Electronic Detection of Delayed Test Result Follow-Up in Patients with Hypothyroidism

Ashley N. D. Meyer, PhD[1], Daniel R. Murphy, MD, MBA[1], Aymer Al-Mutairi, MD[2], Dean F. Sittig, PhD[3], Li Wei, MS[1], Elise Russo, MPH[1], and Hardeep Singh, MD, MPH[1]

[1]Houston VA Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey VA Medical Center and Department of Medicine, Baylor College of Medicine, Houston, TX, USA; [2]Department of Family & Community Medicine, Baylor College of Medicine, Houston, TX, USA; [3]School of Biomedical Informatics and UT-Memorial Hermann Center for Healthcare Quality and Safety, The University of Texas Health Science Center at Houston, Houston, TX, USA.

BACKGROUND: Delays in following up abnormal test results are a common problem in outpatient settings. Surveillance systems that use trigger tools to identify delayed follow-up can help reduce missed opportunities in care.

OBJECTIVE: To develop and test an electronic health record (EHR)-based trigger algorithm to identify instances of delayed follow-up of abnormal thyroid-stimulating hormone (TSH) results in patients being treated for hypothyroidism.

DESIGN: We developed an algorithm using structured EHR data to identify patients with hypothyroidism who had delayed follow-up (>60 days) after an abnormal TSH. We then retrospectively applied the algorithm to a large EHR data warehouse within the Department of Veterans Affairs (VA), on patient records from two large VA networks for the period from January 1, 2011, to December 31, 2011. Identified records were reviewed to confirm the presence of delays in follow-up.

KEY RESULTS: During the study period, 645,555 patients were seen in the outpatient setting within the two networks. Of 293,554 patients with at least one TSH test result, the trigger identified 1250 patients on treatment for hypothyroidism with elevated TSH. Of these patients, 271 were flagged as potentially having delayed follow-up of their test result. Chart reviews confirmed delays in 163 of the 271 flagged patients (PPV = 60.1%).

CONCLUSIONS: An automated trigger algorithm applied to records in a large EHR data warehouse identified patients with hypothyroidism with potential delays in thyroid function test results follow-up. Future prospective application of the TSH trigger algorithm can be used by clinical teams as a surveillance and quality improvement technique to monitor and improve follow-up.

KEY WORDS: thyroid-stimulating hormone (TSH); electronic algorithms; triggers; patient safety; hypothyroidism; test result follow-up.

INTRODUCTION

Failure to follow up abnormal test results in a timely fashion is a common problem in the outpatient setting.[1–9] Test result management in an electronic health record (EHR)-enabled health care system is susceptible to both technological and non-technological vulnerabilities.[10,11] EHR-based test result notification systems are aimed at reducing delays in follow-up of abnormal test results, and likely represent an improvement over communication using paper-based methods.[12] However, their use has not eliminated follow-up delays, even when health care providers acknowledge receipt of notifications, a technological limitation of EHRs.[13] Indeed, our prior work revealed that 7–8% of EHR-based abnormal test result notifications were not acted on within 4 weeks.[14,15] Cognitive limitations of providers also contribute to delays in follow-up: for example, providers are often unable to discern relevant versus irrelevant information when assessing the need for follow-up while also dealing with information overload and alert fatigue.[16,17] Given these and other vulnerabilities in outpatient care, "back-up" surveillance systems are needed to help ensure that follow-up actions proceed as intended.

A recent report from the National Academy of Medicine "Improving Diagnosis in Health Care,"[18] calls for improved EHR-based measurement strategies to identify and reduce follow-up delays. One such method that addresses the technological limitations of existing notification systems and cognitive limitations of providers is the use of electronic trigger tools to systematically mine EHR data to flag medical records with potential delays in care. The goal then is to make that information available to the provider or the associated care team in an effort to prevent subsequent delays. Thus, trigger tools can be used to alert clinicians about possible delays and other adverse events.[19–22] Triggers consist of an algorithm applied to an EHR database to identify, or "trigger," all patient records that match a predefined pattern of care (e.g., a patient with a test result flagged as "abnormal" but with no provider follow-up within 30 days of the result). We had previously developed and validated a trigger that scanned large repositories of clinical data to identify patients with potentially delayed diagnostic clues that warranted an evaluation for prostate, colorectal, or lung cancer. We found that positive predictive

values (PPVs) of the trigger algorithms were all >50%, which appears promising for real-world application.[23] EHR triggers can thus automatically identify medical records at risk for delayed follow-up of abnormal test results and facilitate timely follow-up action.[24]

Thyroid-stimulating hormone (TSH) test results can often be missed or delayed for follow-up,[25] leading to suboptimal patient outcomes. Clinical guidelines recommend follow-up of abnormal TSH levels within 4–8 weeks after testing, with the goal of ensuring a euthyroid state using oral thyroxine (T4).[26] However, not all patients treated with thyroxine achieve normal TSH levels, making subsequent TSH testing and follow-up of those results necessary for appropriate monitoring and for preventing long-term complications such as cardiac conditions,[27] infertility,[28] anxiety and depression,[29] and—most concerning—birth defects[30] in pregnant women, all of which would also cause increased utilization of health care resources. The purpose of the present study was to develop and evaluate an EHR-based trigger algorithm to identify delays in follow-up of abnormal TSH test results in patients with hypothyroidism. Once validated, this trigger could be used prospectively by applying it frequently on near real-time data to help providers know when their treatment plans may not be working as expected.

METHODS

Design and Setting

We conducted a study to develop and validate an electronic trigger to identify delayed follow-up of abnormal TSH results of patients undergoing treatment for hypothyroidism. The Department of Veterans Affairs (VA) has created a large national data warehouse of clinical EHR data that includes inpatient, outpatient, mental health, rehabilitation, and long-term care services.[31] The data are collected from all 144 VA facilities, serving over 8 million veterans. This study was performed using data from January 1, 2011, to December 31, 2011, from two large VA networks, cumulatively encompassing 17 medical centers and associated clinics across 12 states. In 2011 alone, providers in these two networks saw 645,555 unique patients in outpatient settings. The study was approved by our institutional review board and the VA Office of Research.

Trigger Development and Refinement

Our team developed the trigger to be consistent with therapy and monitoring guidelines published by the American Thyroid Association and the American Association of Clinical Endocrinologists.[26] We specifically focused on patients with an existing hypothyroidism diagnosis who were currently receiving therapy, since these patients essentially always require a therapeutic intervention in response to an elevated TSH. In contrast, patients with newly elevated TSH were excluded, as

subsequent action might involve other types of diagnostic evaluation. The trigger was designed first to identify adult patients with elevated TSH values (>10 mIU/L) who were receiving thyroid replacement therapy (TRT) prior to the date of the TSH result. When multiple elevated TSH values occurred in a patient, we used the earliest elevated TSH in the study period (when triggers are used prospectively, the first instance should lead to action, and thus later occurrences would not be expected to occur). The trigger then excluded instances where thyroid replacement therapy would not be appropriate or expected, including (1) a diagnosis of hyperthyroidism on the problem list, (2) where evidence of hospitalization, emergency room, or urgent care visit was detected within 24 h prior to the elevated TSH (because acute illnesses can skew thyroid testing), and (3) patient non-adherence within 30 days after the elevated TSH based on the presence of an International Classification of Diseases, Ninth Revision (ICD-9) code V15.81 (because the provider would be more likely to emphasize medication compliance rather than making changes to thyroid replacement therapy). Finally, the trigger excluded records containing evidence of expected follow-up, which included orders for new prescriptions for TRT within 30 days or repeat TSH testing performed within 30 days after the elevated TSH result. We chose 30 days as a reasonable cut-off to allow plenty of time for typical treatment and communication pathways to proceed naturally, without causing significant clinical impact from delays. The algorithm for detecting these criteria was created using Structured Query Language (SQL) programming designed to extract data in structured data fields such as ICD-9, Logical Observation Identifiers Names and Codes (LOINC)[32], and Current Procedural Terminology (CPT) codes.

After the initial criteria were drafted and programmed into an SQL-based algorithm, we pilot-tested the trigger by applying it to the clinical data warehouse of all patients treated at one study site between January 1 and December 31, 2009. Fifty randomly selected patients identified by the trigger algorithm served as a test cohort. In our prior work, this sample size was more than adequate to identify issues with the trigger logic and make iterative refinements. An actively practicing board-certified primary care physician (AA) reviewed the EHR for each patient to identify evidence to confirm delayed follow-up (including missing follow-up) of abnormal TSH results. Based on this preliminary review of the test cohort, the trigger criteria were refined to produce the final trigger criteria. Specifically, repeat TSH testing was adjusted from 30 days to 60 days after the initial abnormal TSH value date based on clinical guidelines that allow 4–8 weeks for follow-up, to allow ample time for scheduling, and to account for TSH stabilization which often takes 6 weeks after thyroid medication adjustments.[26] True-positive cases were defined as patient records in which the reviewing physician could *not* identify any of the following types of follow-up action: TRT adjustment within 30 days, repeat TSH testing within 60 days, or a plan to delay action. False-positive cases were defined as patient records where,

although the trigger identified no follow-up action, during chart review the reviewing physician identified either a follow-up action, a documented plan to delay follow-up action until a future date, or some other reason why follow-up should not have been expected (e.g., patient saw an external provider). We considered providers the most knowledgeable about each patient's clinical situation, and thus if a provider documented a specific intention to delay follow-up in the free-text notes (i.e., data not accessible to the trigger), we did not consider this a delay in follow-up. The final criteria are listed in Table 1.

Trigger Validation

We applied the finalized trigger to the 2011 data at the study sites. Trigger-positive patient records were reviewed by a physician using a manual data collection instrument developed based on previous work on trigger development and evaluation.[23,24,33] The reviewer (AA) used the review instrument to evaluate whether a delay was truly experienced by the patient, collect reasons for false-positive results, and determine patient and provider characteristics that potentially impacted delays. Characteristics included patient age; gender; race; presence of comorbidities including diabetes, hypertension, hyperlipidemia, obesity, and depression; type of provider seen when being tested for the TSH results (primary care provider [PCP] vs. specialist); who was primarily managing the patient's thyroid disease (PCP vs. an endocrinologist); what type of PCP the patient regularly saw (physician vs. nurse practitioner or physician's assistant); and whether the patient had previous PCP or endocrine visits.

Table 1 Criteria for Electronic Trigger Applied to All Patient Records

Inclusion criteria
• TSH >10 mIU/L*
• Actively taking thyroid replacement therapy, defined by either:
  ○ New or renewed TRT medication order within prior 90 days†
  ○ Presence of TRT on patient's medication list
Clinical Exclusions:
• TSH performed during or within 24 h prior to hospitalization or emergency department/urgent care visit
• Patients with medication non-adherence‡
• Patients with hyperthyroidism diagnosis
• Age <18 or >100 years
Appropriate follow-up criteria
• New TRT prescription within 30 days after TSH result
• Repeat TSH level within 60 days after TSH result

*Found using Logical Observation Identifiers Names and Codes (LOINC) codes: 3016,11579, and 11580
†TRT = thyroid replacement therapy (includes Synthroid, Levothyroxine, Lovoxyl, Armour, Thyrolar, Liotrix, Cytomel, Liothyronine Sodium, Unithroid, L-Thyroxin, Levo-T, Levoxyl, Novothyrox, Levothroid, Levolet, and Triostat)
‡Found using International Classification of Diseases, Ninth Revision (ICD-9) codes: 250.0, 250.00, 250.01, 250.02, 250.03, 250.1, 250.10, 250.11, 250.12, 250.13, 250.2, 250.20, 250.21, 250.22, 250.23, 250.3, 250.30, 250.31, 250.32, 250.33, 250.4, 250.40, 250.41, 250.42, 250.43, 250.5, 250.50, 250.51, 250.52, 250.53, 250.6, 250.60, 250.61, 250.62, 250.63, 250.7, 250.70, 250.71, 250.72, 250.73, 250.8, 250.80, 250.81, 250.82, 250.83, 250.9, 250.90, 250.91, 250.92, 250.93, 401.0, 401.1, 401.9, 405.01, 405.09, 405.11, 405.19, 405.91, 405.99, 272.1, 272.2, 272.3, 272.4, 272.5, 272.6, 272.7, 272.8, 272.9, 278.00, 278.01, 278.03, 290.13, 290.21, 296.20, 296.21, 296.22, 296.23, 296.24, 296.25, 296.26, 296.30, 296.31, 296.32, 296.33, 296.34, 296.35, 296.36, 296.82, 298.0, 301.12, 309.1, and 311

Statistical Analysis

The positive predictive value (PPV) of the TSH trigger was calculated as the number of true-positive cases confirmed by physician chart review divided by the total number of records identified by the EHR trigger. Descriptive statistics were also provided for cases with delayed follow-up and those with appropriate follow-up. Lastly, we compared patient and provider characteristics between records that were confirmed to have delayed versus timely follow-up using independent t tests for continuous variables and Fisher's exact or chi-square tests for categorical variables. We used SPSS version 21 software (IBM Corporation, Armonk, NY) to perform all analyses.

RESULTS

Trigger criteria were applied to the records of all 645,555 patients who had an outpatient visit during the study period, of which 293,554 patients had at least one TSH result (Fig. 1). Of these patients with a TSH result, 5230 patients (1.8%) had a TSH result above 10 mIU/L (abnormally high—these 5230 patients had a mean of 1.5 abnormally high results in the study period; however, only the first one for each patient was included). The trigger algorithm excluded 3980 patients based on new hypothyroidism diagnoses (elevated TSH results in patients with no prior hypothyroidism diagnosis), recent hospitalizations (due to artificially abnormal TSH results in acute illness), age, or hyperthyroidism diagnoses, leaving 1250
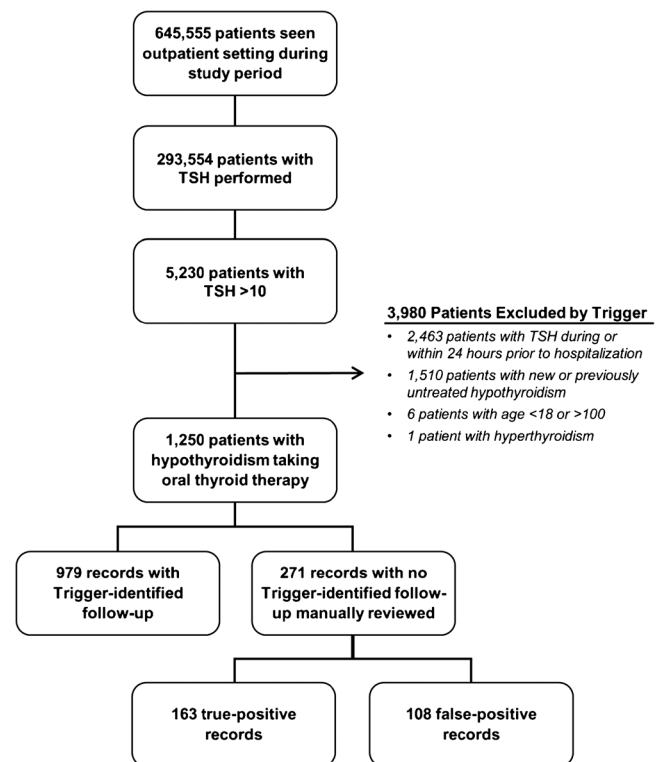
Figure 1 Trigger flowchart and findings.

patients where follow-up action was needed. Of these, the trigger identified expected follow-up action in 979 records and flagged 271 records as having potential follow-up delays (21.7% of all TRT-treated patients with known hypothyroidism). No non-adherence codes were identified by the trigger. The physician reviewer confirmed delayed follow-up in 163 of the 271 triggered patients with elevated TSH, yielding a PPV of 60.1% (95% CI: 54.0–66.0%), meaning that 13.0% of all TRT-treated patients in this population with known hypothyroidism had confirmed delayed follow-up.

In most instances (n = 126 of 163, 77.3%), records lacked any clinician documentation regarding the abnormal TSH result (Table 2). An analysis of the reasons for true-positive and false-positive findings is presented in Table 2. The most common reasons for false-positive findings (n = 108) were free-text documentation (which is not interpretable by the computer algorithm) in the progress note that patients failed to take medications as instructed (n = 58, 53.7%), that appropriate follow-up in the form of an increase in medication dose occurred (n = 21, 19.4%), or that deliberate plans were made to follow up within a time period longer than 60 days (n = 7, 6.5%). There were no significant differences in patient or provider characteristics in the delayed versus timely follow-up groups (see Table 3). For patients identified as having delayed follow-up, repeat TSH testing occurred a median of 218 days after the abnormal test result (interquartile range: 147–419 days).

DISCUSSION

We developed and validated an EHR-based trigger for identifying delays in follow-up of abnormal TSH levels in patients with hypothyroidism and found the trigger to have a PPV of 60.1%. This relatively high PPV translates to an average of 1.7 records needing to be reviewed after trigger application to benefit one patient. Among all patients treated for hypothyroidism, 13% (163 of 1250) had delayed follow-up of an elevated TSH, underscoring the need for a tool to detect such patients in order to optimize treatment. Without such a tool, all 293,554 patient charts with a TSH result would have required examination to find the same 163 patients with delayed follow-up, a process that would be impractical for clinical application. The trigger, on the other hand, narrowed the number of charts needing review to 271 (or 0.9% of the charts), and required no clinician action to do so, greatly improving the efficiency of identifying possible follow-up delays. Thus, our study suggests a potential benefit for this EHR-based TSH trigger to be used prospectively to improve the timeliness and quality of care for patients with elevated TSH levels. The median time to follow-up in these cases was 218 days, providing a large window of opportunity to intervene with prospective trigger tools in these cases. Such tools would have the potential to decrease this follow-up time from 218 days to closer to 60 days, potentially reducing the adverse effects of such delays on patients.

Despite a reasonable PPV (i.e., one which is high enough to be meaningful to a provider and leads to more signal than noise), our trigger led to a high number of false-positive cases. However, the reduction in the number of charts needing review from 293,554 to 271 to find 163 delays (and 108 false positives) may be worth the effort. Additionally, further analysis of these false-positive cases suggests a need for enhanced methods to mine data from EHRs to lower the number of false-positive cases in the future, thus improving the PPV. Justification for not taking additional action in the false positives was often documented in progress notes; therefore, future work could use natural language processing methods to allow triggers to make use of the wealth of useful clinical information locked in free-text data.[34] Alternatively, future trigger systems could rely on additional structured data entry by clinicians. For example, the use of a structured data field by physicians to reflect dosing changes in patients' medication lists instead of documenting them in a free-text note (e.g., "take two tablets instead of one") would decrease the number of false-positive results found with this trigger. The use of existing, routinely collected, structured data by our algorithm enhances portability, as standard ICD codes are used ubiquitously within other institutions. Similar trigger algorithms could be used to identify delayed follow-up action of several types of diagnostic tests in both VA and non-VA settings. However, even though triggers using ICD-9 codes may be portable to other health systems, variation in the presence and structure of clinical data elements contained within data repositories would require modification for use at individual facilities.

There were several limitations to our study. First, our reviews were dependent on the documentation in the patient's

Table 2 Reasons for Definition as True Positive or False Positive

Reasons	n (%)
True-positive records (n = 163)	
No documentation that addressed abnormal TSH (no recognition)	126 (77.3%)
No appropriate action taken (TSH acknowledged, but no action mentioned)	33 (20.2%)
Confusion regarding who should address the test (provider ordering the test vs. PCP)	2 (1.2%)
Provider documented that hypothyroidism was controlled despite abnormal TSH	2 (1.2%)
False-positive records (n = 108)	
Patient failed to take medications as instructed, physician discussed appropriate dosing	58 (53.7%)
Dose increase documented only in free text	21 (19.4%)
Repeat TSH testing ordered with deliberate plan to test after 60 days	7 (6.5%)
Patient followed up with another external provider	6 (5.6%)
Referral made to endocrinology	5 (4.6%)
TSH found trending downward on consecutive testing and later retesting planned	5 (4.6%)
Elevated TSH caused by Thyrogen testing	3 (2.8%)
Inability to reach patient for follow-up despite multiple attempts	3 (2.8%)

PCP primary care provider, TSH thyroid-stimulating hormone

Table 3 Patient and Provider Characteristics by Delay Group

Characteristics	Patients with delayed follow-up (n=163)	Patients with no delayed follow-up (n=108)	P value (Fisher's or chi-square tests where appropriate)
Age [mean (SD)]	62.8 (15.5)	62.2 (16.6)	0.32 (*t* test)
Gender [n (%)]			
Male	147 (90.2%)	93 (86.1%)	0.33
Female	16 (9.8%)	15 (13.9%)	
Race [n (%)]			
American Indian/Alaskan Native	4 (2.5%)	0 (0.0%)	0.09
Black/African American	37 (22.7%)	16 (14.8%)	
Native Hawaiian/Pacific Islander	0 (0.0%)	2 (1.9%)	
White	103 (63.2%)	77 (71.3%)	
White-non Hispanic	4 (2.5%)	1 (0.9%)	
Unknown	15 (9.2%)	12 (11.1%)	
Diabetes [n (%)]			
No	120 (73.6%)	79 (73.1%)	~1.00
Yes	43 (26.4%)	29 (26.9%)	
Hypertension [n (%)]			
No	45 (27.6%)	42 (38.9%)	0.06
Yes	118 (72.4%)	66 (61.1%)	
Hyperlipidemia [n (%)]			
No	66 (40.5%)	42 (38.9%)	0.80
Yes	97 (59.5%)	66 (61.1%)	
Obesity [n (%)]			
No	107 (65.6%)	74 (68.5%)	0.69
Yes	56 (34.4%)	34 (31.5%)	
Depression [n (%)]			
No	141 (86.5%)	99 (91.7%)	0.24
Yes	22 (13.5%)	9 (8.3%)	
TSH ordering provider [n (%)]			
Specialist	22 (13.5%)	9 (8.3%)	0.24
PCP	141 (86.5%)	99 (91.7%)	
Person primarily managing patient's thyroid disease [n (%)]			
PCP	152 (93.3%)	93 (86.1%)	0.15
Endocrinologist	7 (4.3%)	10 (9.3%)	
Unknown or None	4 (2.5%)	5 (4.6%)	
Type of PCP patient regularly sees [n (%)]			
Physician	130 (79.8%)	83 (76.9%)	0.65
Nurse practitioner or physician's assistant	33 (20.2%)	25 (23.1%)	
Prior PCP visit [n (%)]			
No	6 (3.7%)	3 (2.8%)	~1.00
Yes	157 (96.3%)	105 (97.2%)	
Prior endocrine visit [n (%)]			
No	148 (90.8%)	96 (88.9%)	0.68
Yes	15 (9.2%)	12 (11.1%)	

record; thus, undocumented actions would not be accounted for in our review, possibly resulting in an overestimation of PPV and an underestimation of the false-positive rate. However, in our prior work on determining delayed follow-up of abnormal laboratory (including TSH) and imaging results, we found that most instances without documented follow-up were truly instances where no follow-up occurred, as confirmed with the providers.[12,13,15] Second, due to limited resources, our study did not involve a review of trigger-negative records, precluding the ability to assess the trigger's negative predictive value, sensitivity, or specificity. However, the use of the trigger still allowed us to identify 163 delays in care that otherwise would not have been found. Additional evaluation is warranted, especially in other integrated health systems, using comprehensive electronic records. Third, we were similarly unable to have a second reviewer review a subset of records to confirm delays. However, inter-rater reliability for our chart review studies identifying similar types of delays has been quite high (kappas typically > 0.80),[24,35] adding confidence to our findings. Finally, data on clinical outcomes were not

included or analyzed, limiting our ability to understand the clinical impact of care delays in these cases. However, additional research examining the clinical impact as well as downstream impacts of this and similar tools should be a component of future research.

Compared to random manual record reviews for identifying patients at risk for care delays, triggers are able to more efficiently flag patients lost to follow-up, with lower personnel resource use.[23,33] Future prospective application of the TSH trigger algorithm can be used by clinical teams as a surveillance and quality improvement technique to monitor and improve follow-up.

However, we believe trigger development efforts are only the beginning in an effort to positively impact clinical care, namely, by calling attention to delays. Physicians or other team members must take action.[36] We believe future efforts are needed to use triggers prospectively in real-time clinical settings, as well as to explore more advanced algorithms that attempt to glean information from free-text notes in the record.

CONCLUSION

We developed an EHR trigger to identify follow-up delays for abnormal TSH test results, with a PPV of 60.1%, in outpatients with hypothyroidism. Identifying these delays has the potential to prevent health complications resulting from untreated hypothyroidism ranging from cardiac conditions to birth defects in pregnant women. Our trigger, with modifications (such as for data structure or inclusion and exclusion criteria), could be adapted to other EHR systems and other types of diagnostic tests. Thus, this methodology may be useful for measuring, monitoring, and improving follow-up of patients, and could support quality improvement programs for reducing care delays.

Corresponding Author: *Ashley N. D. Meyer, PhD; Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety Michael E. DeBakey Veterans Affairs Medical Center and the Section of Health Services Research, Department of Medicine Baylor College of Medicine, 2002 Holcombe Boulevard (152), Houston, TX 77030, USA (e-mail: ameyer@bcm.edu).*

REFERENCES

1. **Singh H, Naik A, Rao R, Petersen L.** Reducing diagnostic errors through effective communication: harnessing the power of information technology. J Gen Intern Med. 2008;23:489–94.
2. **Poon E, Gandhi T, Sequist T, Murff H, Karson A, Bates D.** "I wish I had seen this test result earlier!": dissatisfaction with test result management systems in primary care. Arch Intern Med. 2004;164:2223–8.
3. **Gandhi TK.** Fumbled handoffs: one dropped ball after another. Ann Intern Med. 2005;142:352–8.
4. **Poon EG, Haas JS, Louise PA, et al.** Communication factors in the follow-up of abnormal mammograms. J Gen Intern Med. 2004;19:316–23.
5. **Schiff GD.** Introduction: communicating critical test results. Jt Comm J Qual Patient Saf. 2005;31:63–5.
6. **Hickner J, Graham DG, Elder NC, et al.** Testing process errors and their harms and consequences reported from family medicine practices: a study of the American Academy of Family Physicians National Research Network. Qual Saf Health Care. 2008;17:194–200.
7. **Singh H, Petersen LA, Thomas EJ.** Understanding diagnostic errors in medicine: a lesson from aviation. Qual Saf Health Care. 2006;15:159–64.
8. **Singh H, Graber M.** Reducing diagnostic error through medical home-based primary care reform. JAMA. 2010;304:463–4.
9. **Moore C, Saigh O, Trikha A, Lin JJ.** Timely follow-up of abnormal outpatient test results: Perceived barriers and impact on patient safety. J Patient Saf. 2008;4.
10. **Wears RL, Berg M.** Computer technology and clinical work: still waiting for Godot. JAMA. 2005;293:1261–3.
11. **Sittig DF, Singh H.** A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care. 2010;19:i68–74.
12. **Singh H, Arora H, Vij M, Rao R, Khan MM, Petersen L.** Communication outcomes of critical imaging results in a computerized notification system. J Am Med Inform Assoc. 2007;14:459–66.
13. **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.
14. **Al-Mutairi A, Meyer AN, Chang P, Singh H.** Lack of timely follow-up of abnormal imaging results and radiologists' recommendations. J Am Coll Radiol. 2015.
15. **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.
16. **Murphy DR, Reis B, Sittig DF, Singh H.** Notifications received by primary care practitioners in electronic health records: a taxonomy and time analysis. Am J Med. 2012;125:209e1–7.
17. **Murphy DR, Meyer AND, Russo E, Sittig DF, Wei L, Singh H.** The burden of inbox notifications in commercial electronic health records. JAMA Intern Med. 2016.
18. Improving diagnosis in health care. *National Academies of Sciences Engineering and Medicine* [serial online]. 2015. Available from: The National Academies Press. Accessed 14 June 2016.
19. Institute for Healthcare Improvement, Kaiser Permanente, Baylor Health Care System. Outpatient adverse event trigger tool. *IHI 2006* [serial online]. 2008.
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. **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.
22. **Kaafarani HMA, Rosen AK, Nebeker JR, et al.** Development of trigger tools for surveillance of adverse events in ambulatory surgery. Qual Saf Health Care. 2010;19:425–9.
23. **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.
24. **Murphy DR, Wu L, Thomas EJ, Forjuoh SN, Meyer AND, Singh H.** Electronic trigger-based intervention to reduce delays in diagnostic evaluation for cancer: a cluster randomized controlled trial. J Clin Oncol. 2015;33:3560–7.
25. **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.
26. **Garber JR, Cobin RH, Gharib H, et al.** Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. Endocr Pract. 2012;18:988–1028.
27. **Klein I, Danzi S.** Thyroid disease and the heart. Circulation. 2007;116:1725–35.
28. **Poppe K, Velkeniers B, Glinoer D.** The role of thyroid autoimmunity infertility and pregnancy. Nat Clin Pract Endocrinol Metab. 2008;4:394–405.
29. **Gulseren S, Gulseren L, Hekimsoy Z, Cetinay P, Ozen C, Tokatlioglu B.** Depression, anxiety, health-related quality of life, and disability in patients with overt and subclinical thyroid dysfunction. Arch Med Res. 2006;37:133–9.
30. **Khoury MJ, Becerra JE, D'Almada PJ.** Maternal thyroid disease and risk of birth defects in offspring: a population-based case–control study. Paediatr Perinat Epidemiol. 1989;3:402–20.
31. **Fihn SD, Francis J, Clancy C, et al.** Insights from advanced analytics at the veterans health administration. Health Aff. 2014;33:1203–11.
32. **Wang Y, Pakhomov S, Dale J, Chen E, Melton G.** Application of HL7/LOINC document ontology to a university-affiliated integrated health

system research clinical data repository. AMIA Summits Transl Sci Proc. 2014;2014:230–4.

33. **Murphy DR, Thomas EJ, Meyer AND, Singh H.** Development and validation of electronic health record-based triggers to detect delays in follow-up of abnormal lung imaging. Radiology. 2015;277:81–7.

34. **Joshi AK.** Natural language processing. Science. 1991;253:1242–9.

35. **Murphy DR, Meyer AND, 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. **Meyer AND, Murphy DR, Singh H.** Communicating findings of delayed diagnostic evaluation to primary care providers. J Am Board Fam Med. 2016;29:469–73.