Journal of the American Medical Informatics Association, 28(10), 2021, 2202–2211
doi: 10.1093/jamia/ocab121
Advance Access Publication Date: 19 July 2021
Research and Applications

Research and Applications

Validation of an electronic trigger to measure missed diagnosis of stroke in emergency departments

Viralkumar Vaghani,[1] Li Wei,[1] Umair Mushtaq,[1] Dean F. Sittig ,[2] Andrea Bradford ,[1] and Hardeep Singh [1]

[1]Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, Texas, USA and [2]University of Texas—Memorial Hermann Center for Healthcare Quality & Safety, School of Biomedical Informatics, University of Texas Health Science Center, Houston, Texas, USA

Corresponding Author : Hardeep Singh, MD, MPH, Michael E. DeBakey Veterans Affairs Medical Center (MEDVAMC), Center for Innovations in Quality, Effectiveness and Safety (152), 2002 Holcombe Boulevard 152, Houston, TX 77030 USA (hardeeps@bcm.edu)

Received 5 January 2021; Revised 26 May 2021; Editorial Decision 30 May 2021; Accepted 23 June 2021

ABSTRACT

Objective: Diagnostic errors are major contributors to preventable patient harm. We validated the use of an electronic health record (EHR)-based trigger (e-trigger) to measure missed opportunities in stroke diagnosis in emergency departments (EDs).

Methods: Using two frameworks, the Safer Dx Trigger Tools Framework and the Symptom-disease Pair Analysis of Diagnostic Error Framework, we applied a symptom–disease pair-based e-trigger to identify patients hospitalized for stroke who, in the preceding 30 days, were discharged from the ED with benign headache or dizziness diagnoses. The algorithm was applied to Veteran Affairs National Corporate Data Warehouse on patients seen between 1/1/2016 and 12/31/2017. Trained reviewers evaluated medical records for presence/absence of missed opportunities in stroke diagnosis and stroke-related red-flags, risk factors, neurological examination, and clinical interventions. Reviewers also estimated quality of clinical documentation at the index ED visit.

Results: We applied the e-trigger to 7,752,326 unique patients and identified 46,931 stroke-related admissions, of which 398 records were flagged as trigger-positive and reviewed. Of these, 124 had missed opportunities (positive predictive value for "missed" = 31.2%), 93 (23.4%) had no missed opportunity (non-missed), 162 (40.7%) were miscoded, and 19 (4.7%) were inconclusive. Reviewer agreement was high (87.3%, Cohen's kappa = 0.81). Compared to the non-missed group, the missed group had more stroke risk factors (mean 3.2 vs 2.6), red flags (mean 0.5 vs 0.2), and a higher rate of inadequate documentation (66.9% vs 28.0%).

Conclusion: In a large national EHR repository, a symptom–disease pair-based e-trigger identified missed diagnoses of stroke with a modest positive predictive value, underscoring the need for chart review validation procedures to identify diagnostic errors in large data sets.

Key words: diagnostic errors, health care quality improvement, health services research, patient safety, stroke

INTRODUCTION

Diagnostic error has emerged as a major contributor to preventable patient harm.[1] A recent National Academies report, *Improving Diagnosis in Healthcare*, concluded that diagnostic errors are difficult to measure, and new measurement approaches are needed to understand and reduce diagnostic errors.[2] Emergency departments (EDs) are a high-risk setting for diagnostic errors due to the time sensitivity and severity of emergency conditions, fast paced

environments, frequent interruptions, incomplete or unreliable data gathering, and high workload.[3–6] However, there are no standardized mechanisms to detect and analyze diagnostic errors in the ED. Factors that complicate measurement include the evolving nature of diagnosis and a lack of clinical data that represent the diagnostic process accurately. Progress in measuring and reducing diagnostic errors remains slow due to limitations in methods to identify diagnostic errors and learn from them.[7]

The use of electronic health records (EHRs) in large health care systems results in vast repositories of clinical data that can be mined to detect evidence of possible safety events, including missed opportunities in diagnosis. However, only a few health care systems actively leverage large data sets to identify opportunities to improve diagnosis. Recently, two conceptual measurement frameworks have been proposed for using large data sets to detect patterns or signals suggestive of missed diagnosis. The Symptom–disease Pair Analysis of Diagnostic Error (SPADE) framework[8] describes a method for using large administrative data sets (billing, insurance claims) to map frequently missed diagnoses (eg, stroke) to one or more previously documented high-risk symptoms (eg, headache). This framework, which uses linked symptom–disease dyads based on biological relationships, has been applied to detect missed myocardial infarction (MI), cerebrovascular accident (CVA), and appendicitis.[9–11] The SPADE approach relies on linkages within administrative datasets but does not use additional clinical review information to confirm events. Conversely, the Safer Dx Trigger Tools Framework[12] uses the concept of an "electronic trigger" (e-trigger) to identify a set of patient records for review that may involve a diagnostic error. It focuses on the analysis of clinical data to identify patterns of care via documentation that may signal a potential diagnostic error (eg, a primary care office visit followed by an unplanned hospitalization), followed by manual reviews of triggered records to confirm presence or absence of diagnostic error. This framework helps to identify areas for learning and improvement. These e-trigger algorithms are now gaining traction in the study of patient safety. Although simple triggers have been applied for more than a decade,[13–17] development of sophisticated triggers for diagnostic errors and delays is relatively new.[13,18–22]

Stroke is commonly misdiagnosed.[23–29] A meta-analysis of 23 studies reported that approximately 9% of cerebrovascular events are missed at initial ED presentation.[30] A recent study using the SPADE framework to analyze administrative data estimated that 1.2% of hospital admissions for stroke were preceded by possible misdiagnosis of headache or dizziness at an initial ED visit.[10] This and other studies suggest that neurological symptoms may be important symptom predictors of missed stroke.[10,31,32] However, administrative data are insufficient to confirm preventable events and have limited potential to inform actionable strategies to reduce error. We therefore aimed to modify the symptom–disease pair approach by querying clinical data sets followed by clinician review to confirm and characterize missed diagnostic opportunities.

In this study, we developed and tested a new e-trigger to detect possible cases of missed stroke in the ED based on two symptom–disease pairs (stroke preceded by previous diagnosis of headache or dizziness). Our objectives were to apply the Safer Dx Trigger Tools Framework to develop an e-trigger for each symptom–disease pair and to examine performance of these e-triggers for identifying missed opportunities in stroke diagnosis. Validating this approach through confirmatory record reviews could advance the science of diagnostic error measurement and reduction.

MATERIALS AND METHODS

Study Setting

We conducted a study of ED encounters at 130 Veterans Affairs (VA) health care facilities. Being an integrated health system that uses a comprehensive homegrown EHR, the VA has longitudinal patient data to track a patient's diagnostic journey over time. To enable e-trigger development and implementation, we accessed the VA's national Corporate Data Warehouse hosted on the VA Informatics and Computing Infrastructure (VINCI).[33] This database contains EHR data from all VA facilities across the US, serving over 9 million veterans annually. The study was approved by the institutional review boards of the local VA facility and academic affiliate institution.

Trigger Development

We developed a stroke e-trigger using the 7 steps of the Safer Dx Trigger Tools Framework[12]: (1) identify and prioritize diagnostic error of interest; (2) operationally define criteria to detect diagnostic error; (3) determine potential electronic data sources; (4) construct an e-trigger algorithm to obtain a cohort of interest; (5) test e-trigger on a data source and review medical records; (6) assess e-trigger algorithm performance; and (7) iteratively refine e-triggers to improve trigger performance. For steps 1–4, we used the previously published "look-back" SPADE method (including all diagnostic codes used in this method)[8,10] to design an electronic query to identify all patients from the VINCI database who met the following criteria:

1. Admitted to a VA hospital with an admission diagnosis of ischemic stroke, hemorrhagic stroke, or transient ischemic attack (TIA) based on ICD-10 codes (Supplementary Appendix 1, Table 1);
2. Date of admission between January 1, 2016 and December 31, 2017;
3. Presence of a "treat-and-release" VA ED visit with ED discharge diagnosis as headache or dizziness within 30 days prior to the admission to an affiliated VA hospital (ie, index ED visit). Benign headache (eg, tension headache) and dizziness (eg, vertigo) diagnoses were identified using ICD-10 codes (Supplementary Appendix 2).

Data Collection Procedures

Next, we tested the e-trigger and assessed its performance (steps 5 and 6 of the framework). After running the query, a trained physician reviewed each triggered record to identify missed opportunities for diagnosis of stroke or TIA at the index ED visit. One physician served as the primary reviewer for all records, and a second physician-reviewer independently evaluated a random subset of 20% of records (n = 80) to assess the reliability of judgements related to missed opportunities for diagnosis. Reviewers used the Compensation and Pension Record Interchange (CAPRI), a VA application that enables national VA EHR access to review the patient's medical record. Prior to the study, the two physician-reviewers received multiple training sessions and pilot-tested the review process on 25 records using operational definitions and procedures developed for the study. Pilot cases were discussed within the research team to develop consensus and to resolve disagreements. Physician-reviewers used a structured electronic data collection instrument to standardize data collection and minimize data entry errors.

Table 1. Baseline Characteristics of Patients in Missed and Non-missed Groups

Factors	Overall (n = 217)	Missed (n = 124)	Non-missed (n = 93)	*p*-value
Age (mean, SD)	68.1 (11.6)	69.3 (11.3)	66.4 (11.8)	.69
Sex				
Male	209 (96.3%)	122 (98.4%)	87 (93.5%)	.077
Female	8 (3.7%)	2 (1.6%)	6 (6.5%)	
Race				
African American	66 (30.4%)	40 (32.3%)	26 (28.0%)	.291
American Indian	2 (0.9%)	0	2 (2.2%)	
Asian	2 (0.9%)	1 (0.8%)	1 (1.1%)	
White	140 (64.5%)	78 (62.9%)	62 (66.7%)	
Other	1 (0.5%)	0	1 (1.1%)	
Decline to Answer	6 (2.8%)	5 (4.0%)	1 (1.1%)	
Provider				
Physician	166 (76.5%)	94 (75.8%)	72 (77.4%)	.975
NP	16 (7.4%)	9 (7.3%)	7(7.5%)	
PA	14 (6.5%)	8 (6.5%)	6 (6.5%)	
Others	21 (9.7%)	13 (10.5%)	8 (8.6%)	
Presenting Symptom at ED index visit				
Headache	87 (40.1%)	40 (32.3%)	47 (50.5%)	.012
Dizziness	109 (50.2%)	73 (58.9%)	36 (38.7%)	
Both	21 (9.7%)	11 (8.9%)	10 (10.8%)	
Top 5 discharge diagnoses noted at ED index visit				
	Headache (23.5%) Dizziness (22.1%) Vertigo (10.1%) Lightheadedness (5.5%) Migraine (5.1%)	Dizziness (28.2%) Headache (12.1%) Vertigo (11.2%) Lightheadedness (8.9%) Benign paroxysmal positional vertigo (4.1%)	Headache (38.7%) Dizziness (13.9%) Vertigo (8.6%) Migraine (6.4%) Dizziness, headache (5.3%)	N/A
Type of Stroke Diagnosis				
Ischemic stroke	146 (67.3%)	88 (71.0%)	58 (62.4%)	.272
Hemorrhagic stroke	27 (12.4%)	16 (12.9%)	11 (11.8%)	
Transient ischemic attack (TIA)	42 (19.4%)	19 (15.3%)	23 (24.7%)	
Unspecified	2 (0.9%)	1 (0.8%)	1 (1.1%)	
Stroke Risk Predictors Score				
Dawson TIA score[a]				
Low risk (<5.4)	87 (40.1%)	44 (35.5%)	43 (46.2%)	.110
High risk (>=5.4)	130 (59.9%)	80 (64.5%)	50 (53.8%)	
ROSIER Scale[b]				
Low risk (<=0)	179 (82.5%)	97 (78.2%)	82 (88.2%)	.056
High risk (>0)	38 (17.5%)	27 (21.8%)	11 (11.8%)	
ABCD2 score[c]				
Low (1–3)	113 (52.1%)	71 (57.3%)	42 (45.2%)	.180
Moderate (4–5)	101 (46.5%)	51 (41.1%)	50 (53.8%)	
High (6–7)	3 (1.4%)	2 (1.6%)	1 (1.1%)	

Abbreviations: SD, standard deviation; TIA, transient ischemic attack;
[a]TIA recognition tool.
[b]Acute stroke recognition tool.
[c]Stroke prediction tool following TIA.

Reviewers searched all the documentation in the record, including notes from other providers, such as nursing and triage notes at the index ED visit, for presence/absence of "red flags" (eg, speech abnormalities, limb weakness) and risk factors associated with stroke or TIA (eg, hypertension, hyperlipidemia).[32,34–36] Reviewers also searched for clinical actions in response to red flags, such as neurological consultation or appropriate imaging. We used information based on American Heart Association/American Stroke Association guidelines[37] to develop our approach to evaluate for missed opportunities when stroke-related red-flags were present. We determined that all patients with red flags at index ED encounter warrant a diagnostic workup. In the absence of red flags at index ED encounter, we used current clinical literature on headache or dizziness assessment to develop our approach on how patients with multiple stroke risk factors should be worked up.[38–44] We defined a *missed opportunity in diagnosis* (MOD) when no additional action or evaluation was undertaken despite stroke-related red flags. We also defined a *potential missed opportunity in diagnosis* (P-MOD) when red flags were absent at the ED visit, but the patient had multiple stroke risk factors, and the neurological examination was abnormal or incomplete. These instances were deemed as opportunities for improvement. We defined *absence of a missed opportunity in diagnosis* (No-MOD) if a patient received appropriate clinical actions in response to stroke-related red flags/risk factors or when a patient with multiple stroke risk factors had a complete normal neurological examination or pursued discharged against medical advice. Essentially,

this implied that nothing different could have been done to pursue a correct or more timely diagnosis given the context of their clinical situation. MODs, P-MODs, and No-MODs were only identified among patients with a definitive diagnosis of stroke or TIA. This approach to identify diagnostic error is similar to that used in previous work.[45–47]

During pilot reviews, we found several errors in coding of diagnoses. These were related either to absence of any new confirmed stroke/TIA findings in the clinical documentation at discharge or absence of benign diagnosis of headache or dizziness at the ED visit. Coding errors included instances of both application of incorrect codes by the prior stroke SPADE algorithm (eg, use of code "impacted cerumen") as well as instances of misapplication in the medical record itself (eg, patient admitted with possible stroke but found to have none upon additional evaluation, or patient did not have any headache or dizziness in the ED). We categorized all such records as coding errors. We labeled records as inconclusive when we could not confidently confirm the diagnosis of stroke/TIA at hospital discharge (eg, when documented as "likely TIA").

In addition to information about red flags, risk factors, neurological examination,[48] and clinical interventions, reviewers collected data about the patient's age, sex, race/ethnicity, and the type of provider at the index ED visit (attending physician, trainee, physician assistant, or nurse practitioner). We calculated the Dawson TIA score,[49] ROSIER scale,[50] and ABCD2[51] score to quantify the risk of stroke or TIA prediction based on EHR documentation at the index ED visit. Finally, we collected information about the quality of clinical documentation using a validated Physician Documentation Quality Instrument (PDQI-9)[52] and a single-item note impression ("please rate the overall quality of this note" with 5-point Likert scale from very poor to excellent), which has been used in prior studies to evaluate the quality of EHR notes.[52,53] We used the ED provider's main clinical note to assess the quality of documentation. Disagreements between reviewers were discussed and resolved by consensus prior to analysis.

Statistical Analysis

After reaching consensus on discordant judgements between reviewers, we compared patients in the "missed" group (records with MODs or P-MODs) to those in the "non-missed" group to assess differences in demographic characteristics and comorbidities. We computed Cohen's kappa to assess interrater reliability for determination of missed opportunities. We used *t*-tests, Fisher's exact tests, and chi-squared analyses to assess between-group differences. We used descriptive statistics to describe commonly missed red flags, stroke risk factors, neurological examination, and documentation quality for both groups. We used SPSS 22 (IBM, Armonk, NY) for our analyses.

RESULTS

Trigger Performance

We applied the e-trigger to 7,752,326 unique patients and identified 46,931 stroke-related admissions. Three hundred ninety-eight records were flagged as trigger-positive and reviewed. Of these, 124 (31.2%) patients were determined to experience either a MOD or P-MOD ("missed" group). Conversely, 93 (23.4%) patients (non-missed group) did not have any evidence of missed opportunity in stroke diagnosis. Additionally, 162 (40.7%) patients were miscoded (eg, had a stroke diagnosis carried forward from the past) and 19 (4.7%) patients were inconclusive (Figure 1); these were omitted from subsequent analyses of patient characteristics. Most common coding errors were related to absence of any new confirmed stroke/TIA findings in the clinical documentation at discharge (82.7%) or absence of benign diagnosis of headache or dizziness at the ED visit (16.1%). Reviewer agreement for determining missed opportunities was good (87.3%, Cohen's kappa = 0.81). Of 46,931 total stroke admissions for 2016–17, the trigger identified 124 patients who experienced any missed opportunity of stroke diagnosis, yielding a minimum 0.3% frequency of missed stroke in this population. The positive predictive value (PPV) of the e-trigger for detecting any missed opportunity of stroke diagnosis was 31.2% (124 of 398). The PPV was 10.8% with more stringent criteria that included only confirmed cases of missed diagnosis (ie, MODs alone). After removal of miscoded patients, the e-trigger PPV improved from 31.2% to 52.5%.

Patient Characteristics

Table 1 summarizes overall and group-level characteristics of patients in missed and non-missed groups. Dizziness was a more frequent presenting symptom during ED index visits in the missed group (n = 73, 58.9%), whereas headache was more frequent in the non-missed group (n = 47, 50.5%). Groups were similar in age, sex, race, provider type, type of stroke diagnosis, and stroke risk predictor scores (Table 1).

Red Flags and Risk Factors

Most patients (n = 183; 84.3%) in both groups presented with multiple stroke risk factors (mean number of risk factors = 3; SD = 1.2), but the mean number of risk factors was greater in the missed group (mean 3.2 vs 2.6 for non-missed group, $p < 0.001$). Hypertension, hyperlipidemia, and diabetes were more frequent risk factors in the missed group (Table 2). Just over a quarter of patients across both groups (n = 57, 26.2%) presented with one or more stroke-related red flags (mean number of red flags = 0.4; SD = 0.7). The mean number of red flags was greater in the missed (0.5) than the non-missed group (0.2), $p = 0.002$. In the missed group, sudden loss of balance was the most common missed red flag, followed by visual field defect and unilateral limb weakness (Table 2).

Neurological Examination and Follow-up

Components of neurological physical examination documented in missed and non-missed groups are listed in Table 3. The least frequently performed (or documented) exam components were reflex testing, coordination testing, nystagmus testing, and gait testing. Most patients in the missed group had multiple (≥3) risk factors (n = 99; 79.8%), or at least 1 red-flag sign or symptom (n = 44; 35.5%), but essential components of neurological examination (mental status, cranial nerves, motor exam, sensory exam, reflex testing, coordination testing, and gait testing) were seldom performed or documented in this high-risk subset (n = 115; 92.7%). Of 44 patients who had at least one red flag, specific imaging tests such as CT scan were ordered for fewer than two-thirds (n = 28; 63.6%); none received an appropriate follow-up imaging test (MRI) or neurology consultation.

Documentation quality

Table 4 shows comparisons of clinical note documentation quality between missed and nonmissed groups at the index ED visit.[52] Compared to documentation in the missed group, all aspects of

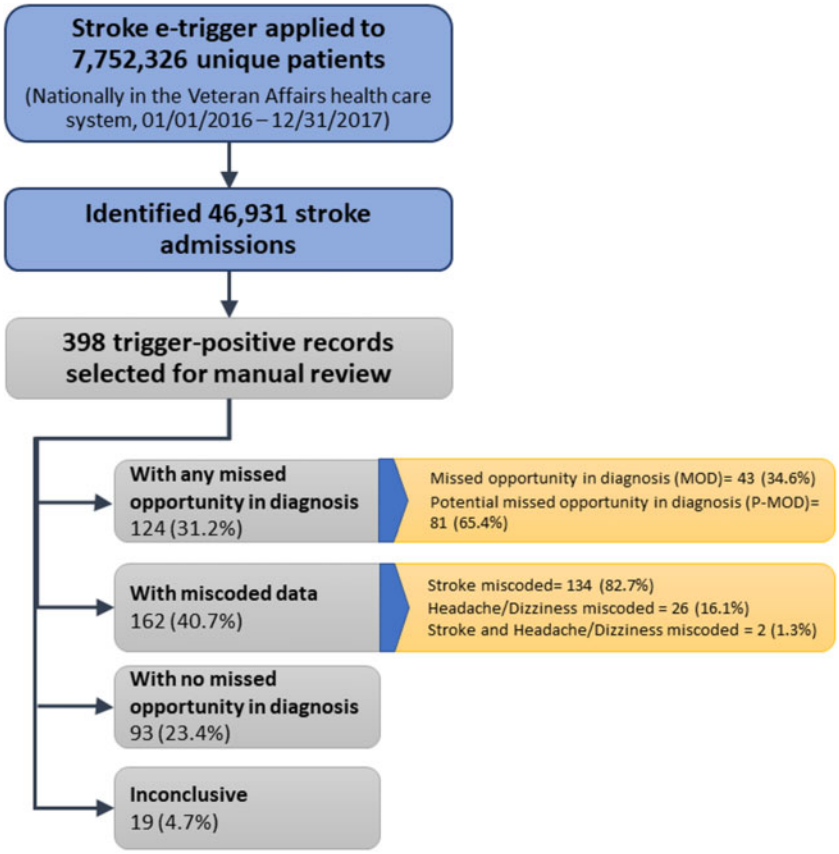

Figure 1. Trigger Validation Process Flow.

Table 2. Risk Factors and Red Flags in Index Visit in Missed and Non-missed Groups

Factors	Overall (n = 217)	Missed (n = 124)	Non-missed (n = 93)	*p*-value
Presence of stroke risk factors				
Hypertension	187 (86.2%)	112 (90.3%)	75 (80.6%)	.041
Hyperlipidemia	150 (69.1%)	93 (75.0%)	57 (61.3%)	.031
Diabetes	106 (48.8%)	75 (60.5%)	31 (33.3%)	<.001
Hx of stroke/TIA	71 (32.7%)	47 (37.9%)	24 (25.8%)	.060
Current smokers	65 (30.0%)	38 (30.6%)	27 (29.0%)	.797
History of atrial fibrillation	24 (11.1%)	15 (12.1%)	9 (9.7%)	.574
History of carotid stenosis	15 (6.9%)	10 (8.1%)	5 (5.4%)	.440
History of aneurysm	13 (6.0%)	7 (5.6%)	6 (6.5%)	.804
History of recent head trauma	14 (6.5%)	5 (4.0%)	9 (9.7%)	.094
Presence of stroke red flags				
Sudden loss of balance/coordination	17 (7.8%)	15 (12.1%)	2 (2.2%)	.007
Visual field defect	21 (9.7%)	14 (11.3%)	7 (7.5%)	.353
Unilateral limb weakness	15 (6.9%)	11 (8.9%)	4 (4.3%)	.189
Speech abnormalities	9 (4.1%)	6 (4.8%)	3 (3.2%)	.410
Unilateral facial weakness	15 (6.9%)	6 (4.8%)	1 (1.1%)	.120
Sudden severe headache	6 (2.8%)	6 (4.8%)	0	.033
Sudden confusion	3 (1.4%)	2 (1.6%)	1 (1.6%)	.607

documentation quality were more likely to be rated highly (4 or 5 on a 5-point scale) in the nonmissed group. The frequency of documentation rated overall as fair, poor, or very poor was higher in the missed group in comparison to the non-missed group. Examples of clinical note characteristics for each criterion of the PDQI-9 instrument can be found in Supplementary Appendix 1, Table 2.

DISCUSSION

We tested an e-trigger to detect missed opportunities for diagnosis of TIA or stroke in the ED based on the presence of two symptom–disease pairs. We found the e-trigger to have a modest positive predictive value (PPV) of 31.2% for missed or potentially missed diagnosis. When extrapolated to the overall large number of patients

Table 3. Neurological Examination in Index Visit in Missed and Non-missed Groups

Type of Examination	Overall (n = 217)	Missed (n = 124)	Non-missed (n = 93)	*p*-value
Mini-mental state examination				
Performed/noted	213 (98.2%)	120 (96.8%)	93 (100%)	.104
Not performed/noted	4 (1.8%)	4 (3.2%)	0	
Cranial nerves				
Performed/noted	142 (65.4%)	71 (57.3%)	71 (76.3%)	.003
Not performed/noted	75 (34.6%)	53 (42.7%)	22 (23.7%)	
Nystagmus				
Performed/noted	36 (16.6%)	25 (20.2%)	11 (11.8%)	.102
Not performed/noted	181 (83.4%)	99 (79.8%)	82 (88.2%)	
Motor exam				
Performed/noted	133 (61.3%)	68 (54.8%)	65 (69.9%)	.024
Not performed/noted	84 (38.7%)	56 (45.2%)	28 (30.1%)	
Sensory exam				
Performed/noted	101 (46.5%)	48 (38.7%)	53 (57.0%)	.008
Not performed/noted	116 (53.5%)	76 (61.3%)	40 (43.0%)	
Reflex testing				
Performed/noted	43 (19.8%)	19 (15.3%)	24 (25.8%)	.055
Not performed/noted	174 (80.2%)	105 (84.7%)	69 (74.2%)	
Coordination testing				
Performed/noted	48 (22.1%)	20 (16.1%)	28 (30.1%)	.014
Not performed/noted	169 (77.9%)	104 (83.9%)	65 (69.9%)	
Gait testing				
Performed/noted	67 (30.9%)	33 (26.6%)	34 (36.6%)	.117
Not performed/noted	150 (69.1%)	91 (73.4%)	59 (63.4%)	
Dix-Hallpike maneuver				
Performed/noted	5 (2.3%)	2 (1.6%)	3 (3.2%)	.653
Not performed/noted	212 (97.7%)	122 (98.4%)	90 (96.8%)	
Head Impulse, Nystagmus, Test of Skew (HINTS) examination				
Performed/noted	1 (0.5%)	0	1 (1.1%)	.429
Not performed/noted	216 (99.5%)	124 (100.0%)	92 (98.9%)	

who present to the ED with symptoms or multiple risk factors of stroke, this represents a potentially valuable approach to identify missed opportunities retrospectively and help design interventions to improve clinical practice.

We found evidence of two frequently occurring problems in the diagnostic evaluation of stroke that may be actionable to improve diagnostic safety. First, documentation quality was low in about half of all trigger-positive cases we reviewed. However, inadequate documentation of the history and physical examination was especially common in cases where we found a lack of action upon stroke-related red flags and risk factors. Clinicians also did not document all essential components of neurological examination in patients presenting with red flags. Second, there was inadequate follow-up action (such as specific neurologic maneuvers, ordering appropriate imaging tests, or initiating referrals) by clinicians when patients presented with red flags or multiple stroke risk factors. For instance, neurologic bedside tests that were more specific to rule out stroke, such as nystagmus testing or HINTS exam were rarely performed. Our findings highlight the need to improve stroke-related diagnostic processes[54–56] and to implement clinical documentation guidelines[57] that focus on capturing key clinical data while minimizing burden for clinicians.

Previously,[10] the frequency of missed strokes in patients with headache or dizziness was estimated to be 1.2% using administrative or claims data.[58,59] Our study raises concern for reliance on diagnostic coding alone for determination of error, as about 4 out of 10 records were miscoded for stroke, dizziness, or headache. Furthermore, just under a quarter of the triggered sample did not have missed or delayed opportunity; as such, the presence of the symptom–disease pair may not be by itself a reliable indicator of diagnostic error. Methodologies (including SPADE) that rely only on administrative billing and coding data to measure safety without any confirmatory medical chart reviews thus may not estimate the problem accurately. Alternatively, the Safer Dx Trigger tool methods use large-scale EHR data to identify a highly selective cohort and allow for rigorous confirmatory chart reviews of this narrow subset to identify diagnostic errors. For instance, for every 100 charts identified as "diagnostic errors" via stroke SPADE methodology, the Safer Dx methodology will only confirm 31 (ie, 69 will be false positive). Our findings underscore the need for confirmatory, manual record review procedures for algorithms used to identify missed opportunities and other care gaps from EHR data. Given that medical record reviews are generally considered a reference standard for determining diagnostic error,[60] they may inform more accurate error frequency estimates. These findings also have implications for studies that rely on large scale EHR data for more automated measurement of quality and safety, including those using big data, machine learning, and artificial intelligence-based approaches.

Emerging evidence from malpractice claims suggests that diagnostic errors related to stroke diagnosis in ED are increasing over time.[29,30,61] Most missed opportunities in our cohort resulted from breakdowns of processes related to the patient–provider encounter (ie, information gathered during history and physical examination). There may be certain sociodemographic or clinical factors that

Table 4. Comparison of Documentation Quality at Index Visit in Missed and Non-missed Groups

Description	Overall (n = 217)	Missed group (n = 124)	Non-missed group (n = 93)	*P* value
Up to date: The note contains the most recent test results and recommendations.				
Not at all, slightly, moderately	85 (39.2%)	62 (50.0%)	23 (24.7%)	<.001
Very, extremely	132 (60.8%)	62 (50.0%)	70 (75.3%)	
Accurate: The note is true. It is free of incorrect information.				
Not at all, slightly, moderately	94 (43.3%)	73 (58.9%)	21 (22.6%)	<.001
Very, extremely	123 (56.7%)	51 (41.1%)	72 (77.4%)	
Thorough: The note is complete and documents all of the issues of importance to the patient.				
Not at all, slightly, moderately	178 (82.0%)	112 (90.3%)	66 (71.0%)	<.001
Very, extremely	39 (18.0%)	12 (9.7%)	27 (29.0%)	
Useful: The note is extremely relevant, providing valuable information and/or analysis.				
Not at all, slightly, moderately	101 (46.5%)	74 (59.7%)	27 (29.0%)	<.001
Very, extremely	116 (53.5%)	50 (40.3%)	66 (71.0%)	
Organized: The note is well-formed and structured in a way that helps the reader understand the patient's clinical course.				
Not at all, slightly, moderately	92 (42.4%)	65 (52.4%)	27 (29.0%)	.001
Very, extremely	125 (57.6%)	59 (47.6%)	66 (71.0%)	
Comprehensible: The note is clear, without ambiguity or sections that are difficult to understand.				
Not at all, slightly, moderately	82 (37.8%)	62 (50.0%)	20 (21.5%)	<.001
Very, extremely	135 (62.2%)	62 (50.0%)	73 (78.5%)	
Succinct: The note is brief, to the point, and without redundancy.				
Not at all, slightly, moderately	77 (35.5%)	59 (47.6%)	18 (19.4%)	<.001
Very, extremely	140 (64.5%)	65 (52.4%)	75 (80.6%)	
Synthesized: The note reflects the author's understanding of the patient's status and ability to develop a plan of care.				
Not at all, slightly, moderately	82 (37.8%)	63 (50.8%)	19 (20.4%)	<.001
Very, extremely	135 (62.2%)	61 (49.2%)	74 (79.6%)	
Internally consistent: No part of the note ignores or contradicts any other part.				
Not at all, slightly, moderately	78 (35.9%)	58 (46.8%)	20 (21.5%)	<.001
Very, extremely	139 (64.1%)	66 (53.2%)	73 (78.5%)	
Overall documentation quality*				
Fair, poor, very poor	109 (50.2%)	83 (66.9%)	26 (28.0%)	<.001
Excellent, good	108 (49.8%)	41 (33.1%)	67 (72.0%)	

make certain patients vulnerable to missed opportunities, but we were unable to determine that based on our study design and sample size. Mixed methods approaches could reveal additional insights that are not available from record reviews. Clinicians in ED settings make complex diagnostic decisions during brief encounters. High cognitive load, patient acuity, and decision density are factors that have been implicated in diagnostic errors in the ED setting.[55,62] Clinicians in our study noted red flags and risk factors for stroke, but this did not reliably lead to further diagnostic evaluation. Previous studies have found similar gaps during patient–provider encounters in the ED setting for other acute conditions.[46,47,63] Work-system factors such as cognitive workload, frequent interruptions, and time pressures encountered by providers likely contribute to this problem.[3,4]

In addition to interventions focused on improving the ED work-system, our study suggests the need for technology-based interventions to support ED clinicians in the diagnostic evaluation of stroke. Clinical guidelines such as Early Management of Patients with Acute Stroke from the American Heart Association/American Stroke Association[37] could be incorporated into EHR-based clinical documentation templates, which might help clinicians in rapid identification of red-flag symptoms and appropriate diagnostic evaluation, including appropriate imaging.[64] Stroke risk prediction algorithms could be used within EHRs to stratify patients into low, intermediate, or high risk categories based on history and physical examination at the initial encounter.[49–51] Use of EHR-based clinical decision support tools that inform diagnostic decision-making could be beneficial in early identification of TIA/stroke.[65] Alternatively, portable video-oculography goggles with HINTS examination can be used in real time in patients with dizziness, which helps to differentiate stroke from vestibular disorders.[66] Efforts are needed to bolster implementation of existing tools into ED practice.[67]

Our e-trigger offers an efficient method to detect missed opportunities for stroke diagnosis for patients who present with dizziness or headache in the ED setting. For instance, the computer algorithm scanned 46,931 stroke-related admissions to identify just 398 for human review, hence doing most of the work. Furthermore, it would be impossible for a human to review that many admissions. Currently, health systems are not using any sophisticated detection methods for diagnostic error and are finding them occasionally and passively through rudimentary incident reporting systems. While our trigger performed modestly, PPVs to identify events of interest have been traditionally lower in the area of patient safety.[68–74] With additional testing, this approach could be applied to other EHR data warehouses to retrospectively identify diagnostic errors for learning and improvement purposes. E-trigger enhanced review procedures overcome several limitations of other safety measurement methods.[8] For example, highly selective record reviews make error detection efforts far more efficient than random record reviews or reliance on incident reporting systems, which have revealed very little data to address diagnostic errors.[75] E-triggers could strengthen patient safety improvements efforts in health systems with limited resources and competing demands on quality measurement.

The e-trigger had a higher PPV and interrater reliability to detect missed opportunities than e-triggers used in several previous studies on diagnostic errors.[1,45,63,76,77] Further refinement of diagnostic

codes and removal of admission–discharge diagnosis discrepancies at the time of trigger development stage could substantially improve the PPV of the e-trigger. Reliance on standardized structured data codes (eg, ICD codes), and Structured Query Language for data mining makes this trigger highly portable between health systems. Potential future uses for this e-trigger include safety monitoring with feedback of diagnostic performance to providers to help them better calibrate.[78,79] However, currently these methods are mostly useful for additional review and learning and should not be used as quality measures for accountability purposes. We also recommend additional validation and performance evaluation at non-VA settings before implementation of such methods in practice. Because very few methods focus on ED diagnostic errors, future efforts using similar e-triggers could be useful to identify and understand contributory factors associated with missed opportunities of diagnosis in ED setting. Use of free-text data using natural language processing could potentially improve the efficiency and yield of the e-trigger and reduce the effort required for confirmatory record reviews. Future informatics applications could be used proactively to either detect or prevent errors before they cause harm. In fact, e-triggers have recently shown effectiveness in proactively detecting delays in care after an abnormal test result suspicious for cancer,[18–20,22,80–82] kidney failure,[83,84] infection,[83] and thyroid dysfunction.[85] Similar, more prospective approaches could be applied to reduce diagnostic error risk for patients presenting to the ED.

There are several limitations of our study. Our study population was within a single national health care system in the US that treats a predominantly male population, and findings might not generalize to other health care institutions. Study methods may overlook patients who receive care at urgent care or EDs not affiliated with VA, died before returning for a second visit for stroke, sought care elsewhere (eg, outside the VA), or did not ultimately receive a diagnosis of stroke (either because their symptoms were too subtle, were missed even upon return, or the patient did not seek follow-up). Thus, it is likely that the study underestimates the number of people who may be missed. In addition, it was not feasible to do facility-level analysis to assess quality of coding or note documentation due to a relatively small number of patients at each of the 130 VA facilities. However, the study included 130 health care facilities, and similar diagnostic errors have been described elsewhere.[1,45–47,63,76,86] Retrospective chart reviews rely on EHR documentation by clinicians, which may not always represent the actual care delivered.[87] However, determination of many safety-related outcomes, including diagnostic errors, often relies on medical record reviews to evaluate transpired events.[88–90] Such reviews also introduce the possibility of hindsight bias affecting the reviewer's clinical judgement.[91] To minimize this, we designed a data collection instrument based on objective criteria (American Heart Association/American Stroke Association Guidelines) and published literature to avoid relying solely on subjective clinical judgment.

CONCLUSION

E-triggers based on symptom–disease pairs for high-risk conditions are a potentially valuable approach to identify missed opportunities for diagnosis. Given the high frequency of coding errors and cases without errors, our findings underscore the need to validate the output of algorithmically identified diagnostic errors in large data sets. Nevertheless, we find that a significant number of patients who presented with red flags and multiple stroke risk factors did not receive appropriate diagnostic evaluation. Our study calls for multifaceted interventions to address contributory factors related to the ED work-system and clinicians' diagnostic performance in order to reduce harm from missed stroke diagnosis.

ACKNOWLEDGMENTS

None.

FUNDING

Dr Singh is funded by the Veterans Affairs (VA) Health Services Research and Development Service (HSR&D) (IIR-17-127 and the Presidential Early Career Award for Scientists and Engineers USA 14-274), the Agency for Healthcare Research and Quality (R01HS27363 and R18HS017820), the Houston VA HSR&D Center for Innovations in Quality, Effectiveness, and Safety (CIN 13-413), and the Gordon and Betty Moore Foundation. Dr Sittig was supported in part by the Agency for Health Care Research and Quality (P30HS024459).

AUTHOR CONTRIBUTIONS

Study concept and design: VV, HS, DFS.

Acquisition of data: VV, LW.

Statistical analysis: VV.

Analysis and interpretation of data: VV, HS.

Drafting of the manuscript: VV.

Critical revision of the manuscript for important intellectual content: VV, DFS, AB, UM, LW, HS.

Administrative, technical or material support: VV, DFS, AB, UM, LW, HS.

Study supervision: VV, DFS, HS.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

DATA AVAILABILITY STATEMENT

The data underlying this article cannot be shared publicly.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Singh H, Meyer AND, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014; 23 (9): 727–31.
2. Ball JR, Balogh E. Improving diagnosis in health care: highlights of a report from the national academies of sciences, engineering, and medicine. *Ann Intern Med* 2016; 164 (1): 59–61.
3. Berg LM, Kallberg A-S, Goransson KE, Ostergren J, Florin J, Ehrenberg A. Interruptions in emergency department work: an observational and interview study. *BMJ Qual Saf* 2013; 22 (8): 656–63.
4. Hamden K, Jeanmonod D, Gualtieri D, Jeanmonod R. Comparison of resident and mid-level provider productivity in a high-acuity emergency department setting. *Emerg Med J* 2014; 31 (3): 216–9.
5. Schnapp BH, Sun JE, Kim JL, Strayer RJ, Shah KH. Cognitive error in an academic emergency department. *Diagnosis (Berlin, Germany)* 2018; 5 (3): 135–42.

6. Stiell A, Forster AJ, Stiell IG, van Walraven C. Prevalence of information gaps in the emergency department and the effect on patient outcomes. *CMAJ* 2003; 169 (10): 1023–8.
7. Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. *BMJ Qual Saf* 2017; 26 (6): 484–94.
8. Liberman AL, Newman-Toker DE. Symptom–disease Pair Analysis of Diagnostic Error (SPADE): a conceptual framework and methodological approach for unearthing misdiagnosis-related harms using big data. *BMJ Qual Saf* 2018; 27 (7): 557–66.
9. Moy E, Barrett M, Coffey R, Hines AL, Newman-Toker DE. Missed diagnoses of acute myocardial infarction in the emergency department: variation by patient and facility characteristics. *Diagnosis (Berlin, Germany)* 2015; 2 (1): 29–40.
10. Newman-Toker DE, Moy E, Valente E, Coffey R, Hines AL. Missed diagnosis of stroke in the emergency department: a cross-sectional analysis of a large population-based sample. *Diagnosis (Berlin, Ger)* 2014; 1 (2): 155–66.
11. Mahajan P, Basu T, Pai C-W, *et al.* Factors Associated With Potentially Missed Diagnosis of Appendicitis in the Emergency Department. *JAMA Netw Open* 2020; 3 (3): e200612. doi:10.1001/jamanetworkopen.2020.0612.
12. Murphy DR, Meyer AN, Sittig DF, Meeks DW, Thomas EJ, Singh H. Application of electronic trigger tools to identify targets for improving diagnostic safety. *BMJ Qual Saf* 2019; 28 (2): 151–9.
13. Kennerly DA, Kudyakov R, da Graca B, *et al.* Characterization of adverse events detected in a large health care delivery system using an enhanced global trigger tool over a five-year interval. *Health Serv Res* 2014; 49 (5): 1407–25.
14. Naessens JM, O'Byrne TJ, Johnson MG, Vansuch MB, McGlone CM, Huddleston JM. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Heal Care J Int Soc Qual Heal Care* 2010; 22 (4): 266–74.
15. Chapman SM, Fitzsimons J, Davey N, Lachman P. Prevalence and severity of patient harm in a sample of UK-hospitalised children detected by the Paediatric Trigger Tool. *BMJ Open* 2014; 4 (7): e005066.
16. Mull HJ, Nebeker JR, Shimada SL, Kaafarani HMA, Rivard PE, Rosen AK. Consensus building for development of outpatient adverse drug event triggers. *J Patient Saf* 2011; 7 (2): 66–71.
17. 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 (5): 425–9.
18. Murphy DR, Meyer AND, Vaghani V, *et al.* Development and validation of trigger algorithms to identify delays in diagnostic evaluation of gastroenterological cancer. *Clin Gastroenterol Hepatol* 2018; 16 (1): 90–8.
19. Murphy DR, Meyer AND, Vaghani V, *et al.* Application of electronic algorithms to improve diagnostic evaluation for bladder cancer. *Appl Clin Inform* 2017; 8 (1): 279–90.
20. 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 (3): 613–20.
21. 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 (1): 8–16.
22. Murphy DR, Meyer AND, Vaghani V, *et al.* Electronic triggers to identify delays in follow-up of mammography: harnessing the power of big data in health care. *J Am Coll Radiol* 2018; 15 (2): 287–95.
23. Venkat A, Cappelen-Smith C, Askar S, *et al.* Factors associated with stroke misdiagnosis in the emergency department: a retrospective case-control study. *Neuroepidemiology* 2018; 51 (3–4): 123–7.
24. Richoz B, Hugli O, Dami F, Carron P-N, Faouzi M, Michel P. Acute stroke chameleons in a university hospital: risk factors, circumstances, and outcomes. *Neurology* 2015; 85 (6): 505–11.
25. Kowalski RG, Claassen J, Kreiter KT, *et al.* Initial misdiagnosis and outcome after subarachnoid hemorrhage. *JAMA* 2004; 291 (7): 866–9.
26. Vermeulen MJ, Schull MJ. Missed diagnosis of subarachnoid hemorrhage in the emergency department. *Stroke* 2007; 38 (4): 1216–21.
27. Arch AE, Weisman DC, Coca S, Nystrom KV, Wira CR 3rd, Schindler JL. Missed ischemic stroke diagnosis in the emergency department by emergency medicine and neurology services. *Stroke* 2016; 47 (3): 668–73.
28. Schiff GD, Hasan O, Kim S, *et al.* Diagnostic error in medicine: analysis of 583 physician-reported errors. *Arch Intern Med* 2009; 169 (20): 1881–7.
29. Kachalia A, Gandhi TK, Puopolo AL, *et al.* Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. *Ann Emerg Med* 2007; 49 (2): 196–205.
30. Tarnutzer AA, Lee S-H, Robinson KA, Wang Z, Edlow JA, Newman-Toker DE. ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: a meta-analysis. *Neurology* 2017; 88 (15): 1468–77.
31. Rosenman MB, Oh E, Richards CT, *et al.* Risk of stroke after emergency department visits for neurologic complaints. *Neurol Clin Pract* 2020; 10 (2): 106–14.
32. Kim AS, Fullerton HJ, Johnston SC. Risk of vascular events in emergency department patients discharged home with diagnosis of dizziness or vertigo. *Ann Emerg Med* 2011; 57 (1): 34–41.
33. Fihn SD, Francis J, Clancy C, *et al.* Insights from advanced analytics at the Veterans Health Administration. *Health Aff (Millwood)* 2014; 33 (7): 1203–11.
34. Aigner A, Grittner U, Rolfs A, Norrving B, Siegerink B, Busch MA. Contribution of established stroke risk factors to the burden of stroke in young adults. *Stroke* 2017; 48 (7): 1744–51.
35. Guzik A, Bushnell C. Stroke epidemiology and risk factor management. *Continuum (Minneap Minn)* 2017; 23 (1, Cerebrovascular Disease): 15–39.
36. Hankey GJ. Stroke. *Lancet (London, England)* 2017; 389 (10069): 641–54.
37. Powers WJ, Rabinstein AA, Ackerson T, *et al.* 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2018; 49 (3): e46–110.
38. Davenport R. Acute headache in the emergency department. *J Neurol Neurosurg Psychiatry* 2002; 72 (Suppl 2): ii33.
39. Hainer BL, Matheson EM. Approach to acute headache in adults. *Am Fam Physician* 2013; 87 (10): 682–7.
40. Kerber KA. Vertigo and dizziness in the emergency department. *Emerg Med Clin North Am* 2009; 27 (1): 39–50, viii.
41. Muncie HL, Sirmans SM, James E. Dizziness: approach to evaluation and management. *Am Fam Physician* 2017; 95 (3): 154–62.
42. Liberman AL, Prabhakaran S. Stroke chameleons and stroke mimics in the emergency department. *Curr Neurol Neurosci Rep* 2017; 17 (2): 15.
43. Wallace EJC, Liberman AL. Diagnostic challenges in outpatient stroke: stroke chameleons and atypical stroke syndromes. *Neuropsychiatr Dis Treat* 2021; 17: 1469–80.
44. Wu V, Beyea MM, Simpson MT, Beyea JA. Standardizing your approach to dizziness and vertigo. *J Fam Pract* 2018; 67 (8): 490–2, 495, 498.
45. Singh H, Giardina TD, Forjuoh SN, *et al.* Electronic health record-based surveillance of diagnostic errors in primary care. *BMJ Qual Saf* 2012; 21 (2): 93–100.
46. Singh H, Giardina TD, Meyer AND, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings. *JAMA Intern Med* 2013; 173 (6): 418–25.
47. Bhise V, Meyer AND, Singh H, *et al.* Errors in diagnosis of spinal epidural abscesses in the era of electronic health records. *Am J Med* 2017; 130 (8): 975–81.
48. UC San Diego School of Medicine. Practical Guide to Clinical Medicine. https://meded.ucsd.edu/clinicalmed/neuro2.htmAccessed April 14, 2020
49. Dawson J, Lamb KE, Quinn TJ, *et al.* A recognition tool for transient ischaemic attack. *QJM* 2009; 102 (1): 43–9.
50. Nor AM, Davis J, Sen B, *et al.* The Recognition of Stroke in the Emergency Room (ROSIER) scale: development and validation of a stroke recognition instrument. *Lancet Neurol* 2005; 4 (11): 727–34.

51. Johnston SC, Rothwell PM, Nguyen-Huynh MN, *et al.* Validation and refinement of scores to predict very early stroke risk after transient ischaemic attack. *Lancet (London, England)*. 2007; 369 (9558): 283–92.
52. Stetson PD, Bakken S, Wrenn JO, Siegler EL. Assessing electronic note quality using the Physician Documentation Quality Instrument (PDQI-9). *Appl Clin Inform* 2012; 3 (2): 164–74.
53. Edwards ST, Neri PM, Volk LA, Schiff GD, Bates DW. Association of note quality and quality of care: a cross-sectional study. *BMJ Qual Saf* 2014; 23 (5): 406–13.
54. Clark BW, Derakhshan A, Desai SV. Diagnostic errors and the bedside clinical examination. *Med Clin North Am* 2018; 102 (3): 453–64.
55. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005; 165 (13): 1493–9.
56. Verghese A, Charlton B, Kassirer JP, Ramsey M, Ioannidis JPA. Inadequacies of physical examination as a cause of medical errors and adverse events: a collection of vignettes. *Am J Med* 2015; 128 (12): 1322–4. e3.
57. Payne TH, Corley S, Cullen TA, *et al.* Report of the AMIA EHR-2020 Task Force on the status and future direction of EHRs. *J Am Med Inform Assoc* 2015; 22 (5): 1102–10.
58. Bates DW, O'Neil AC, Petersen LA, Lee TH, Brennan TA. Evaluation of screening criteria for adverse events in medical patients. *Med Care* 1995; 33 (5): 452–62.
59. Zhan C, Miller MR. Administrative data-based patient safety research: a critical review. *Qual Saf Health Care*. 2003; 12 (Suppl 2): ii58–63.
60. Thomas EJ, Studdert DM, Burstin HR, *et al.* Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000; 38 (3): 261–71.
61. Glick TH, Cranberg LD, Hanscom RB, Sato L. Neurologic patient safety: an in-depth study of malpractice claims. *Neurology* 2005; 65 (8): 1284–6.
62. Croskerry P, Sinclair D. Emergency medicine: a practice prone to error? *CJEM* 2001; 3 (4): 271–6.
63. Medford-Davis L, Park E, Shlamovitz G, Suliburk J, Meyer AND, Singh H. Diagnostic errors related to acute abdominal pain in the emergency department. *Emerg Med J* 2016; 33 (4): 253–9.
64. Wang SV, Rogers JR, Jin Y, Bates DW, Fischer MA. Use of electronic healthcare records to identify complex patients with atrial fibrillation for targeted intervention. *J Am Med Inform Assoc* 2017; 24 (2): 339–44.
65. Ranta A, Dovey S, Weatherall M, O'Dea D, Gommans J, Tilyard M. Cluster randomized controlled trial of TIA electronic decision support in primary care. *Neurology* 2015; 84 (15): 1545–51.
66. Newman-Toker DE, Saber Tehrani AS, Mantokoudis G. Quantitative video-oculography to help diagnose stroke in acute vertigo and dizziness: toward an ECG for the eyes. *Stroke* 2013; 44 (4): 1158–61.
67. Graber ML, Sorensen AV, Biswas J, *et al.* Developing checklists to prevent diagnostic error in Emergency Room settings. *Diagnosis (Berlin, Ger)* 2014; 1 (3): 223–31.
68. Silva M das DG, Martins MAP, Viana L de G, *et al.* Evaluation of accuracy of IHI Trigger Tool in identifying adverse drug events: a prospective observational study. *Br J Clin Pharmacol* 2018; 84 (10): 2252–9.
69. Lipitz-Snyderman A, Classen D, Pfister D, *et al.* Performance of a trigger tool for identifying adverse events in oncology. *J Oncol Pract* 2017; 13 (3): e223–30.
70. Menat U, Desai C, Shah M, Shah A. Evaluation of trigger tool method for adverse drug reaction reporting by nursing staff at a tertiary care teaching hospital. *Int J Basic Clin Pharmacol* 2019; 8 (6): 1139.
71. Klein DO, Rennenberg RJMW, Koopmans RP, Prins MH. The ability of triggers to retrospectively predict potentially preventable adverse events in a sample of deceased patients. *Prev Med Reports* 2017; 8: 250–5.
72. Kolc BS, Peterson J, Harris SS, Vavra JK. Evaluation of trigger tool methodology related to adverse drug events in hospitalized patients. *Patient Saf* 2019; 1 (2): 14–23. SE-Original Research and Articles
73. Davis J, Harrington N, Bittner Fagan H, Henry B, Savoy M. The accuracy of trigger tools to detect preventable adverse events in primary care: a systematic review. *J Am Board Fam Med* 2018; 31 (1): 113–25.
74. Musy SN, Ausserhofer D, Schwendimann R, *et al.* Trigger tool-based automated adverse event detection in electronic health records: systematic review. *J Med Internet Res* 2018; 20 (5): e198.doi:10.2196/jmir.9901
75. Singh H, Bradford A, Goeschel C. Operational measurement of diagnostic safety: state of the science. *Diagnosis (Berlin, Ger)* 2021; 8 (1): 51–65.
76. Singh H, Thomas EJ, Khan MM, Petersen LA. Identifying diagnostic errors in primary care using an electronic screening algorithm. *Arch Intern Med* 2007; 167 (3): 302–8.
77. Eisenberg M, Madden K, Christianson JR, Melendez E, Harper MB. Performance of an automated screening algorithm for early detection of pediatric severe sepsis. *Pediatr Crit Care Med* 2019; 20 (12): e516–23.
78. Meyer AND, Singh H. The path to diagnostic excellence includes feedback to calibrate how clinicians think. *JAMA* 2019; 321 (8): 737–8.
79. Meyer AND, Upadhyay DK, Collins CA, *et al.* A program to provide clinicians with feedback on their diagnostic performance in a learning health system. *Jt Comm J Qual Patient Saf* 2021; 47 (2): 120–6. doi:10.1016/j.jcjq.2020.08.014.
80. 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 findings. *Radiology* 2015; 277 (1): 81–7.
81. 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. *JCO* 2015; 33 (31): 3560–7.
82. Wandtke B, Gallagher S. Reducing delay in diagnosis: multistage recommendation tracking. *AJR Am J Roentgenol* 2017; 209 (5): 970–5.
83. Danforth KN, Smith AE, Loo RK, *et al.* Electronic clinical surveillance to improve outpatient care: diverse applications within an integrated delivery system. *eGEMs* 2014; 2 (1): 1056.
84. Sim JJ, Rutkowski MP, Selevan DC, *et al.* Kaiser permanente creatinine safety program: a mechanism to ensure widespread detection and care for chronic kidney disease. *Am J Med* 2015; 128 (11): 1204–1211.e1.
85. Meyer AND, Murphy DR, Al-Mutairi A, *et al.* Electronic detection of delayed test result follow-up in patients with hypothyroidism. *J Gen Intern Med* 2017; 32 (7): 753–9.
86. Singh H, Hirani K, Kadiyala H, *et al.* Characteristics and predictors of missed opportunities in lung cancer diagnosis: an electronic health record-based study. *JCO* 2010; 28 (20): 3307–15.
87. Weiner SJ, Wang S, Kelly B, Sharma G, Schwartz A. How accurate is the medical record? A comparison of the physician's note with a concealed audio recording in unannounced standardized patient encounters. *J Am Med Inform Assoc* 2020; 27 (5): 770–5. doi:10.1093/jamia/ocaa027.
88. Classen D, Li M, Miller S, Ladner D. An electronic health record-based real-time analytics program for patient safety surveillance and improvement. *Health Aff (Millwood)* 2018; 37 (11): 1805–12.
89. Forster AJ, Worthington JR, Hawken S, *et al.* Using prospective clinical surveillance to identify adverse events in hospital. *BMJ Qual Saf* 2011; 20 (9): 756–63.
90. Zwaan L, de Bruijne M, Wagner C, *et al.* Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med* 2010; 170 (12): 1015–21.
91. Thomas EJ, Lipsitz SR, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. *Ann Intern Med* 2002; 136 (11): 812–6.