An electronic trigger based on care escalation to identify preventable adverse events in hospitalised patients

Viraj Bhise,[1,2] Dean F Sittig,[3] Viralkumar Vaghani,[1,2] Li Wei,[1,2] Jessica Baldwin,[1,2] Hardeep Singh[1,2]

[1]Center for Innovations in Quality, Effectiveness, and Safety (IQuESt), Michael E DeBakey Veterans Affairs Medical Center, Houston, Texas, USA
[2]Department of Medicine, Baylor College of Medicine, Houston, Texas, USA
[3]School of Biomedical Informatics, University of Texas Health Science Center, Houston, Texas, USA

Correspondence to
Dr Hardeep Singh, Department of Medicine, Baylor College of Medicine, 2002 Holcombe Blvd, 152, Houston, Texas 77030, USA; hardeeps@bcm.edu

ABSTRACT

Background Methods to identify preventable adverse events typically have low yield and efficiency. We refined the methods of Institute of Healthcare Improvement's Global Trigger Tool (GTT) application and leveraged electronic health record (EHR) data to improve detection of preventable adverse events, including diagnostic errors.

Methods We queried the EHR data repository of a large health system to identify an 'index hospitalization' associated with care escalation (defined as transfer to the intensive care unit (ICU) or initiation of rapid response team (RRT) within 15 days of admission) between March 2010 and August 2015. To enrich the record review sample with unexpected events, we used EHR clinical data to modify the GTT algorithm and limited eligible patients to those at lower risk for care escalation based on younger age and presence of minimal comorbid conditions. We modified the GTT review methodology; two physicians independently reviewed eligible 'e-trigger' positive records to identify preventable diagnostic and care management events.

Results Of 88 428 hospitalisations, 887 were associated with care escalation (712 ICU transfers and 175 RRTs), of which 92 were flagged as trigger-positive and reviewed. Preventable adverse events were detected in 41 cases, yielding a trigger positive predictive value of 44.6% (reviewer agreement 79.35%; Cohen's kappa 0.573). We identified 7 (7.6%) diagnostic errors and 34 (37.0%) care management-related events: 24 (26.1%) adverse drug events, 4 (4.3%) patient falls, 4 (4.3%) procedure-related complications and 2 (2.2%) hospital-associated infections. In most events (73.1%), there was potential for temporary harm.

Conclusion We developed an approach using an EHR data-based trigger and modified review process to efficiently identify hospitalised patients with preventable adverse events, including diagnostic errors. Such e-triggers can help overcome limitations of currently available methods to detect preventable harm in hospitalised patients.

BACKGROUND

Measuring adverse events accurately is foundational for patient safety improvement efforts, but all existing measurement tools have limitations.[1] Many hospitals use trigger tools such as the Institute of Healthcare Improvement's (IHI) Global Trigger Tool (GTT) to monitor adverse events and patient harm in inpatient settings.[2–4] Reviewers of GTTs are explicitly instructed to not make judgements about preventability during the review process.[2 5] A recent systematic review by Hibbert *et al*[5] suggests that rather than being used primarily for counting adverse events, application of GTT should be reframed as an opportunity to understand events and to determine the most frequent event type for purposes of quality improvement. The review also recommends the need for using preventability scores for setting local priorities and for including 'omission' adverse events.[5] For example, similar to many other measurement methods, current applications of GTT usually are unable to find 'omission' events related to diagnostic errors.[6–8]

Because trigger tools help identify an at-risk patient cohort that needs confirmatory reviews to determine adverse events, the yield of the trigger and the efficiency of the application processes are important considerations for anyone using them. In the recent Hibbert *et al* review, the yield of GTT was found to vary between 7% and 40%.[5] Previous application has involved a manual review of a large number of patient charts to look for the presence of triggers, followed by a detailed review among triggered records to identify adverse events.[2–4 9–11] Conversely, newly available clinical data from electronic health records (EHRs) provide a unique opportunity to select which records to review.[12 13] Methods to focus and optimise current trigger tools and improve the

efficiency and yield of detecting adverse events would increase the percentage of reviewed medical records where an adverse event was identified, lowering the burden of record reviews at the same time.

Our study objective was to refine the methods of GTT application and leverage EHR data to improve detection of preventable adverse events, including diagnostic errors. More efficient methods to measure preventable events could lead to focused learning and quality improvement efforts, help facilitate analysis to understand contributory factors for these events, and help inform interventions for improvement.[14]

METHODS

We queried the EHR data repository at a large health system to identify an 'index hospitalization' associated with escalation of care (defined as transfer to the medical intensive care unit (ICU) or initiation of rapid response team (RRT) within 15 days of admission) between March 2010 and August 2015. We used expert input to identify automated inclusion and exclusion criteria that could be used to enrich the triggered cohort such that on review we might be more likely to find a preventable adverse event. We then used an iterative process using expert input to finalise our trigger by conducting pilot chart reviews of triggered records. We focused on patients at *lower* risk for escalation of care during hospitalisation based on two criteria: (1) age 65 years or younger when admitted to an adult inpatient service, and (2) presence of minimal comorbid conditions (Charlson Comorbidity Index[15] <2). For such patients, escalation of care if it occurred would more likely be unexpected and more likely to be preventable. To further increase the yield of preventable events, we electronically excluded patients transferred for postprocedure care (eg, after surgery and procedures like percutaneous coronary intervention), who were frequently admitted (three or more prior hospitalisations in the past year), or who were transferred to a hospice or palliative care within a 6-month time-period prior to the index hospitalisation. Using the automated inclusion and exclusion criteria above facilitated the use of electronic data to refine the GTT algorithm and led to development of an 'e-trigger'.

We also refined the GTT review methodology. We only reviewed charts that were identified by the e-trigger. After undergoing training, two physicians independently reviewed all eligible records to identify events related to errors in diagnostic assessment and care management. Both physician reviewers were experienced in patient safety-related electronic medical record reviews and received additional training for this study. They were asked to spend 20 minutes or less per chart to ensure broader practical application of review techniques in the future. Reviewers used the Safer-Dx instrument for assessment of diagnostic errors and for collecting information about process breakdowns

using the five diagnostic process dimensions (patient factors, patient–provider encounter, test performance and interpretation, test follow-up and consultations).[16 17] To capture care management events, we identified adverse drug events, healthcare-acquired infections, post-operative complications, fall-related injuries and other adverse events. Potential harm was captured using the AHRQ Common Format Harm Scale V.1.2.[18] Disagreements among reviewers were discussed and resolved by consensus prior to analysis.

RESULTS

Of 88 428 hospitalisations during the study period, 887 were associated with escalation of care (712 ICU transfers and 175 RRTs). Of these 887 index hospitalisations, 92 (10.4%) involved unique patients in a low-risk cohort who encountered an unexpected escalation of care. The positive predictive value (PPV) for detecting any preventable adverse event in this cohort was 44.6% (41 of 92), with reviewer agreement of 79.35% (Cohen's kappa 0.573, CI 0.409 to 0.747).

We detected 7 (7.6%) diagnostic errors and 34 (37.0%) care management-related preventable adverse events: 24 (26.1%) adverse drug events, 4 (4.3%) patient falls, 4 (4.3%) procedure-related complications and 2 (2.2%) hospital-associated infections. Diagnostic errors included missed diagnoses of deep vein thrombosis, haemothorax, sepsis and alcohol withdrawal (examples in table 1). Errors occurred from breakdowns in the patient–provider encounter (ie, history, exam, test ordering; n=6, 85.7%), including failures in information gathering and interpretation (eg, history of alcohol use was missed, leg pain in an immobilised patient was not evaluated during patient assessment) and delays in test follow-up and tracking (eg, chest X-ray ordered but abnormal finding missed). In most of the events (73.1%; 30 of 41), there was potential for temporary harm. Also in all seven cases of diagnostic error, there was potential for serious harm.

DISCUSSION

We developed a new approach, based on an e-trigger and modified review methods, to identify patients with preventable adverse events in inpatient settings. The approach leveraged EHR data and used a modified GTT algorithm and chart review methodology to increase the yield for preventable events. We were also able to identify inpatient diagnostic errors, which is a limitation of other currently available tools. Modified e-triggers that use increasingly available clinical data through EHRs could improve identification of preventable adverse events in hospitals and set a stronger foundation for quality improvement and learning efforts.[14] Relatively more efficient measurement methods could lead to better understanding of contributory factors for these events, and help inform interventions for improvement.

Table 1 Examples of diagnostic errors and other adverse events identified in the study

Case history	Type of error	Dimension of care	Anticipated harm and duration
Patient with known alcohol abuse presented with inability to walk. History of leg pain and past immobilisation was missed. Patient experienced pulmonary embolism.	Diagnostic error	Patient–provider encounter	Severe Temporary
Patient presented with hip pain after a recent fall. Also, patient experienced chest pain and shortness of breath for 1 week. Early signs of infection (leucocytosis, cough with coffee-ground sputum) were missed. 3 days later patient developed septic shock and died 2 days later.	Diagnostic error	Patient–provider encounter	Death Permanent
Patient is admitted with multiple rib fractures. Chest X-ray on admission showed a haemothorax, which was missed. 2 days later, patient developed sudden shortness of breath and RRT is called. Repeat chest X-ray confirmed the haemothorax.	Diagnostic error	Follow-up and tracking	Severe Temporary
Patient presented with altered mental status. During initial assessment history of alcohol use was missed. Patient developed alcohol-withdrawal hallucinations and seizures 2 days later.	Diagnostic error	Patient–provider encounter	Severe Temporary
2 patients had allergic reactions/anaphylaxis to CT contrast agent.	Adverse drug reaction	NA	Mild Temporary
Patient was accidentally given multiple hypotensive agents (dose of prazosin was increased and amlodipine was also prescribed). Patient suddenly became hypotensive.	Adverse drug reaction	NA	Mild Temporary

RRT, rapid response team.

The 44.6% PPV of the escalation e-trigger was achieved relatively more efficiently in relation to prior comparable efforts. In our study, of the 88 428 hospitalisations observed, EHR data helped us identify the 887 associated with escalation of care and then helped further select just the 92 care escalations (ie, 0.1% of all hospitalisations) that were unexpected because of their low a priori risk. Thus, EHR data greatly helped increase the yield and efficiency, identifying just the 10% of 'enriched' patient care escalations (92 of 887) we needed to review, of which more than two-fifths were found to contain an error. The e-trigger thus compares favourably with 'unenriched' random review methodologies. The refinement illustrates how organisations can leverage their EHR data to detect and focus on preventable adverse events including diagnostic errors using a lens of learning and quality improvement. Because manual record reviews are resource-intensive, they should be reserved for records that are highly likely to reveal learning opportunities. Future use of 'free-text' data using natural language processing could potentially improve the yield and efficiency further by making the reasoning behind specific patient transfers clearer. With additional development and evaluation, a portfolio of EHR-enhanced 'smart' e-triggers could help hospitals improve the efficiency of their current patient safety monitoring activities.

The methods proposed herein advance prior scientific knowledge on application and use of GTTs. Table 2 compares our findings with those from previous studies using both GTT and escalation of care triggers.[9–11 19–22] Only a few of these studies focused on preventable adverse events and errors as learning opportunities.[5 9 11 19 21] Overall, the PPV in this study compares well to prior work, being superior to two other large non-surgical studies.[9 11] Our PPV for preventable events for care escalation is slightly less than in the Naessens et al's study,[20] where the study investigators used a substantially different manual review methodology involving random reviews of completed charts to identify any of the 55 IHI triggers. Rather than random reviews, the investigators themselves recommend a more focused review of records known to contain triggers with higher yields to get better insight into problems with care delivery. They also recommend developing automated techniques to identify triggers followed by record review to allow focus on contributory causes of events, rather than just identifying events. These recommendations are consistent with our enhancements. We were also able to improve on previous studies that have used initiation of RRT as a trigger to identify preventable adverse events.[19 21] Thus, our study methods and focus on preventable events and diagnostic errors advance the body of knowledge on use of trigger methods for hospitalised patients.

The escalation e-trigger selected events that were more likely to be unexpected than originally proposed in GTT and potentially more likely to be associated with error. To focus on preventable adverse events, healthcare institutions could use similar strategies to refine and improve efficiency of trigger tools. An iterative chart review process under expert guidance could help in further refinement and customisation. Nevertheless,

Table 2 Comparison of findings from a sample of prior GTT studies with escalation e-trigger

Authors, year	Method of record selection	Total records reviewed	Total records identified as GTT-positive	Reported GTT yield (% adverse events discovered per GTT trigger positive charts)	Records identified as positive for escalation of care (manually in all except our study)	Reported yield (% adverse events discovered per 'escalation of care' trigger positive charts)	Distinction between preventable and non-preventable (if reported)
Iyengar et al, 2009[19]	All rapid response team consults in a 4-week period	65	N/A	N/A	23	35% (23/65)	69.6% preventable
Naessens et al, 2010[20]	10 random charts every 2 weeks for 2 years	1138	913	33.6% (307/913)	56	66% (37/56)	No distinction
Kennerly et al, 2013[9]	10–35 random charts every month for 2 years	16172	14182	19.5% (2772/14 182)	Not reported	3.9%	12.5% preventable
O'Leary et al, 2013[10]	250 randomly selected records	250	Not reported	Not reported	18	13% (2/18)	No distinction
Unbeck et al, 2013[22]	350 random orthopaedic patients	350	Not reported	28% (98/350)	5	80% (4/5)	79% preventable
Hwang et al, 2014[11]	30 random charts per week for 6 months	629	Not reported	Not reported	18	11.1% (2/18)	61% preventable
Amaral et al, 2015[21]	247 rapid response team consults over a 7-month period	247	N/A	N/A	247	17.8% (44/247)	79.5% preventable
Escalation e-trigger	Electronic trigger applied to 88 428 hospitalisations	N/A	N/A	N/A	92 high-risk records identified from 887	44.6% (41/92)	All preventable

GTT, Global Trigger Tool.

we note that introducing selective review and sample enrichment is more useful for quality improvement, learning and research purposes and may not be used for calculating event rates.

We see several advantages of using an enriched patient sample. For example, a health system using this trigger (and similar future triggers) will find that the reviewers will need to review a much smaller number of records to find the few that need to be analysed in detail for learning and improvement. This could bolster patient safety improvement efforts in health systems with constrained resources and competing demands on quality measurement. Contributory factors uncovered through a more detailed postreview safety analysis could provide the impetus for solutions, including non-punitive feedback to the front-line care team. Because very few methods focus on inpatient diagnostic errors, future efforts using similar triggers could be useful to identify and understand contributory factors associated with diagnostic adverse events in inpatient settings.[6–8] While this trigger cannot be used for estimating frequency, a combination of various types of electronic triggers could be refined and tested and if found useful can be used to calculate frequency of inpatient diagnostic errors, a number that remains elusive and yet to be defined in US hospitals.[23]

Several limitations merit discussion. Our study was performed at one site and our findings might not be necessarily generalisable to others. However, the trigger uses a common query language and relatively standard criteria (ICD-9 codes and event-specific codes for ICU transfer, RRT and hospice) that could be replicated easily. We were unable to report sensitivity and specificity of the trigger tool or calculate the prevalence of preventable adverse events, due to inability to perform a larger number of additional record reviews necessary to find false-negative cases or calculate prevalence estimates. However, this refinement is the first step towards additional development and application. Determination of preventability is subject to reviewer judgement,[24] but we took measures to make record reviews more objective. Also, as in most other retrospective evaluations of adverse events, we cannot rule out hindsight bias.

In conclusion, we developed an EHR data-based trigger and modified review processes to efficiently identify hospitalised patients with preventable adverse events, including diagnostic errors. Such e-triggers can help overcome limitations of currently available methods and inform the future development of robust measurement systems to detect and prevent harm from diagnostic errors and adverse events in hospitalised settings.

Twitter @HardeepSinghMD

Contributors Study concept and design: VB, HS, DFS. Acquisition of data: VB, LW. Statistical analysis: VB. Analysis and interpretation of data: VB, HS. Drafting of the manuscript: VB. Critical revision of the manuscript for important intellectual content: VB, DFS, VV, LW, JLB, HS. Administrative, technical or material support: VB, DFS, VV, LW, JLB, HS. Study supervision: VB, DFS, HS.

REFERENCES

1 Shojania KG. The elephant of patient safety: what you see depends on how you look. *Jt Comm J Qual Patient Saf* 2010;36:399–AP3.

2 Griffin FA, Resar RK. *IHI global trigger tool for measuring adverse events*. Cambridge, MA: Institute for Healthcare Improvement Innovation Series White Paper, 2009.

3 Adler L, Denham CR, McKeever M, *et al*. Global trigger tool: Implementation basics. *J Patient Saf* 2008;4:245–9.

4 Classen DC, Lloyd RC, Provost L, *et al*. Development and Evaluation of the Institute for Healthcare Improvement Global Trigger Tool. *J Patient Saf* 2008;4:169–77.

5 Hibbert PD, Molloy CJ, Hooper TD, *et al*. The application of the Global Trigger Tool: a systematic review. *Int J Qual Health Care* 2016;28:640–9.

6 Shenvi EC, El-Kareh R. Clinical criteria to screen for inpatient diagnostic errors: a scoping review. *Diagnosis* 2015;2:3–19.

7 Bhise V, Singh H. Measuring diagnostic safety of inpatients: time to set sail in uncharted waters. *Diagnosis* 2015;2:1–2.

8 Balogh EP, Miller BT, Ball JR. Improving diagnosis in health care. *National Academies Press* 2016.

9 Kennerly DA, Saldaña M, Kudyakov R, *et al*. Description and evaluation of adaptations to the global trigger tool to enhance value to adverse event reduction efforts. *J Patient Saf* 2013;9:87–95.

10 O'Leary KJ, Devisetty VK, Patel AR, *et al*. Comparison of traditional trigger tool to data warehouse based screening for identifying hospital adverse events. *BMJ Qual Saf* 2013;22:130–8.

11 Hwang JI, Chin HJ, Chang YS. Characteristics associated with the occurrence of adverse events: a retrospective medical record review using the Global Trigger Tool in a fully digitalized tertiary teaching hospital in Korea. *J Eval Clin Pract* 2014;20:27–35.

12 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.

13 Murphy DR, Wu L, Thomas EJ, *et al*. Electronic Trigger-Based Intervention to Reduce Delays in Diagnostic Evaluation for Cancer: A Cluster Randomized Controlled Trial. *J Clin Oncol* 2015;33:3560–7.

14 Wright J, Shojania KG. Measuring the quality of hospital care. *BMJ* 2009;338:b569.

15 Quan H, Sundararajan V, Halfon P, *et al*. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care* 2005;43:1130–9.

16 Al-Mutairi A, Meyer AN, Thomas EJ, *et al*. Accuracy of the Safer Dx Instrument to Identify Diagnostic Errors in Primary Care. *J Gen Intern Med* 2016;31:602–8.

17 Singh H, Sittig DF. Advancing the science of measurement of diagnostic errors in healthcare: the Safer Dx framework. *BMJ Qual Saf* 2015;24:103–10.

18 Williams T, Szekendi M, Pavkovic S, *et al*. The reliability of AHRQ Common Format Harm Scales in rating patient safety events. *J Patient Saf* 2015;11:52–9.

19 Iyengar A, Baxter A, Forster AJ. Using Medical Emergency Teams to detect preventable adverse events. *Crit Care* 2009;13:R126.

20 Naessens JM, O'Byrne TJ, Johnson MG, *et al*. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care* 2010;22:266–74.

21 Amaral AC, McDonald A, Coburn NG, *et al*. Expanding the scope of Critical Care Rapid Response Teams: a feasible approach to identify adverse events. A prospective observational cohort. *BMJ Qual Saf* 2015;24:764–8.

22 Unbeck M, Schildmeijer K, Henriksson P, *et al*. Is detection of adverse events affected by record review methodology? an evaluation of the "Harvard Medical Practice study" method and the "Global Trigger Tool". *Patient Saf Surg* 2013;7:10.

23 Singh H, Zwaan L. Reducing diagnostic error—A new horizon of opportunities for hospital medicine. *Ann Intern Med* 2016;165.

24 Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. *JAMA* 2001;286:415–20.