



Electronic health record-based triggers to detect potential delays in cancer diagnosis

Daniel R Murphy,^{1,2} Archana Laxmisan,^{1,2} Brian A Reis,^{1,2} Eric J Thomas,³ Adol Esquivel,⁴ Samuel N Forjuoh,⁵ Rohan Parikh,⁶ Myrna M Khan,^{1,2} Hardeep Singh^{1,2}

For numbered affiliations see end of article.

Correspondence to

Dr Daniel R Murphy, Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC), HSR&D Center of Excellence (152), 2002 Holcombe Boulevard, Houston, TX 77030 USA; drmurphy@bcm.edu

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 US government.

Received 2 February 2013
Revised 20 June 2013
Accepted 21 June 2013
Published Online First 19 July 2013



► <http://dx.doi.org/10.1136/bmjqs-2013-002426>

To cite: Murphy DR, Laxmisan A, Reis BA, et al. *BMJ Qual Saf* 2014;**23**:8–16.

ABSTRACT

Background Delayed diagnosis of cancer can lead to patient harm, and strategies are needed to proactively and efficiently detect such delays in care. We aimed to develop and evaluate ‘trigger’ algorithms to electronically flag medical records of patients with potential delays in prostate and colorectal cancer (CRC) diagnosis.

Methods We mined retrospective data from two large integrated health systems with comprehensive electronic health records (EHR) to iteratively develop triggers. Data mining algorithms identified all patient records with specific demographics and a lack of appropriate and timely follow-up actions on four diagnostic clues that were newly documented in the EHR: abnormal prostate-specific antigen (PSA), positive faecal occult blood test (FOBT), iron-deficiency anaemia (IDA), and haematochezia. Triggers subsequently excluded patients not needing follow-up (eg, terminal illness) or who had already received appropriate and timely care. Each of the four final triggers was applied to a test cohort, and chart reviews of randomly selected records identified by the triggers were used to calculate positive predictive values (PPV).

Results The PSA trigger was applied to records of 292 587 patients seen between 1 January 2009 and 31 December 2009, and the CRC triggers were applied to 291 773 patients seen between 1 March 2009 and 28 February 2010. Overall, 1564 trigger positive patients were identified (426 PSA, 355 FOBT, 610 IDA and 173 haematochezia). Record reviews revealed PPVs of 70.2%, 66.7%, 67.5%, and 58.3% for the PSA, FOBT, IDA and haematochezia triggers, respectively. Use of all four triggers at the study sites could detect an estimated 1048 instances of delayed or missed follow-up of abnormal findings annually and 47 high-grade cancers.

Conclusions EHR-based triggers can be used successfully to flag patient records lacking

follow-up of abnormal clinical findings suspicious for cancer.

BACKGROUND

Identifying and preventing delays in cancer diagnosis have proved elusive and challenging to overcome.^{1 2} For certain cancers, delays are common and lead to poor outcomes and increased malpractice litigation.^{3–8} While root causes of such delays are multifactorial,^{2 9–11} many delays arise when abnormal cancer screening results or other ‘red flags’ are missed by providers.^{3 5 12–21} These missed opportunities can result in delays in diagnosis and treatment, and reduce the chances of early, potentially curative therapy.²² While the root causes of these delays are still being uncovered and addressed, there is a pressing need to intervene to prevent harm. However, detection of diagnostic delays across the fragmented, visit-based nature of ambulatory care is challenging, and tracking all patients with suspected cancer across the diagnosis continuum is inefficient and cost-prohibitive.

Comprehensive electronic health records (EHR) that contain data across the longitudinal continuum of care and facilitate data mining²³ make detection of diagnostic delays potentially possible. However, use of simple search tools to identify patients with positive test results (eg, ‘all patients with a positive faecal occult blood test’ (FOBT)) will likely meet resistance from providers who are already overloaded by the amount of data they receive each day through the EHR,^{24–26} and require a large amount of time expended on frequent false positive results. Thus, novel methods are needed

to create a back-up system to detect delays efficiently. One potential method is the use of 'trigger' tools^{27–31} to identify specific patterns within clinical data so that only a selective set of records is targeted for confirmatory review. Triggers are defined as a specific set of clues used to flag records of patients at higher risk of harm, so that they can be reviewed for possible safety events.³² Thus far, they have been used primarily to retrospectively identify errors of commission, such as those related to adverse drug events and nosocomial infections.^{33–39} Although we have applied triggers to detect diagnostic errors,⁴⁰ their use in the outpatient setting remains limited.

Lack of follow-up of 'red flags' or 'alarm' features⁴¹ of cancer, such as abnormal clinical findings (eg, haematochezia, also referred to as bright red blood per rectum) or test results (eg, positive FOBT or iron deficiency anaemia (IDA)), offers an opportunity to design such triggers.⁴² For instance, almost a third of patients with colorectal cancer (CRC) can experience a missed opportunity in the initiation of colonoscopy referral, leading to diagnostic delays.⁵ Similarly, lack of follow-up on abnormal prostate-specific antigen (PSA) results is not uncommon.^{13 16 43} Although controversy exists regarding guidelines for PSA utilisation,^{44 45} and PSA testing and follow-up have become increasingly dependent on patient characteristics and preferences, follow-up of abnormal results, including routine monitoring with repeat testing, persists as a problem. Our study objective was to design and evaluate EHR-based triggers to identify diagnostic delays related to lack of follow-up of key alarm features of prostate and CRC diagnosis.⁴¹ Successful development of such triggers could help create a future back-up system to prospectively detect patients at risk of prolonged delays in cancer diagnosis.

DESIGN

Setting

We used EHR data repositories from two large geographically disparate healthcare systems in the USA to develop four electronic trigger algorithms (one for prostate and three for CRC) that identify potential delays in cancer diagnosis. Both healthcare systems provide inpatient and outpatient care, and each has used a different comprehensive EHR for over a decade. Each institution possessed a data warehouse (data repository) where trigger queries were conducted. Institutional review boards at both institutions approved the study.

Trigger development

We developed triggers with retrospective data using a framework for mining complex clinical data for patient safety research⁴⁶ and employed an iterative process of development and testing. We first performed literature reviews and obtained expert opinions from primary care physicians and specialists to identify *a priori* criteria for each trigger in four clinical categories: (1) 'demographic'

criteria, (2) 'red flag' criteria, that is, presence of diagnostic clues related to prespecified abnormal findings or test results, (3) 'clinical exclusion' criteria that would deem follow-up as unnecessary (eg, terminal illness or known prostate cancer or CRC) and (4) 'expected follow-up' criteria used to exclude patients who had already received appropriate and timely follow-up (see table 1 for operational definitions of all criteria). Criteria available in structured fields (ie, not within any narrative free-text formats, such as progress notes) were then translated into computer logic and incorporated into a preliminary trigger for application to the EHR data repositories. The trigger was designed to evaluate each criterion in a step-wise manner, first by collecting all patients with appropriate demographic and red flag criteria, then excluding those with clinic exclusion and expected follow-up criteria. Since no standard definition of a 'delay' currently exists, timeliness of follow-up action was determined by expert opinion through discussions with primary care providers and specialists who considered the relative urgency of a potential colon or prostate cancer diagnosis, as well as a Veterans Affairs directive⁴⁷ that recommended a 60-day window for colonoscopy performance, and a UK National Health Service recommendation that action be taken within 62 days of a suspected cancer diagnosis.⁴⁸ Diagnoses, procedures and visits types were identified using all relevant International Classification of Diseases, Ninth Revision (ICD-9) and Current Procedural Terminology (CPT) codes.

For each criterion (see table 1 for a list of all finalised trigger criteria), we used data mining algorithms to separate all patients seen at either facility during a prespecified 1-year period (1 January to 31 December 2008) into two groups, those that met the criterion and those that did not. From each group, 10 records were randomly chosen for review, and a team of clinicians determined whether the algorithm appropriately evaluated all pertinent data (eg, ICD-9/CPT codes and order status) to indicate whether a diagnosis was made, procedure was performed, or event occurred, while minimising false positive records (eg, we found that prostatitis was frequently represented with the ambiguous 'urinary tract infection' ICD-9 code; thus, this code was not used as a criterion). Based on the findings from these reviews and clinical knowledge, the programming logic used was iteratively modified as necessary (eg, CPT codes for colectomy were expanded to include all codes between 44150 and 44158). This was followed by extraction and review of additional records as needed. Criteria were incorporated into each of the four triggers, and all final trigger algorithms were reviewed by primary care providers and specialists prior to testing of performance.

Trigger performance

We applied the final trigger algorithms to cohorts of all patients with a visit to their respective healthcare system during a 1-year period ('test cohorts'). We

Table 1 Criteria used for each trigger algorithm

	Colorectal cancer triggers			
	Prostate cancer trigger	Faecal occult blood test (FOBT)	Iron deficiency anaemia (IDA)	Haematochezia
Demographic criteria	Age 40–70 years Alive* 90 days after the PSA result	Age 40–75 years Alive* 60 days after positive FOBT	Age 40–75 years Alive 60 days after Haemoglobin result*	Age 40–75 years Alive* 60 days after haematochezia diagnosis
Red flag criteria	PSA result of 4.1–15 ng/mL No PSA \geq 4.1 ng/mL in prior 2 years	Positive FOBT test	Lab results consistent with IDA: ▶ Haemoglobin \leq 11 g/dL, ▶ Mean corpuscular volume \leq 81 fL, and ▶ No ferritin \geq 100 ng/mL in prior 12 months	New diagnosis of haematochezia
Clinical exclusion criteria	Prior history of prostate cancer Prior prostate biopsy within prior 2 years Prostatitis diagnosis within 30 days prior to or 90 days after red flag criteria met Terminal illness diagnosis Hospice/palliative care enrolment	Colonoscopy within 5 years prior to red flag criteria met Prior history of colorectal cancer Prior total colectomy Terminal illness diagnosis Hospice/palliative care enrollment Upper GI bleed within 1 year before or 60 days after red flag criteria met	Colonoscopy within 3 years prior to red flag criteria met Prior thalassaemia diagnosis (ICD-9 code) Hospitalisation within 14 days prior to red flag criteria met Surgery within 14 days prior to red flag criteria met Menorrhagia diagnosis within 5 years prior or 60 days after red flag criteria met Other cause of bleeding within 1 year prior or 60 days after red flag criteria met Pregnant within 1 year prior to or 60 days after red flag criteria met Prior history of colorectal cancer Prior total colectomy Terminal illness diagnosis Hospice/palliative care enrollment Upper GI bleed within 1 year before or 60 days after red flag criteria met	Colonoscopy within 3 years prior to red flag criteria met Prior history of colorectal cancer Prior total colectomy Terminal illness diagnosis Hospice/palliative care enrollment Upper GI bleed within 1 year before or 60 days after red flag criteria met
Expected follow-up criteria	Any of the following within 90 days after red flag criteria met: ▶ Repeat PSA performed ▶ Urology consult requested or completed ▶ Prostate biopsy performed	Colonoscopy performed within 60 days after red flag criteria met	Colonoscopy performed within 60 days after red flag criteria met	Colonoscopy performed within 60 days after red flag criteria met

*As identified from social security number-based death index.

FOBT, faecal occult blood test; GI, gastrointestinal; IDA, iron deficiency anaemia; PSA, prostate specific antigen.

randomly selected trigger-positive records (ie, records identified by the trigger as having a high risk for missed or delayed diagnosis) to determine each trigger's positive predictive value (PPV) which was defined as the number of records identified by the trigger lacking follow-up divided by the total number of records identified by the trigger. Based on prior research on diagnostic triggers, we expected these triggers might be able to achieve a PPV of 35%.⁴⁰ In real-world application, each chart identified by the trigger would require further review, consuming providers' already limited time. For triggers to be practical, they must achieve a PPV of at least this percentage, and we therefore powered the study accordingly. A minimum sample size of 60 records per trigger at each site was calculated as sufficient to identify a rate of at least 35%, with a power of at least 90%, and a two-sided α of 0.05.

We trained three reviewers and developed standardised data collection instruments. Each reviewer was initially required to independently review 20 charts. Once a 90% inter-rater agreement with other reviewers was achieved, reviewers would proceed to perform additional reviews for the study. Reviewers were provided with operational definitions of all trigger criteria, and cases where uncertainty existed were decided by consensus between reviewers. Reviewers used all available data within the EHR to evaluate two outcomes: (1) whether the trigger correctly identified the structured data criteria (evaluation of trigger internal validity) and (2) whether the patient truly had a delay in follow-up (evaluation of trigger performance related to 'true positives'). Reviewers also collected information on time to documented follow-up, and whether justification for lack of follow-up was documented (eg, progress note detailing that a patient declined follow-up). Data were also collected to determine whether patients with potential delays were subsequently diagnosed with a cancer or precancerous lesion. Because 2 years had elapsed since the end of the study period, a 2-year follow-up period was chosen as a uniform cut-off for which to assess cancer outcomes for each trigger. If any patients had failed to receive follow-up by the time of chart review, we informed the respective providers.

Data analysis

Data were analysed using Excel (Microsoft, Redmond, Washington, USA) and SAS V.9.2 (SAS Institute, Cary, North Carolina, USA). Trigger performance, time to follow-up, reasons for lack of follow-up, and cancer outcomes were reported using descriptive statistics.

RESULTS

Trigger development

Four trigger algorithms designed to identify patients at high risk for delayed prostate cancer and CRC

diagnosis were developed after iterative review of 214 records (88 prostate cancer and 126 CRC). The final criteria for each trigger are shown in table 1. Based on prior work, we defined abnormal PSA as between 4.0 and 15.0 ng/mL. We found that including levels above 15 ng/mL only reduced trigger specificity due to a high likelihood of follow-up from more robust provider-alerting protocols in the EHRs under study, and because most of these patients already had known cancer.²¹ We were able to exclude most 'clinical exclusion' criteria, such as known diagnosis of prostate cancer, prostatitis, terminal illness, or recent prostate biopsy. We defined expected follow-up as a repeat PSA, prostate biopsy performance, or urology referral placed within 90 days after the diagnostic clue. Thus, the final output of the trigger algorithm included only patients at high risk for prostate cancer diagnosis without any evidence of appropriate follow-up care.

Each CRC trigger shared some 'clinical exclusion' criteria (eg, terminal illness, known CRC) and used some unique trigger-specific criteria (eg, thalassaemia for the IDA trigger) to exclude patients where follow-up was not needed. Expected follow-up was defined as a colonoscopy performed within 60 days after the red flag diagnostic clue.

Trigger application

The final PSA trigger was applied to the records of 292 587 patients who visited their respective facilities between 1 January and 31 December 2009 (30.5% from site 1 and 69.5% from site 2). A total of 1082 (0.4%) records had demographic and red flag data criteria, of which 168 records were excluded due to clinical exclusion criteria and 488 were excluded because of expected follow-up criteria. The remaining 426 flagged records (0.15% of all patients seen, 39.4% of charts with demographic and red flag criteria) were deemed as high risk for delayed diagnosis ('trigger-positive') and were manually reviewed for confirmation (figure 1).

During development of the CRC triggers, we discovered that one healthcare system used the haematochezia ICD-9 code interchangeably with 'melena,' preventing us from distinguishing lower from upper gastrointestinal bleeding. Thus, the haematochezia trigger was tested only at one site. The final FOBT and IDA triggers were applied to 291 773 patients seen at both sites between 1 March 2009 and 28 February 2010 (30.6% from site 1 and 69.4% from site 2), while the haematochezia trigger was applied to 202 553 records at one site. Overall, 3246 records met demographic and red flag criteria: 516 (0.2% of patients seen) FOBT, 1753 (0.6%) IDA, and 977 (0.5%) haematochezia. From these, 1812 records were excluded based on clinical exclusion criteria, and 296 records were excluded because they contained expected follow-up criteria (figure 1). This resulted in 1138 records identified as high risk for missed

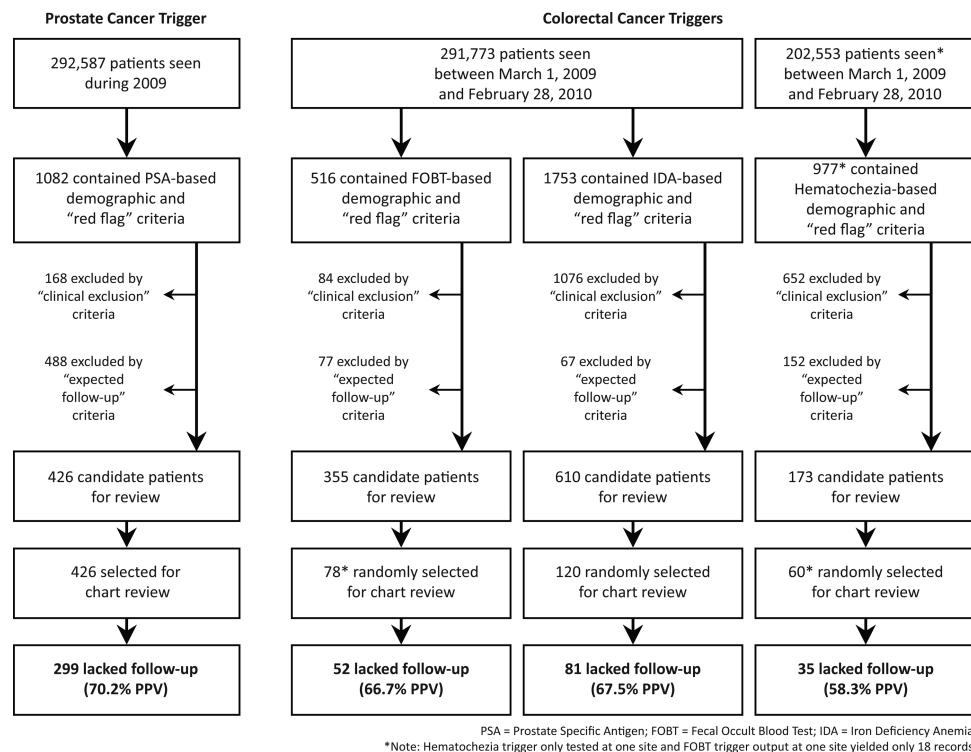


Figure 1 Overview of study results.

follow-up: 355 (0.1% of all patients seen) FOBT, 610 (0.2%) IDA, and 173 (0.1%; one site only) haematochezia.

Trigger performance

Because we found that elevated PSA results were commonly attributed to benign prostatic hypertrophy (BPH),

we chose to review the full sample in order to determine meaningful 2-year outcomes (cancerous/precancerous lesions vs BPH) for potential lack of follow-up. Thus, all 426 records identified to have a potential delay were reviewed, of which 299 (70.2%; 95% CI 65.7% to 74.3%) truly lacked expected follow-up. In 63.5% of these, no documented reason for delaying follow-up was

Table 2 Reasons for missed or delayed prostate cancer follow-up as ascertained by record review

Reason	PSA-based trigger n=426 Total records # (%)
Justification documented (false positives)	
Patient declined further prostate-related care	27 (21.3)
Patient pursued outside care	22 (17.3)
Unable to reach patient despite multiple phone calls/mailings	5 (3.9)
Record identified by trigger despite not meeting red flag criteria (data miscoded)	0 (0.0)
Patient identified by trigger despite meeting exclusion or expected follow-up criteria	73 (57.5)
Treated for prostatitis	41
Prior recent biopsy at outside institution	19
Appropriate care performed	10
Known history of prostate cancer	2
Terminal illness/hospice care	1
Total with justification documented	127 (29.8)
No justification documented (true positives)	
No reason/rationale documented	190 (63.5)
Appropriate care ordered but not performed within 90 days (eg, PSA repeated beyond 90 days)	109 (36.5)
Total without justification documented (PPV)	299 (70.2)

PPV, positive predictive value; PSA, prostate specific antigen.

Table 3 Reasons for missed or delayed colorectal cancer follow-up as ascertained by record review

Colorectal cancer triggers	FOBT-based trigger	IDA-based trigger	Haematochezia-based trigger
	n=78	n=120	n=60
Reason	Total records # (%)	Total records # (%)	Total records # (%)
Justification documented (false positives)			
Patient declined colonoscopy	3 (11.5)	2 (5.1)	7 (28.0)
Patient pursued outside care	2 (7.7)	5 (12.8)	1 (4.0)
Record identified by trigger despite not meeting Red flag criteria (data miscoded)	0 (0.0)	0 (0.0)	13 (52.0)
Patient identified by trigger despite meeting exclusion or expected follow-up criteria	21 (80.8)	32 (82.1)	4 (16.0)
Upper GI bleed identified as bleeding source	4	17	0
Prior colonoscopy performed at outside institution	10	8	2
Patient hospitalised at time of red flag diagnostic clue	0	6	0
Terminal illness/hospice care	2	0	2
Other	5	1	0
Total with justification documented	26 (33.3)	39 (32.5)	25 (41.7)
No justification documented (true positives)			
No reason/rationale documented	15 (28.8)	72 (88.9)	18 (51.4)
Colonoscopy ordered by not performed by 60 days	34 (65.4)	4 (4.9)	12 (34.3)
Inadequate follow-up (eg, subsequent FOBT only)	3 (5.8)	5 (6.2)	5 (14.3)
Total without justification documented (PPV)	52 (66.7)	81 (67.5)	35 (58.3)

FOBT, faecal occult blood test; GI, gastrointestinal; IDA, iron deficiency anaemia; PPV, positive predictive value.

provided even though providers acknowledged PSA elevation in their notes in 55 (29%) of these records (eg, providers documented the PSA value, but did not include a differential or alternate diagnosis, reason not to pursue action, or follow-up plan; table 2 describes reasons for delayed follow-up in the remaining cases). About a third (29.8%) of the charts reviewed were inappropriately flagged by the trigger, most often because information regarding 'clinical exclusion' or 'expected follow-up' criteria existed in progress note narrative, but not in searchable structured fields. This most frequently occurred when a patient was treated for prostatitis, but a non-specific ICD-9 code, such as for 'elevated PSA' or 'urinary tract infection' was used as a diagnosis field rather than 'prostatitis.'

Of 299 records with delayed or missed follow-up of abnormal findings, 33 (11.0%) were found to have a diagnosis of prostate cancer or focal high-grade neoplasia at 2 years. Of these, 21 (63.6%; median follow-up of 160 days) had a Gleason score of at least 7, or disease that had spread beyond the prostate capsule (TNM stage \geq T3).

The FOBT trigger identified only 18 records at one of the sites, and the haematochezia trigger could not be run at one site; thus, 258 records (78 FOBT, 120 IDA and 60 haematochezia) were reviewed (figure 1). Reviewers identified 52 (66.7%; 95% CI 55.6% to 76.2%) records from the FOBT trigger, 81 (67.5%; 95% CI 58.7% to 75.2%) from the IDA trigger, and 35 (58.3%; 95% CI 45.7% to 70.0%) from the

haematochezia trigger as lacking follow-up. Among these, no rationale was documented for delayed follow-up care in 89% of IDA and 51% of haematochezia triggers; therefore, it is not clear whether providers missed this information or were aware but chose to delay follow-up for some specific reason. Delayed care in the FOBT trigger most often (65.4%) occurred when expected follow-up was ordered, but not performed within 60 days (table 3). In the haematochezia trigger, 52.0% of inappropriately flagged records resulted from melena being coded as haematochezia. Conversely, most inappropriately flagged records from the FOBT (80.8%) and IDA (82.1%) triggers resulted from clinical exclusion criteria documented in free-text progress notes but not in a structured format accessible to the trigger (eg, care received outside the system documented in a progress note).

Of the 168 patient charts reviewed with delayed care, 6 (3.6%; median time to follow-up of 74 days) were eventually identified to have a diagnosis of CRC by 2 years after the red flag diagnostic clue. An additional 15 (8.9%) were found to have at least one pre-neoplastic polyp removed during a colonoscopy. Using findings from this sample to estimate the outcomes of all 1138 charts identified by the CRC triggers, we would expect the triggers to enable earlier identification of approximately 749 instances of delayed or missed follow-up of abnormal findings, and 26 delayed CRC diagnoses by 2 years.

DISCUSSION

We developed, applied and evaluated four electronic triggers to search large EHR repositories for patients at high risk for delayed diagnosis of prostate and CRC. Each of the triggers achieved a PPV between 58% and 70%, and together would allow detection of an estimated 1048 instances of delayed or missed follow-up of abnormal findings and 47 high-grade cancers annually at the study facilities. Because there are no current methods to harness electronic data to identify these types of delays, our trigger-based methods are more efficient than non-selective chart reviews. We anticipate that their future use will involve prospective and proactive application to detect delays in care. Within the entire delayed or missed follow-up group, 11.6% patients were subsequently diagnosed with either cancerous or precancerous lesions. Thus, these triggers could potentially allow appropriate action to be taken earlier in the disease progression continuum.

With the growing adoption of EHRs,^{49 50} triggers are now being increasingly proposed for mining data repositories to identify patients with (or at risk for) adverse outcomes.^{29 31 38} However, novel approaches are needed to develop algorithms that account for multiple levels of inclusion and exclusion criteria and achieve a reasonable PPV. We also carefully considered definitions of delays to balance the trade-off between enabling providers to intervene before disease progresses to a more advanced stage, versus overloading them with information by unnecessarily alerting them before a delay in care has actually occurred. This type of rationale could reduce the burden on providers' time and effort if trigger information were being sent to them directly for review.

A recent American Medical Association report outlined several key areas for improving patient safety in the ambulatory setting, including improving follow-up of abnormal test results.⁵¹ Although developed and tested on retrospective data, such triggers could be used prospectively as a back-up system to inform providers about their patients who have not yet received appropriate follow-up in response to an abnormal test result. For example, trigger algorithms could be incorporated into EHR-based provider notification systems or into panel management programmes allowing the care team or dedicated individuals at an institution to efficiently identify and address delays in follow-up of abnormal findings. Further evaluation should determine whether this strategy will actually lead to improving timeliness of diagnosis and patient outcomes, including improving stage at diagnosis and morbidity and mortality, and how such interventions could be implemented within the context of the outpatient setting.

Although we were able to achieve over 50% PPV for each trigger, additional useful information remains inaccessible inside free-text progress notes. Future incorporation of text mining or natural language processing methods⁵² could potentially be used to extract

information to further improve trigger PPV. Strategies to identify patients at higher risk, or with more risk factors for lack of timely follow-up, could include the integration of statistical predictive models into triggers. At this time, however, such statistical models are not known to exist, and must first be developed and validated.

Several limitations merit discussion. First, our study was performed at only two sites, both of which were integrated health systems and used a comprehensive EHR. Thus, our results might not necessarily be generalised to other sites. However, our triggers use a common query language and rely on fairly standard data criteria used across the USA, such as lab test and ICD-9 codes, and it is likely that only minor changes would be needed to implement triggers at new sites. This also allows triggers to be modified for site-specific needs or as guidelines change, such as with shifting PSA recommendations. As terminology standards become increasingly adopted, such algorithms should become even more portable. Second, we were unable to report sensitivity and specificity of the triggers due to the vast number of records requiring review to identify a single false negative, and thus, our results are affected by the low prevalence of missed follow-up given the large number of patients who receive diagnostic testing, as well as the paucity of adverse events that occur even when care is delayed. This is a commonly cited limitation to data mining and trigger development where the outcome of interest has a low prevalence⁵³; however, the use of these triggers allows automation of a process that would otherwise be extremely difficult, and potentially involve manual reviews of thousands of records. Third, the study was not designed to identify the root cause of the delayed care or missed diagnosis. For example, reviewers noted many cases where delays in follow-up were beyond the control of primary care providers, such as difficulty obtaining timely appointments with specialists, or patients failing to show up at scheduled appointments. However, trigger information could still facilitate delivery of timely healthcare. Additionally, evaluation of follow-up was based on chart reviews and may not fully reflect the care delivered or provider's rationale. For example, many instances of appropriate PSA follow-up were performed after the 90-day period, indicating some providers may have intentionally chosen to delay care, but did not document their reasoning to do so. Finally, this study was designed to assess development feasibility and performance metrics of triggers to detect potential delays in diagnosis. Future work is planned to evaluate the impact of using triggers prospectively to impact clinical outcomes.

In conclusion, we successfully developed electronic triggers to identify patients at high risk for delayed diagnosis for prostate and CRC. Triggers had reasonable predictive values and could be useful for others trying to develop measurement systems to detect

delays in diagnostic care. This study serves as a basis for future research to evaluate the effect of prospective application of triggers on patient outcomes.

Author affiliations

¹Houston VA Health Services Research & Development Center of Excellence, and The Center of Inquiry to Improve Outpatient Safety Through Effective Electronic Communication, both at the Michael E. DeBakey Veterans Affairs Medical Center and the Section of Health Services Research and Development, Houston, Texas, USA

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

³University of Texas Houston Medical School, Department of Internal Medicine; and the UT Houston-Memorial Hermann Center for Healthcare Quality & Safety, Houston, Texas, USA

⁴Department of Clinical Effectiveness and Performance Measurement, St. Luke's Health System, Houston, Texas, USA

⁵Department of Family and Community Medicine, Scott & White Healthcare, Texas A&M Health Science Center, College of Medicine, Temple, Texas, USA

⁶University of Texas School of Public Health, Houston, Texas, USA.

Acknowledgements The authors thank Louis Wu, Dawn Begaye, Anne Robertson, Harvinder Arora and Kenneth Hung for their efforts in completing this study.

Contributors DM, ET, AE, SF, RP and HS participated in the conception and design of this study. DM, AL, BR, EJ, MK and HS participated in the analysis data and interpretation of results. DM, AL and HS drafted the article, and all other authors performed critical review of the articles. All authors have approved the submitted version of this article.

Funding This study is supported by an NIH K23 career development award (K23CA125585), the Agency for Health Care Research and Quality (R18HS017820) and in part by the Houston VA HSR&D Center of Excellence (HFP90-020). These sources had no role in the preparation, review, or approval of the manuscript.

Competing interests None.

Ethics approval Baylor College of Medicine.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All authors had full access to all the data in the study, and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

- Macleod U, Mitchell ED, Burgess C, *et al.* Risk factors for delayed presentation and referral of symptomatic cancer: evidence for common cancers. *Br J Cancer* 2009;101: S92–S101.
- Poon EG, Kachalia A, Puopolo AL, *et al.* Cognitive errors and logistical breakdowns contributing to missed and delayed diagnoses of breast and colorectal cancers: a process analysis of closed malpractice claims. *J Gen Intern Med* 2012;27:1416–23.
- Singh H, Sethi S, Raber M, *et al.* Errors in cancer diagnosis: current understanding and future directions. *J Clin Oncol* 2007;25:5009–18.
- 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.
- Singh H, Daci K, Petersen LA, *et al.* Missed opportunities to initiate endoscopic evaluation for colorectal cancer diagnosis. *Am J Gastroenterol* 2009;104:2543–54.
- Wahls TL, Peleg I. Patient- and system-related barriers for the earlier diagnosis of colorectal cancer. *BMC Fam Pract* 2009;10:65.
- 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.
- Phillips RL Jr, Bartholomew LA, Dovey SM, *et al.* Learning from malpractice claims about negligent, adverse events in primary care in the United States. *Qual Saf Health Care* 2004;13:121–6.
- The Lancet Oncology. Cancer detection and primary care... revisited. *Lancet Oncol* 2012;13:559.
- Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005;165:1493–9.
- Singh H, Davis Giardina T, Petersen LA, *et al.* Exploring situational awareness in diagnostic errors in primary care. *BMJ Qual Saf* 2012;21:30–8.
- McFall SL, Smith DW. Lack of follow-up of prostate-specific antigen test results. *Public Health Rep* 2009;124:718–25.
- Zeliadt SB, Hoffman RM, Etzioni R, *et al.* What happens after an elevated PSA test: the experience of 13,591 Veterans. *J Gen Intern Med* 2010;25:1205–10.
- Yabroff KR, Washington KS, Leader A, *et al.* Is the promise of cancer-screening programs being compromised? Quality of follow-up care after abnormal screening results. *Med Care Res Rev* 2003;60:294–331.
- Bastani R, Yabroff KR, Myers RE, *et al.* Interventions to improve follow-up of abnormal findings in cancer screening. *Cancer* 2004;101:1188–200.
- Nepple KG, Joudi FN, Hillis SL, *et al.* Prevalence of delayed clinician response to elevated prostate-specific antigen values. *Mayo Clinic Proc* 2008;83:439–45.
- Singh H, Thomas EJ, Mani S, *et al.* Timely follow-up of abnormal diagnostic imaging test results in an outpatient setting: are electronic medical records achieving their potential? *Arch Intern Med* 2009;169:1578–86.
- Fisher DA, Jeffreys A, Coffman CJ, *et al.* Barriers to full colon evaluation for a positive fecal occult blood test. *Cancer Epidemiol Biomarkers Prev* 2006;15:1232–5.
- Etzioni DA, Yano EM, Rubenstein LV, *et al.* Measuring the quality of colorectal cancer screening: the importance of follow-up. *Dis Colon Rectum* 2006;49:1002–10.
- Singh H, Kadiyala H, Bhagwath G, *et al.* Using a multifaceted approach to improve the follow-up of positive fecal occult blood test results. *Am J Gastroenterol* 2009;104: 942–52.
- Singh H, Thomas EJ, Sittig DF, *et al.* Notification of abnormal lab test results in an electronic medical record: do any safety concerns remain? *Am J Med* 2010;123:238–44.
- Shin DW, Cho J, Kim SY, *et al.* Delay to Curative Surgery Greater than 12 Weeks Is Associated with Increased Mortality in Patients with Colorectal and Breast Cancer but Not Lung or Thyroid Cancer. *Ann Surg Oncol* 2013;20:2468–76.
- Frankovich J, Longhurst CA, Sutherland SM. Evidence-based medicine in the EMR era. *N Engl J Med* 2011;365:1758–9.
- Murphy DR, Reis B, Sittig DF, *et al.* Notifications received by primary care practitioners in electronic health records: a taxonomy and time analysis. *Am J Med* 2012;125:209.e1–7.
- Murphy DR, Reis B, Kadiyala H, *et al.* Electronic health record-based messages to primary care providers: valuable information or just noise? *Arch Intern Med* 2012;172:283.
- McDonald CJ, McDonald MH. INVITED COMMENTARY—electronic medical records and preserving primary care physicians' time: comment on 'electronic health record-based messages to primary care providers'. *Arch Intern Med* 2012;172:285.

- 27 Jha AK, Classen DC. Getting moving on patient safety—harnessing electronic data for safer care. *N Engl J Med* 2011;365:1756–8.
- 28 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 Affairs* 2011;30:581–9.
- 29 Resar RK, Rozich JD, Classen D. Methodology and rationale for the measurement of harm with trigger tools. *Qual Saf Health Care* 2003;12(Suppl 2):ii39–45.
- 30 De Wet C, Bowie P. The preliminary development and testing of a global trigger tool to detect error and patient harm in primary-care records. *Postgrad Med J* 2009;85:176–80.
- 31 Schiff GD, Kim S, Krosnjar N, *et al.* Missed hypothyroidism diagnosis uncovered by linking laboratory and pharmacy data. *Arch Intern Med* 2005;165:574–7.
- 32 Szekendi MK, Sullivan C, Bobb A, *et al.* Active surveillance using electronic triggers to detect adverse events in hospitalized patients. *Qual Saf Health Care* 2006;15:184–90.
- 33 Honigman B, Light P, Pulling RM, *et al.* A computerized method for identifying incidents associated with adverse drug events in outpatients. *Int J Med Inform* 2001; 61:21–32.
- 34 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.
- 35 Classen DC, Pestotnik SL, Evans RS, *et al.* Computerized surveillance of adverse drug events in hospital patients. 1991. *Qual Saf Health Care* 2005;14:221–5.
- 36 Mull HJ, Nebeker JR. Informatics tools for the development of action-oriented triggers for outpatient adverse drug events. *AMIA Annu Symp Proc* 2008;505–9.
- 37 Knirsch CA, Jain NL, Pablos-Mendez A, *et al.* Respiratory isolation of tuberculosis patients using clinical guidelines and an automated clinical decision support system. *Infect Control Hosp Epidemiol* 1998;19:94–100.
- 38 Muething SE, Conway PH, Kloppenborg E, *et al.* Identifying causes of adverse events detected by an automated trigger tool through in-depth analysis. *Quality and Safety in Health Care* 2010;19:435–9.
- 39 Roos LL Jr, Cageorge SM, Austen E, *et al.* Using computers to identify complications after surgery. *Am J Public Health* 1985;75:1288–95.
- 40 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.
- 41 Jones R, Latinovic R, Charlton J, *et al.* Alarm symptoms in early diagnosis of cancer in primary care: cohort study using General Practice Research Database. *BMJ* 2007;334:1040–1040.
- 42 Singh H, Petersen LA, Daci K, *et al.* Reducing referral delays in colorectal cancer diagnosis: is it about how you ask? *Qual Saf Health Care* 2010;19:e27.
- 43 Henning J, Waxman S. Legal aspects of men’s genitourinary health. *Int J Impot Res* 2009;21:165–70.
- 44 Lin K, Croswell JM, Koenig H, *et al.* Prostate-specific antigen-based screening for prostate cancer: an evidence update for the U.S. preventive services task force. Agency for Healthcare Research and Quality, 2011.
- 45 Schröder FH, Hugosson J, Roobol MJ, *et al.* Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med* 2012; 366:981–90.
- 46 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;36:120–30.
- 47 VHA Directive 2007-004: Colorectal Cancer Screening. 2007. http://www1.va.gov/vhapublications/ViewPublication.asp?pub_ID=1530
- 48 Cancer Waiting Times | National Cancer Action Team. <http://ncat.nhs.uk/our-work/ensuring-better-treatment/cancer-waiting-times#> (accessed 20 Jun 2013).
- 49 Reddy SG, Babbott SF, Beasley BW, *et al.* Prevalence and functionality of electronic health records in internal medicine continuity clinics. *Acad Med* 2010;85:1369–77.
- 50 Jha AK, DesRoches CM, Campbell EG, *et al.* Use of electronic health records in U.S. hospitals. *N Engl J Med* 2009;360:1628–38.
- 51 Lorincz C, Drazen E, Sokol P, *et al.* *Research in ambulatory patient safety 2000–2010: a 10-year review*. Chicago, IL: American Medical Association, 2011.
- 52 Joshi AK. Natural language processing. *Science* 1991;253:1242–9.
- 53 Nebeker J, Stoddard G, Rosen A. Considering sensitivity and positive predictive value in comparing the performance of triggers systems for iatrogenic adverse events. Abstract presented at the Agency for Healthcare Research and Quality Annual Conference. Bethesda, MD, September 2009.