Grant Final Report

Grant ID: 5R18HS017244

Using Electronic Records to Detect and Learn from Ambulatory Diagnostic Errors

Inclusive Dates: 09/30/08 - 09/30/10

Principal Investigator:
Eric J. Thomas, MD, MPH[*]

Team Members:
Hardeep Singh, MD, MPH[†]
Traber Davis Giardina, MA, MSW[†]
Myrna Khan, PhD[†]
Michael Reis, MD[‡]
Samuel Forjuoh, MD[‡]
Ricky O'Banon[‡]
Sonia Holleman, RA[‡]
Laura Wueste, RN[*]
Steven Kosmach, MSN, RN, CCRC[*]

Performing Organizations:
[*] Memorial Hermann Center for Healthcare Quality & Safety, and Division of General Medicine, Department of Medicine, University of Texas Medical School at Houston, Houston, Texas
[†] Michael E. DeBakey Veterans Affairs Medical Center and the Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, Texas
[‡] Department of Family & Community Medicine, Scott & White Healthcare Systems, Texas A&M Health Science Center, College of Medicine, Temple, Texas

Project Officer:
Kevin Chaney

Submitted to:
The Agency for Healthcare Research and Quality (AHRQ)
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Abstract

Purpose: Our main study objectives were to develop, refine, and test methods to detect diagnostic errors in primary care in several types of practice settings, (internal medicine and family practice, academic and nonacademic). Additionally, we sought to estimate the prevalence of diagnostic errors and describe the most common clinical conditions associated with these errors.

Scope: This study leveraged the electronic health record (EHR) infrastructure of two large healthcare systems to test whether electronic triggers can be used for large-scale measurement and surveillance of diagnostic errors in primary care.

Methods: We designed two types of electronic "triggers" (i.e., signals that prompt record review) to detect unusual visit patterns that may be associated with diagnostic error: Trigger 1: A primary care visit followed by unplanned hospitalization within 14 days; and Trigger 2: A primary care visit followed by 1 or more unscheduled primary care visits, urgent care visit, or emergency room visit within 14 days (excluding Trigger 1 visits). We applied these queries to electronic health record (EHR) repositories at two large integrated health systems to identify triggered visits between October 1, 2006 and September 30, 2007. Trained physician-reviewers determined presence or absence of diagnostic errors in random samples of triggered visits and control (non-triggered) visits.

Results: In 212,165 primary care visits at the two study sites, we found diagnostic errors in 141 of 674 Trigger 1-positive records and 36 of 669 Trigger 2-positive records (PPVs = 20.9% (95% CI,17.9%-24.0%) and 5.4%(95% CI,3.7%-7.1%), respectively). In contrast, only 13 of 614 control records contained evidence of diagnostic error (PPV = 2.1%(95% CI,0.1%-3.3%); $P < 0.01$). Prevalence of diagnostic error was estimated to be 2.8%. Pneumonia, decompensated heart failure, urinary tract infection and acute renal failure were most often missed.

Key Words: electronic health record, diagnostic errors, error surveillance, patient safety

Final Report

Purpose

The increasing use of electronic health records (EHRs) facilitates creation of health data repositories that contain longitudinal patient information in a far more integrated fashion than in previous record systems. In this study, we leveraged the EHR infrastructure of two large health care systems to test whether electronic triggers can be used for large-scale measurement and surveillance of diagnostic errors in primary care. Our main study objectives were to develop, refine, and test methods to detect diagnostic errors in primary care in several types of practice settings (internal medicine and family medicine, academic and nonacademic) in order to describe their prevalence as well as the most common clinical conditions associated with these errors. Such triggers could also be a useful way to detect and learn from diagnostic errors in future primary care practices.

Scope

Despite an increasing complexity and severity of illness seen in primary care, medical errors in this setting are understudied.[1,2] Because patients seek care for a diverse set of conditions through one or more relatively brief encounters, certain types of medical errors (such as diagnostic errors) are likely to be prevalent in the primary care setting.[1,3-7] Data from outpatient malpractice claims[2,8,9] consistently rank missed, delayed, and wrong diagnoses as the most common identifiable error. Other types of studies have also documented the magnitude and significance of diagnostic errors in outpatient settings.[10-15] Although these data point to an important problem, diagnostic errors have not been studied as well as other types of errors.[16,17] Not only are these errors difficult to identify,[9] but the fragmented outpatient environment also complicates tracking of diagnostic processes.

Better measurement and surveillance has been proposed as one potential solution to reduce the burden of diagnostic errors.[17,18] However, most existing error measurement methods (chart review, voluntary reporting, claims file review, etc.) are inefficient, biased, or unreliable.[19] In our preliminary work, we developed two computerized triggers to identify primary care patient records that may contain evidence of trainee-related diagnostic errors.[20] Triggers are signals that can alert providers to review the record to determine whether a patient safety event occurred.[21,22] Our triggers were based on the rationale that an unexpected hospitalization or return clinic visit after an initial primary care visit may indicate a missed diagnosis during the first visit. Although the positive predictive value (PPV) was modest (16.1% and 9.7% for the two triggers, respectively), it was comparable to that of previously designed electronic triggers to detect potential ambulatory medication errors.[23,24] Our previous findings were limited by a lack of generalizability outside of the study setting (a Veterans Affairs [VA] internal medicine trainee clinic) and by a significantly high rate of false positive cases that led to unnecessary record reviews.

Methods

Design and Settings

We designed electronic queries (triggers) to detect patterns of visits that could have been precipitated by diagnostic errors and applied these queries to all primary care visits at two large health systems over a one-year time period. We performed chart reviews to determine the presence or absence of diagnostic errors in triggered visits (i.e., those that met trigger criteria) and controls.

Both study sites provided longitudinal care in a relatively closed system and had integrated and well-established EHRs. Each site's electronic data repository contained administrative and clinical data extracted monthly from the EHR. At Site A, a large VA facility, about 35 full-time primary care providers (PCPs) saw patients in scheduled primary care follow-up clinic visits and "drop-in" unscheduled or urgent care clinic visits. Some staff physicians supervised residents and allied health providers. Emergency room (ER) staff provided care after hours and on weekends. Of approximately 50,000 patients, about 90% were assigned to staff for direct care and the remaining were distributed among residents. At Site B, a large private integrated health care system, 34 PCPs (family medicine physicians) provided care to nearly 50,000 patients in 4 community-based clinics and some supervised family practice residents. Clinic patients sought care at the ER of the parent hospital after-hours. To minimize after-hours attrition to hospitals outside our study settings, we did not apply the trigger to patients assigned to remote satellite clinics of the parent facilities. Both settings included ethnically and socioeconomically diverse patients from rural and urban areas (site A: Asian or Pacific Islander 2%, Black 38%, Hispanic < 1%, white 56%, unknown 3%; site B: Asian or Pacific Islander 1%, Black 8%, Hispanic 15%, white 55%, unknown 21%). Local IRB approval was obtained at both sites.

Trigger Application

Using a Structured Query Language (SQL) based program, we applied three queries to electronic repositories at the two sites to identify primary care index visits (defined as scheduled or unscheduled visits to physicians, physician-trainees, and allied health professionals that did not lead to immediate hospitalizations) between October 1, 2006 and September 30, 2007 that met the following criteria:

- Trigger 1: A primary care visit followed by an unplanned hospitalization that occurred between 24 hours and 14 days after the visit.

- Trigger 2: A primary care visit followed by 1 or more unscheduled primary care visits, an urgent care visit, or an ER visit that occurred within 14 days (excluding Trigger 1-positive index visits).

- Controls: All primary care visits from the study period that did not meet either trigger criterion.

The triggers above were based on our previous work and refined to improve their performance (Table 1).

Table 1. Rationale of Trigger Modifications from Previous Work[20]

Trigger Characteristics	Previous trigger	New trigger	Rationale
Time period	10 days	14 days	Previous findings showed that diagnostic errors continued to be discovered at the 10 day cut-off.
False positive generators	Did not account for planned hospitalizations or elective surgeries	Electronically excluded admissions related to Day surgery, Scheduled Ambulatory Admit, Pre-op clinics, Cardiology Invasive Procedure clinic	Lower rate of false positives triggered to increase PPV as well as efficiency of record reviews.
False positive generators	Did not account for admissions to units considered as " non acute"	Electronically excluded admissions to (e.g., long term and intermediate care, rehab)	Lower rate of false positives triggered to increase PPV as well as efficiency of record reviews.
False positive generators	Included all hospitalizations	Includes only hospitalizations electronically designated as "acute care" (e.g. acute medicine, acute surgery, acute mental health)	Lower rate of false positives triggered to increase PPV as well as efficiency of record reviews.

Our pilot reviews showed less well-defined associations between reasons for return visits and index visits when they were separated out by 3 or 4 weeks, and thus a 14-day cut-off was chosen. In addition, we attempted to remove records with false positive index visits, such as those associated with planned hospitalizations. The flexibility provided by SQL allowed for complex search parameters including dates, times, and locations of clinic visits and hospitalizations.

Data Collection and Error Assessment

We performed detailed chart reviews on random samples of control visits and visits that met both trigger criteria. If patients met a trigger criteria more than once, only one (earliest) index visit was included (unique record). Some records did not meet criteria for detailed review because the probability of an error at the index visit would be highly unlikely for one of the following reasons: no documentation of clinical notes, patients left clinic without being seen, patients saw a non-provider (e.g. nurse), or patients were advised admission for further evaluation but refused. These records were categorized as false positives.

Eligible records were randomly assigned to five trained reviewers, who were chief or senior internal medicine residents from outside institutions. Reviewers were blinded to the goals of the study and to the presence or absence of triggers. All reviewers underwent stringent quality

control and pilot testing procedures and reviewed 25-30 records each before they started collecting study data.

Each record was initially reviewed by two independent reviewers. The reviewers determined the presence or absence of a diagnostic error at the index visit by examining the EHR for details about the index and subsequent visits. A structured data collection form, adapted from our previous work,[20] was used to record documented signs and symptoms, diagnostic tests and procedures, clinician assessment and management decisions. We also collected documented information about encounters outside our study settings. To reduce hindsight bias,[25,26] we did not ask reviewers to assess patient harm. Reviewers were asked to make reasonable judgments of diagnostic performance based strictly on data either already available or easily available at the time of the index clinic visit. If reviewers disagreed on the presence of diagnostic error, a third, blinded reviewer made the final determination. Although we used an explicit definition from the literature[27] and a structured training process based upon our previous record review studies, the process involved implicit judgments. We computed Cohen's kappa (κ) to assess inter-reviewer agreement prior to the tiebreaker decision.

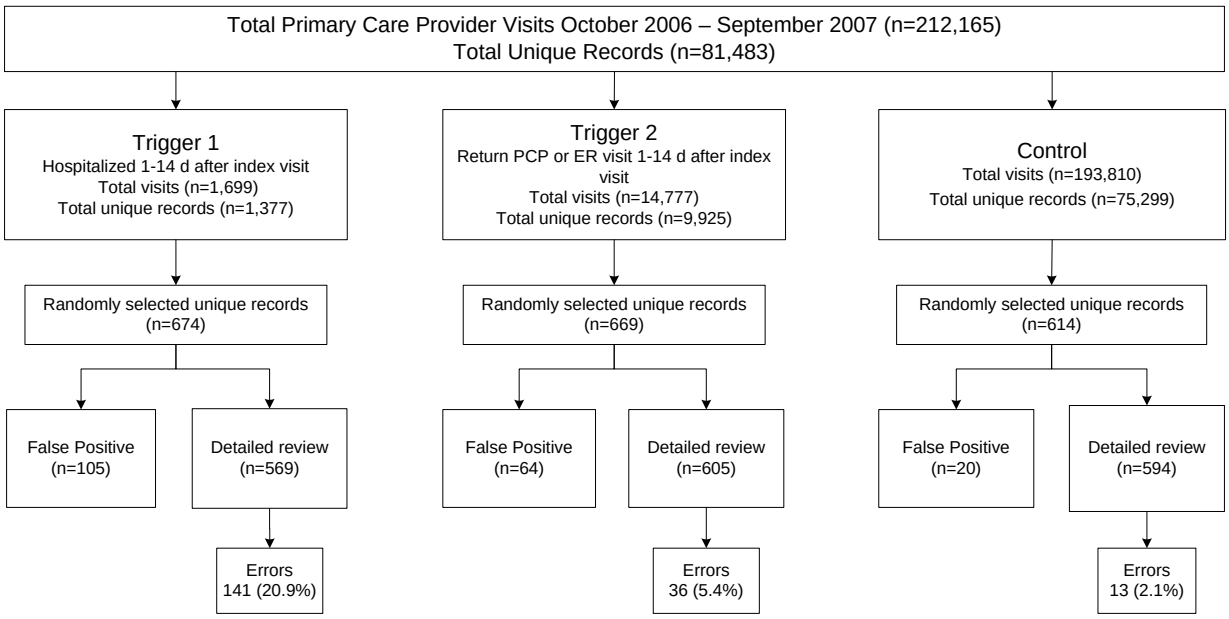
We calculated our sample size based on previous descriptive data and the anticipated reduction of false positive visits using our new triggers. Assuming a lower false positive rate for Trigger 1 (from 34.1% in our previous study to an estimated 20% using the refined trigger), we estimated a minimum sample size of 310 to show significant differences with 80% power at a *P* value of .05. Additionally, based on a previous study,[27] we sought to obtain a sample of at least 100 diagnostic errors in order to meaningfully describe the types of diseases/conditions associated. Given a previous PPV of 16.1% for Trigger 1, we thus estimated a sample size of at least 630 patient visits meeting criteria for Trigger 1. We reviewed comparable numbers of Trigger 2-positive and control records.

We calculated PPVs for both triggers and compared these with the PPV of controls. We also calculated false positive rates for each trigger and compared them to our previously used methods. We tallied the frequency of clinical conditions and presenting symptoms associated with the diagnostic errors that we discovered. Finally, we estimated the prevalence of diagnostic error in our study settings by extrapolating the PPVs to the larger cohort of patients from where our random samples were derived and then to the populations at both facilities. For this calculation, we assumed that record review would reveal the same PPV in the larger cohort of each random sample we chose to review at each site.

Results

We applied the triggers to 212,165 primary care visits (106,143 at Site A and 106,022 at Site B) that contained 81,483 unique records. We then randomly selected 674 unique records positive for Trigger 1, 669 unique records positive for Trigger 2, and 614 unique control records for review (Figure 1).

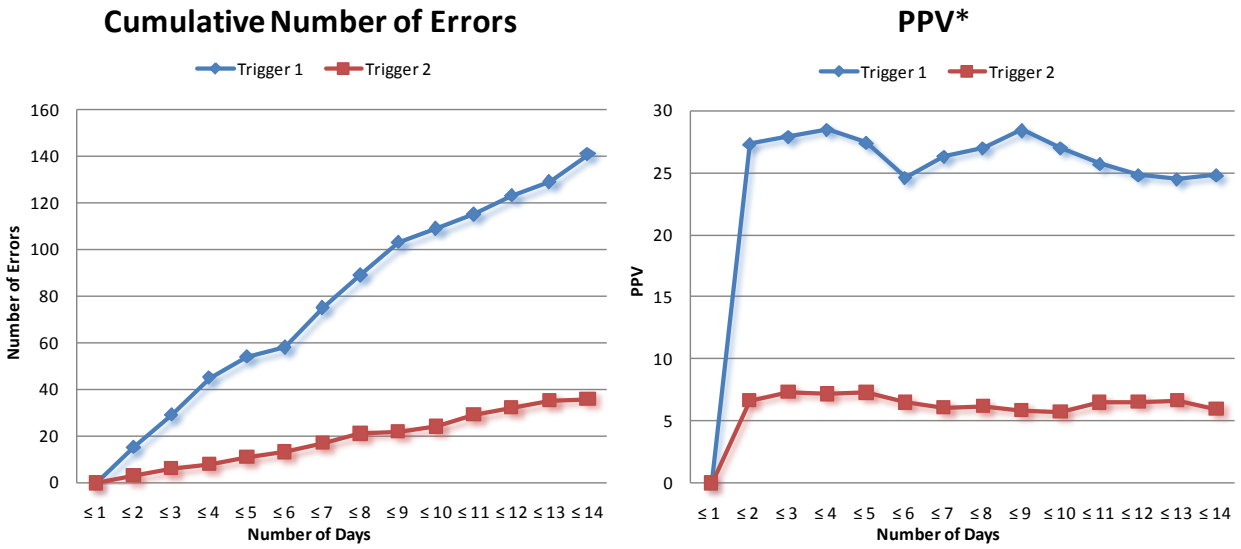
Figure 1. Study flowchart



On detailed review, diagnostic errors were found in 141 Trigger 1 positive records and 36 Trigger 2 positive records, yielding a PPV of 20.9% for Trigger 1 (95% CI,17.9%-24.0%) and 5.4% for Trigger 2 (95% CI,3.7%-7.1%). Errors were found in 13 control records. The control PPV of 2.1% (95% CI,0.1%-3.3%) was significantly lower than that of both Trigger 1 ($P < .001$) and Trigger 2 ($P = .002$). Trigger PPVs were equivalent between sites ($P = .9$ for both triggers).

Reviewers had fair agreement on presence or absence of a diagnostic error in triggered charts prior to the tiebreaker, $\kappa = 0.33$ (95% CI=0.27-0.39). Of 285 charts reviewed by a third reviewer, the third reviewer indicated the presence of a diagnostic error in 29.8%. Figure 2 shows the distribution of diagnostic errors in the inclusion sample over time interval (between index and return dates) for both Trigger 1 and Trigger 2 records. We found no significant trend for PPVs of either trigger through day 14 of the time interval.

Figure 2. Cumulative frequency of errors and PPV* corresponding to increasing time intervals between index and return date



The overall rate of false positives was 15.6% for Trigger 1 and 9.6% for Trigger 2. These rates were significantly lower than those found in our previous study (34.1% and 25.0%, respectively, $P < .001$ for both comparisons). False positive rates differed between sites, which may have been attributable to differences in information systems and documentation practices (Table 2).

Table 2. Expected Positive Predictive Values (PPVs) in an "Ideal" information system that could electronically exclude false positive records

Site/Trigger	Total number of errors n	PPV of diagnostic errors in current EHR system %	False positive rate* n	False positive rate* %	Expected PPV of diagnostic errors in future EHRs with capabilities to identify all false positives∞ %
Site A T1	95	20.9%	74	16.3%	25.0%
Site A T2	23	5.3%	50	11.6%	6.0%
Site A Control	11	2.7%	7	1.7%	2.7%
Site B T1	46	20.9%	31	14.1%	24.3%
Site B T2	13	5.5%	14	7.9%	5.8%
Site B Control	2	1.0%	13	5.9%	1.1%
Overall	190				

Because many false positives (no documentation, telephone or non-PCP encounters etc.) could potentially be coded accurately and identified electronically through information systems, we estimated the highest PPV potentially achievable by an ideal system (Table 2). Our estimates suggest that Trigger 1 PPV would increase from 20.9% to 24.8% if electronic exclusion of false positives was possible, although this did not reach statistical significance (*P*=.06 and .28 for Sites A and B respectively).

Table 3 shows diagnoses that were missed three or more times in Trigger 1 and Trigger 2 records with evidence of error. In some cases, more than one diagnosis was missed. Pneumonia (6.6%) and decompensated congestive heart failure (5.7%) were the two most common missed diagnoses, followed by urinary tract infection/pyelonephritis (5.2%) and acute renal failure (4.7%). Table 4 lists the most common chief complaints noted by the index visit providers for the patients with diagnostic errors. Common chief complaints included cough (often with another related symptom) (11.1%), followed by abdominal pain (8.9%) and shortness of breath (7.4%). Notably, 25 (13.1%) patients did not have any chief complaint and were either established patients following up on chronic medical issues or patients new to the system.

Table 3. Frequencies of most commonly missed diagnoses in 190 unique patient charts (both triggered and controls)

Missed or Delayed Diagnoses	Total (n=212)*	
Pneumonia	14	(6.6%)
Decompensated Congestive Heart Failure	12	(5.7%)
Urinary Tract Infection/Pyelonephritis	11	(5.2%)
Acute Renal Failure	10	(4.7%)
Spinal cord compression or stenosis	9	(4.2%)
Symptomatic anemia	9	(4.2%)
Angina/Myocardial Infarction/Acute Coronary Syndrome	7	(3.3%)
Cancer (new malignancy)	7	(3.3%)
Medication side effect	7	(3.3%)
Complicated Peripheral Vascular Disease and/or Arterial Occlusion	6	(2.8%)
Transient Ischemic Attack/Stroke	5	(2.4%)
Cellulitis	5	(2.4%)
Metastasis of known cancer	5	(2.4%)
Osteomyelitis	4	(1.9%)
Pulmonary Embolism	4	(1.9%)
Deep Venous Thrombosis	4	(1.9%)
Hypertension	4	(1.9%)
Cirrhosis and portal hypertension	3	(1.4%)
Atrial Fibrillation (New onset)	3	(1.4%)
Hyperglycemia	3	(1.4%)
Otitis	3	(1.4%)
Bacteremia	3	(1.4%)
Electrolyte disturbance	3	(1.4%)
Leukemia/lymphoma	3	(1.4%)
Renal Calculus	3	(1.4%)
Hypotension	3	(1.4%)

n>190 because in some cases two diagnoses were missed

Table 4. Frequencies of one or more chief complaints in 190 unique patients with diagnostic error

Chief Complaint		Total[†]	n (%)
Cough	7	**21**	**(11.1%)**
Cough + chest pain	1		
Cough +congestion	4		
Cough +fever	5		
Cough +medication refill	1		
Cough +shortness of breath	2		
Cough +flu like symptoms	1		
Abdominal pain		**17**	**(8.9%)**
Follow-up of routine medical issues or no chief complaint		**16**	**(8.4%)**
Shortness of breath	12	**14**	**(7.4%)**
Shortness of breath +cough	1		
Shortness of breath +generalized weakness, +cough, and dizziness	1		
Establish care		**9**	**(4.7%)**
Back pain	7	**8**	**(4.2%)**
Back pain +leg pain	1		
Dizziness	3	**6**	**(3.2%)**
Dizziness +generalized weakness	1		
Dizziness +vomiting	1		
Dizziness +toe soreness	1		
Chest pain		**5**	**(2.6%)**
Foot pain	4	**5**	**(2.6%)**
Foot pain +swelling	1		
Leg edema/swelling		**5**	**(2.6%)**
Leg pain	2	**5**	**(2.6%)**
Leg pain +swelling	3		
Fatigue		**4**	**(2.1%)**
Knee pain and/or swelling		**4**	**(2.1%)**

[†] Percentages refer to the proportion of these complaints in the total number of diagnostic error cases (190). Some complaints not listed occurred at low frequencies.

To estimate the prevalence of diagnostic errors, we extrapolated the PPVs for triggered visits and controls to the larger population of all primary care patient records at each site. At Site A we estimated 163 errors in 781 Trigger 1 patients, 295 errors in 5566 Trigger 2 patients, and 1056 errors in 39,118 controls. At Site B we estimated 125 errors in 596 patients, 240 errors in 4359 Trigger 2 patients, and 362 in 36,181 controls. Across sites, we predicted an overall diagnostic error prevalence of 2.8% (2241 errors in 81,483 unique records).

Discussion

We tested whether computerized "trigger" methods could be used to identify primary care visit records with higher than expected potential to contain evidence of diagnostic errors. Our electronic triggers had a higher PPV than any other known method of large scale diagnostic error detection. Based on the rates of diagnostic error in our samples, we estimated the overall prevalence of diagnostic error to be 2.8% of primary care records in our study population. Most diagnostic errors involved common conditions such as pneumonia and congestive heart failure, and they were associated with chief complaints that are relatively typical in primary care. However, a significant number of errors occurred in patients without any documented chief complaints, highlighting the challenges of diagnostic error prevention efforts in primary care.

Previously used methods to study diagnostic errors have notable limitations: autopsies are now infrequent,[28] self-report methods (eg, surveys) are prone to bias, and malpractice claims, although useful, shed light on a narrow and non-representative spectrum of medical error.[2,14,19,29,30] Medical record review is often considered the gold standard for studying diagnostic errors, but it is time consuming and costly, with a relatively low yield.[31] Our study provides a methodology that is far more efficient than conducting random record reviews. We identified errors that were more consequential than many "routine" errors,[14] but also less recognizable than obvious, critical errors that might be detected by self-reports.[32] Additionally, our findings differ qualitatively from those resulting from other methods. For instance, studies that examine malpractice claims exclusively suggest that the condition most frequently associated with outpatient diagnostic error is cancer.[2,29]

Our study is one of the largest to address diagnostic error in routine practice (ie, outside of malpractice claims) and the first to estimate diagnostic error prevalence in primary care. It also triangulates a growing body of knowledge on the prevalence and significance of diagnostic errors to give a more robust picture of the problem. Errors in our study were most often related to common presenting symptoms and common diseases (versus rare or unusual diseases or presenting complaints). Although some of these may have been challenging to diagnose initially, diseases such as pneumonia, decompensated CHF, and urinary tract infection do not typically represent "diagnostic dillemas" of medicine. Given high patient volumes, rushed office visits, and multiple competing demands for PCPs' attention, our findings are not surprising and call for a multi-pronged intervention effort for error prevention.[18] Provider-focused interventions might help but are unlikely to effect significant change without a corresponding system-level approach to improve the diagnostic process.[33] Future efforts to help reform primary care practice and facilitate team-based care could offload the current heavy burden faced by current PCPs and significantly reduce these errors.[18] Meanwhile, research in this area should inform the design of new triggers and other surveillance methods to predict and prevent these errors.

Inadequate error feedback mechanisms and low numbers of event reports that relate to diagnostic errors are major impediments to achieving patient safety.[32,34] Existing peer review systems to assess PCPs' competence are also of questionable efficacy.[35] We thus recommend that primary care reform initiatives consider using our triggers, especially Trigger 1, as an approach to conduct routine monitoring and surveillance for diagnostic errors. This approach should be akin to patient safety initiatives related to electronic surveillance for medication errors and nosocomial infections.[22,36,37] Although comprehensive data repositories such as those we used may not be available to many primary care practices, future HIT integration and health information exchange data will facilitate techniques proposed in our study. For instance, these

techniques can be used by patient centered medical homes for oversight and performance monitoring with feedback. Detecting and avoiding preventable hospitalizations will also be necessary to become successful Accountable Care Organizations and to manage bundled payments, two key features of the Patient Protection and Affordable Care Act of 2010.[38] Although further refinement in EHR systems might help reduce false positives, we believe Trigger 1 currently identifies diagnostic errors at sufficiently high rates to be useful. A review of triggered cases by medical home teams to ensure that all contributing factors are identified – not just those related to provider performance – will foster interdisciplinary quality improvement.

Our triggers might be generalizable because our queries contained information that is available in almost all EHRs. Furthermore, they were tested in two large systems where two types of primary care specialties practiced. However, our methods may not apply to primary care practices that are not part of integrated health care systems. Practices unable to dedicate reviewer time for diagnostic error confirmation may not also find them useful. Another use of our triggers will be as a tool for patient navigators – typically nurses who help patients with complex, chronic diseases navigate the healthcare system. Site B is using our methodology to develop triggers to help navigators identify cancer patients with delayed diagnoses or therapies. Co-Investigator Dr. Singh is working with Site A to improve the test result notification process (based upon the American Journal of Medicine paper noted below). Site A will also be interested in using trigger 1 in routine quality measurement activities, although discussions are preliminary.

While the kappa agreement between our reviewers was not very high, similar values have been observed in other error studies.[39] The moderate level of agreement suggests that our estimate of the prevalence of diagnostic errors is relatively imprecise. We also might have underestimated error rates because we relied solely on patterns of primary care visits to trigger record reviews.[40] Some misdiagnosed patients, unknown to us, might have sought care in another system or recovered without any further care. Finally, we may have missed errors that occurred beyond 14 days as well as those not detected by our triggers.

In summary, our findings demonstrate that EHR-based trigger methods can enable more meaningful measurement and surveillance of diagnostic errors in primary care. Most errors involved common conditions seen in primary care and given the volume of primary care visits in the US, an estimated prevalence rate of about 2.8% entails a substantial patient safety risk. Primary care reform initiatives should redesign delivery systems as well as implement techniques for active error surveillance. Improved measurement and detection methods could provide useful feedback about errors to frontline providers and promote prevention-related learning.

Reference List

1. Gandhi TK, Lee TH. Patient Safety beyond the Hospital. *New England Journal of Medicine.* 2010;363:1001-1003.

2. Phillips R, Jr., Bartholomew LA, Dovey SM, Fryer GE, Jr., Miyoshi TJ, Green LA. Learning from malpractice claims about negligent, adverse events in primary care in the United States. *Qual Saf Health Care.* 2004;13:121-126.

3. Bhasale A. The wrong diagnosis: identifying causes of potentially adverse events in general practice using incident monitoring. *Fam Pract.* 1998;15:308-318.

4. Elder NC, Dovey SM. Classification of medical errors and preventable adverse events in primary care: a synthesis of the literature. *J Fam Pract.* 2002;51:927-932.

5. Woods DM, Thomas EJ, Holl JL, Weiss KB, Brennan TA. Ambulatory care adverse events and preventable adverse events leading to a hospital admission. *Quality and Safety in Healthcare.* 2007.

6. Wachter RM. Is ambulatory patient safety just like hospital safety, only without the "stat"? *Ann Intern Med.* 2006;145:547-549.

7. Bhasale AL, Miller GC, Reid SE, Britt HC. Analysing potential harm in Australian general practice: an incident-monitoring study. *Med J Aust.* 1998;169:73-76.

8. Chandra A, Nundy S, Seabury SA. The growth of physician medical malpractice payments: evidence from the National Practitioner Data Bank. *Health Aff (Millwood).* 2005;Suppl Web Exclusives:W5-240-W5-249.

9. Graber M. Diagnostic errors in medicine: a case of neglect. *Jt Comm J Qual Patient Saf.* 2005;31:106-113.

10. Casalino LP, Dunham D, Chin MH et al. Frequency of Failure to Inform Patients of Clinically Significant Outpatient Test Results. *Arch Intern Med.* 2009;169:1123-1129.

11. Schiff GD, Hasan O, Kim S et al. Diagnostic Error in Medicine: Analysis of 583 Physician-Reported Errors. *Arch Intern Med.* 2009;169:1881-1887.

12. 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;169:1578-1586.

13. 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;28:3307-3315.

14. Singh H, Thomas EJ, Wilson L et al. Errors of diagnosis in pediatric practice: a multisite survey. *Pediatrics.* 2010;126:70-79.

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;123:238-244.

16. Schiff GD, Kim S, Abrams R, Cosby K, Lambert B, Elstein AS. Diagnosing diagnosis errors: Lessons from a multi-institutional collaborative project. *Advances in Patient Safety.* 2005.

17. Wachter RM. Why diagnostic errors don't get any respect--and what can be done about them. *Health Aff (Millwood).* 2010;29:1605-1610.

18. Singh H, Graber M. Reducing Diagnostic Error through Medical Home-Based Primary Care Reform. *JAMA.* 2010.

19. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med.* 2003;18:61-67.

20. Singh H, Thomas E, Khan M, Petersen L. Identifying diagnostic errors in primary care using an electronic screening algorithm. *Arch Intern Med.* 2007;167:302-308.

21. Agency for Healthcare Research and Quality. Patient Safety Network. Agency for Healthcare Research and Quality . 2008. http://psnet.ahrq.gov/

22. 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;15:184-190.

23. Field TS, Gurwitz JH, Harrold LR et al. Strategies for detecting adverse drug events among older persons in the ambulatory setting. *J Am Med Inform Assoc.* 2004;11:492-498.

24. Honigman B, Lee J, Rothschild J et al. Using computerized data to identify adverse drug events in outpatients. *J Am Med Inform Assoc.* 2001;8:254-266.

25. Fischhoff B. Hindsight not equal to foresight: the effect of outcome knowledge on judgment under uncertainty 1975. *Qual Saf Health Care.* 2003;12:304-311.

26. McNutt, R., Abrams, R., and Hasler, S. Diagnosing diagnostic mistakes: AHRQ web morbidity and mortality rounds. AHRQ . 2005. 9-25-2006.

27. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med.* 2005;165:1493-1499.

28. Shojania KG, Burton EC. The Vanishing Nonforensic Autopsy. *New England Journal of Medicine.* 2009;358:873-875.

29. 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;145:488-496.

30. McAbee GN, Donn SM, Mendelson RA, McDonnell WM, Gonzalez JL, Ake JK. Medical Diagnoses Commonly Associated With Pediatric Malpractice Lawsuits in the United States. *Pediatrics.* 2008;122:e1282-e1286.

31. Thomas EJ, Studdert D.M., Burstin HR et al. Incidence and types of adverse events and negligent care in Utah and Colorado in 1992. *Medical Care.* 2000;38:261-271.

32. Sevdalis N, Jacklin R, Arora S, Vincent CA, Thomson RG. Diagnostic error in a national incident reporting system in the UK. *J Eval Clin Pract.* 2010.

33. Singh, H, Davis, T, Petersen, LA, Wilson, LA, Dismukes, K, Bhagwath, G, and Thomas, EJ. A New Paradigm to Understand Errors of Diagnosis in the Primary Care Setting. 2010. Under Review.

34. Schiff GD. Minimizing Diagnostic Error: The Importance of Follow-up and Feedback. *The American Journal of Medicine.* 2008;121:S38-S42.

35. United States General Accounting Office. VA Health Care: Physician Peer Review Identifies Quality of Care Problems but Actions to Address Them Are Limited. Report to Ranking Minority Member, Committee on Veterans' Affairs, U.S. Senate. GAO/HEHS-95-121. 1995. Washington, D.C., United States. General Accounting Office.

36. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *J Am Med Inform Assoc.* 2003;10:115-128.

37. Jha AK, Kuperman GJ, Teich JM et al. Identifying adverse drug events: development of a computer-based monitor and comparison with chart review and stimulated voluntary report. *J Am Med Inform Assoc.* 1998;5:305-314.

38. Orszag PR, Emanuel EJ. Health care reform and cost control. *N Engl J Med.* 2010;363:601-603.

39. Thomas EJ, Lipsitz SR, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. *Ann Intern Med.* 2002;136:812-816.

40. Shojania K. The Elephant of Patient Safety: What You See Depends on How You Look. *Joint Commission Journal on Quality and Patient Safety.* 2010;36:399-401.

List of Publications and Products

Singh, H., Thomas, E., Sittig, D., et al. Notification of Laboratory Test Results in an Electronic Medical Record: Do Any Safety Concerns Remain? *American Journal of Medicine*. 2010;123(3):238-244.

We're planning to submit three additional manuscripts for publication.