Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Systems Demonstrations. *Association for Computational Linguistics* 2011:139–44.
- 65 Divita G, Carter ME, Tran LT, *et al*. v3NLP Framework: tools to build applications for extracting concepts from clinical text. *EGEMS* 2016;4:1228.
- 66 Wu J, Roy J, Stewart WF. Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. *Med Care* 2010;48(6 Suppl):S106–13.
- 67 Russo E, Sittig DF, Murphy DR, *et al*. Challenges in patient safety improvement research in the era of electronic health records. *Healthc* 2016;4:285–90.
- 68 Schwartz A, Weiner SJ, Weaver F, *et al*. Uncharted territory: measuring costs of diagnostic errors outside the medical record. *BMJ Qual Saf* 2012;21:918–24.
- 69 Weiner SJ, Schwartz A. Directly observed care: can unannounced standardized patients address a gap in performance measurement? *J Gen Intern Med* 2014;29:1183–7.