Automated Screening of Emergency Department Notes for Drug-Associated Bleeding Adverse Events Occurring in Older Adults

Richard D. Boyce[1]   Jeremy Jao[1]   Taylor Miller[2]   Sandra L. Kane-Gill[3]

[1] Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania, United States
[2] Department of Pharmacy, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, United States
[3] School of Pharmacy, University of Pittsburgh, Pittsburgh, Pennsylvania, United States

Address for correspondence Richard D. Boyce, PhD, Suite 419, 5607 Baum Boulevard, University of Pittsburgh, Pittsburgh, PA 15206, United States (e-mail: rdb20@pitt.edu).

Abstract

Objective   To conduct research to show the value of text mining for automatically identifying suspected bleeding adverse drug events (ADEs) in the emergency department (ED).

Methods   A corpus of ED admission notes was manually annotated for bleeding ADEs. The notes were taken for patients $\geq$ 65 years of age who had an ICD-9 code for bleeding, the presence of hemoglobin value $\leq$ 8 g/dL, or were transfused > 2 units of packed red blood cells. This training corpus was used to develop bleeding ADE algorithms using Random Forest and Classification and Regression Tree (CART). A completely separate set of notes was annotated and used to test the classification performance of the final models using the area under the ROC curve (AUROC).

Results   The best performing CART resulted in an AUROC on the training set of 0.882. The model's AUROC on the test set was 0.827. At a sensitivity of 0.679, the model had a specificity of 0.908 and a positive predictive value (PPV) of 0.814. It had a relatively simple and intuitive structure consisting of 13 decision nodes and 14 leaf nodes. Decision path probabilities ranged from 0.041 to 1.0. The AUROC for the best performing Random Forest method on the training set was 0.917. On the test set, the model's AUROC was 0.859. At a sensitivity of 0.274, the model had a specificity of 0.986 and a PPV of 0.92.

Conclusion   Both models accurately identify bleeding ADEs using the presence or absence of certain clinical concepts in ED admission notes for older adult patients. The CART model is particularly noteworthy because it does not require significant technical overhead to implement. Future work should seek to replicate the results on a larger test set pulled from another institution.

Keywords
► hemorrhage
► quality improvement
► text mining
► adverse drug event
► emergency department

Background and Significance

While medications help millions of people live longer and healthier lives, they are also the cause of approximately 280,000 hospital admissions and one-third of all hospital adverse events in the United States each year.[1] Older adults are particularly vulnerable and have been observed to experience a greater number of potential unsafe medication situations than younger adults.[2] Bleeding events are among the most common medication-related causes of hospitalization and

Copyright © 2017 Schattauer

death.[3] Drug-induced bleeding events have been the focus of recent national attention due to their association with preventable serious adverse events[1,4] and an increase in recent years in the number of antithrombotic therapies approved by the Food and Drug Administration.

Tools are needed to help accurately identify suspected bleeding adverse drug events (ADEs) for the purposes of evaluating causes and finding opportunities for future prevention. One use for such tools is to track the rate of potential occurrences over time to assess the impact of interventions to reduce hospital-acquired bleeding ADEs.[5] The identification of the ADEs in the clinical setting has proven to be incomplete, time-consuming, and costly. Despite the fact that voluntary reporting by clinical staff is known to underestimate the actual rate of ADE,[6] it remains a commonly employed approach to ADE detection.[7] Manual medical record review is a potentially more complete method of ADE detection with a greater number of ADEs identified; however, it is associated with a greater than 10-fold increase in person-hours per week,[8] making it impractical for routine use.

Objectives

Research has shown that text mining clinical notes can accurately and rapidly identify suspected adverse events.[9,10] Algorithm development using keyword matching[11] and trigger-word–based ADE detection[12] has already shown promise for ADE identification. Furthermore, protocols have been proposed for using natural language processing of various electronic data to identify adverse events that occur to hospitalized patients.[13] However, further research is needed to show the value of text mining for automatically identifying bleeding ADEs in the emergency department (ED). Moreover, to achieve broader adoption, we believe that it is important to create a validated ADE detection tool that uses text mining, but does not require significant technical overhead. In this article, we report on a new, highly interpretable and easy to implement Classification and Regression Tree (CART)[14] model. We compare the performance of the model to a Random Forest model that we created using the same dataset. We show that both models accurately identify bleeding ADEs using the presence or absence of certain clinical concepts in ED admission reports for older adult patients.

Methods

Study Design

A retrospective chart review was conducted at three academic-affiliated hospitals with sites including an urban, 800-bed tertiary care academic medical center; a 249-bed acute care and teaching hospital; and a 399-bed tertiary care hospital situated in a suburban setting, respectively. Data from ED admission reports were obtained using a cross-site hospital-based information medical records archiving system. Institutional review board approval was obtained before initiating the study.

Patient Population

Eligible patients were $\geq$ 65 years of age and had any of the following: an ICD-9 code for bleeding (see ►Appendix), the presence of hemoglobin value $\leq$ 8 g/dL, or were transfused > 2 units of packed red blood cells in the ED. These selection criteria are very similar to those reported by Mayo Clinic as having a sensitivity of approximately 94% and a positive predictive value (PPV) of approximately 42%.[5] The dataset used to train the algorithm consisted of ED reports for patients who met these criteria during an ED visit occurring between July 1, 2011, and June 30, 2012. A separate dataset used to test the classification performance of the final model included ED reports for patients who met the same criteria during a visit between June 3, 2013, and May 23, 2014.

Protocol

A practicing pharmacist (T.M.) annotated each de-identified ED report as positive or negative for the mention of a bleeding ADE. All reports were then processed using the open source NOBLE program[15] (downloaded from http://slidetutor.upmc.edu/ on November 11, 2013, source code available at GitHub: https://github.com/dbmi-pitt/noble-tools) to identify mentions of concepts from the Unified Medical Language System (UMLS) Metathesaurus (version 2014AB).[16] NOBLE was run using the Best Match Algorithm[15] with the following four terminologies:

1. RxNorm—a standardized nomenclature for clinical drugs maintained by the National Library of Medicine.[17]
2. SNOMED-CT—a comprehensive clinical terminology commonly used for clinical information exchange and research with electronic health records data.[18]
3. Medical Dictionary for Regulatory Activities (MedDRA)—a standardized medical terminology used for reporting adverse events.[19]
4. Medical Subject Headings (MeSH)—a standardized terminology for indexing and cataloging published biomedical research.[20]

The tool was configured to label both positive and negated concepts (e.g., "dizziness" and "no dizziness" were coded separately). NOBLE implements the NegEx algorithm[21] which focuses on directly negated phrases such as "absence of" and "negative for" rather than pseudonegated phrases such as "not certain if" and "not rule out." We also postprocessed NOBLE output to drop concepts that we thought would be noninformative for bleeding ADE classification. Specifically, the tool dropped all concepts tagged with the following UMLS Semantic Network types: *Quantitative Concept*, *Organism*, *Cell Component*, *Cell*, *Cell Function*, *Spatial Concept*, and *Phenomenon or Process*.

The open source Waikato Environment for Knowledge Analysis (WEKA) version 3.6.12 was used to conduct all machine learning experiments. Our main goal was to build a drug-related bleeding ADE classification model that would be easy to understand by clinicians. For that reason, we chose to use the CART machine learning approach[14] which outputs a model as a readable binary decision tree. We recognize that, in some cases, a highly interpretable model such as CART

might not perform as well as a less interpretable model such as Random Forest. So, we also applied Random Forest classification to the same input that was provided to the CART algorithm. For this study, the input to both the CART and Random Forest algorithms was a matrix with each column representing a unique concept present in one or more ED reports, and each row representing all concepts present for a single ED report. Each cell in the matrix held a 1 or a 0 depending on if the concept represented by the cell's column was present in the report represented by the cell's row. An exception was the last column in the matrix which held a 1 or 0 depending if the pharmacist identified a bleeding ADE mention within the text of the report.

There can be several thousands of clinical concepts present in a corpus of ED reports, most of which likely have no connection to the mention of a bleeding ADE. For example, ►Fig. 1 shows a hypothetical example excerpt from an ED note with mention of a bleeding ADE and concepts identified by the NOBLE program. The majority of the recognized concepts are descriptive of the patient context including his/her allergies, mode of arrival, and history of present illness. This situation, present in all ED reports, increased the possibility that a machine learning model would over-fit the training dataset. To address this issue, we tested several different strategies for feature selection using two established methods, namely, Correlation-based Feature Subset Selection[22] and Information Gain Attribute Evaluation. The former method "evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them,"[23] while the latter "evaluates the worth of an attribute by measuring the information gain with respect to the class."[24]

Both of the feature selection methods mentioned earlier use a 10-fold cross-validation approach to evaluate feature selection performance. WEKA's implementation of cross-validation splits a given dataset into 10 subsets using an integer-valued random number generator seed. During initial testing, we found that the relative importance of concepts selected by each method was influenced by which seed was selected. Since our focus was on ADEs, we decided that it would be important to ensure that all potentially important drug features would have a chance to be picked up by the final classification model. To accomplish that, we used a set union operation to forcibly combine all drug-related concepts present in the ED report corpus with the features selected automatically using one of the methods mentioned earlier. Operationally, drug-related concepts were those concepts mentioned in the ED with the following UMLS Semantic Network type: *Pharmacologic Substance*, *Biologically Active Substance*, *Organic Chemical*, *Hazardous or Poisonous Substance*, *Indicator*, *Reagent*, or *Diagnostic Aid*, *Biomedical or Dental Material*.

Chief Complaint (As documented by Nursing)
Nosebleed

Mode of Arrival: Ambulance
Pre-hospital Intervention(s): None
Triage Intervention(s): None
Triage Disposition: Treatment Room

History of Present Illness
This is an a 89-year-old male with history of multiple medical problems including CHF, A. fib on Coumadin, MI, pacemaker, St. Jude's valve who reports she's had intermittent right-sided nosebleed since this morning that worsened over the past hour and he has been unable to control it. He reports he feels mild associated shortness of breath at rest is worse with any exertion. He reports some blood has been coming out of the medial canthus of his right eye He denies any injuries or trauma to the nose. He denies any dizziness or lightheadedness. He denies any chest pain. He does have a previous history of CHF. No fevers or chills. No previous significant nosebleeds

Histories

Allergies

cephalosporins [Unknown (origin)]; IVP Contrast [Unknown (origin)] penicillin [Swelling, NOS]; povidone iodine topical [flare up]; ciprofloxacin Unknown (origin)]; sulfADIAZINE

Home Medications (from "Document Medication by Hx")

Fig. 1 A hypothetical example excerpt from an emergency department note with mention of a bleeding adverse drug event and concepts (shown in blue) identified by the NOBLE program.

Our testing strategy was to compare the performance of the CART and Random Forest algorithms on the training set using common sense feature selection variations, including no feature selection, drug-related features alone, and then various combinations of concepts produced by the feature selection methods using three different seeds (e.g., the union or intersection of concepts across seeds). The Area under the ROC curve (AUROC) metric was generated for each test and used to make a decision about further testing until we arrived at what we thought was the best performing approach. The best performing CART and Random Forest model was applied to a separate dataset that was not used for model development to assess AUROC. A probability of 50% was used as the threshold for predicting a bleeding ADE when calculating precision, recall, and balanced F measure because it was the default setting in WEKA.

Results

►Fig. 2 shows flow diagrams of the pharmacist's review for both the training and test sets. A total of 1,716 reports were identified and reviewed for the training dataset, with 1,035 (60.3%) reports describing bleeding events. These 1,035 reports were used for algorithm development and are represented in the denominator of the following percentages. Bleeding events discussed in 250 (24.2%) reports had no certain relationship with drug exposure, 196 (18.9%) reports indicated a bleeding event that was clearly not drug related, and 589 (56.9%) were suspected drug-related bleeding event. Of these, the types of drug-related bleeding were epistaxis ($n = 243$, 41.3%), gastrointestinal ($n = 241$, 40.9%), intracranial ($n = 38$, 6.5%), and other ($n = 67$, 11.4%) (►Table 1).

The test set contained a total of 228 reports, all of which were used for testing the final models. Of these, 120 (52.6%) reports described bleeding events. Bleeding events discussed in 10

(8.3%) reports had no certain relationship with drug exposure, and 24 (20%) reports indicated a bleeding event that was clearly not drug related. Among the reports that clearly described bleeding events, 86 (72%) were suspected drug-related bleeding events. Of these, the types of drug-related bleeding were epistaxis ($n = 32$, 26.7%), gastrointestinal ($n = 28$, 23.3%), intracranial ($n = 15$, 12.5%), and other ($n = 11$, 9.2%) (►Table 1).

Ten feature selection strategies were tested using the training set:

- No feature selection.
- Drug-related features only.
- The union of concepts identified in 10, 30, and 60% of cross-folds using Correlation-based Feature Subset Selection.[22] We refer to these strategies as CFS-Union-1, CFS-Union-2, CFS-Union-3.
- The intersection of concepts identified in 10 and 30% of cross-folds using Correlation-based Feature Subset Selection (CFS-Intersection-1 and CFS-Intersection-2).
- The union and intersection of the 300 top-ranked concepts using Information Gain Attribute Evaluation. This cut-off was chosen because it resulted in what we thought would be a reasonable ∼25%/∼75% split between nondrug concepts and drug-related concepts. We refer to these strategies as IG-Union-1 and IG-Intersection-1.
- The union of the 1,000 top-ranked concepts using Information Gain Attribute Evaluation. This cut-off was chosen to test our assumption that a ∼25%/∼75% split between nondrug concepts and drug-related concepts was reasonable. We refer to this strategy as IG-Union-2.

►Table 2 shows the performance on the training set of the CART and Random Forest algorithms across all 10 strategies. For CART, the three Information Gain feature selection strategies led to better performance than the five Correlation-based Feature Subset Selection strategies, all features, or the drug
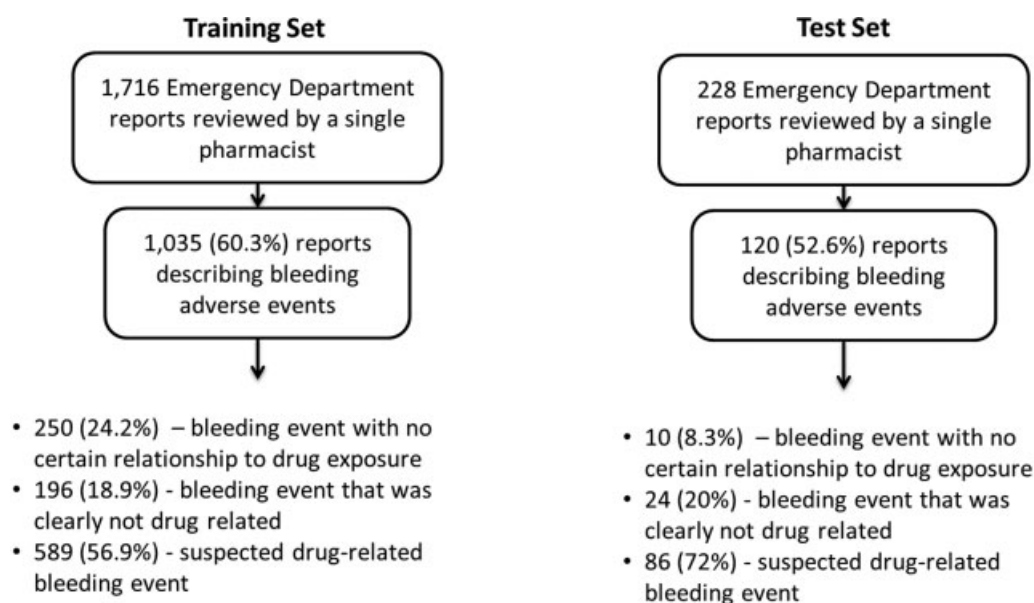
Fig. 2 Flow diagrams showing the pharmacist's review results for both the training and test sets. ►Table 1 provides further information about reports that described bleeding adverse events.

Table 1 Count of emergency department reports mentioning bleeding AEs and ADEs in the training and test sets

Training dataset n = 1,035	Type of bleeding	AE count (% of AEs) n = 446	ADE count (% of ADEs) n = 589
	Epistaxis	123 (27.6)	243 (41.3)
	Gastrointestinal	228 (51.1)	241 (40.9)
	Intracranial	48 (10.8)	38 (6.5)
	Other	47 (10.5)	67 (11.4)
Test dataset n = 120	Type of bleeding	AE count (%) n = 34	ADE count (%) n = 86
	Epistaxis	7 (20.1)	32 (26.7)
	Gastrointestinal	12 (35.3)	28 (23.3)
	Intracranial	13 (38.2)	15 (12.5)
	Other	2 (5.9)	11 (9.2)

Abbreviations: ADEs, adverse drug events; AEs, adverse events; CART, Classification and Regression Tree; PPV, positive predictive value.

features alone. For Random Forest, the five Correlation-based Feature Subset Selection and the three Information Gain feature selection strategies led to better performance than using all features or the drug features alone.

The best performing strategy for CART (IG-Union-1 in ►Table 2) is shown in ►Fig. 3. It resulted from taking the union of the top 300 concepts for the Information Gain feature method using three different seeds. The AUROC for this method on the training set was 0.882. At a sensitivity of 0.827, the model's specificity was 0.858 and PPV was 0.752. The model's AUROC on the test set was 0.827 (►Table 3). At a sensitivity of 0.679, the model had a specificity of 0.908 and a PPV of 0.814. It has a relatively simple and intuitive structure

consisting of 13 decision nodes and 14 leaf nodes. Leaf node probabilities range from 0.041 to 1.0. The lowest probability (based on test set data) occurs for notes that do not mention Coumadin (warfarin) or any of the following concepts: aspirin, Plavix (clopidogrel), aspirin 81 MG delayed release oral tablet, or Pradaxa (dabigatran etexilate mesylate). The greatest probability occurs for notes that mention Coumadin and hematoma but not hemorrhage, gastrointestinal hemorrhage, or nostril. ►Supplementary Table (available in the online version) shows the concepts used in the model and the exact string text used by NOBLE to identify each concept.

The best performing strategy for Random Forest (CFS-Union-1 in ►Table 2) resulted from taking the union of the set of concepts identified as important in 10% of cross-folds by the Correlation-based Feature Subset Selection method across three different 10-fold cross-validation runs, each using a different seed. The algorithm generated a Random forest of 100 trees, each constructed while considering 11 random features. The AUROC for this method on the training set was 0.917. At a sensitivity of 0.556, the model's specificity was 0.953 and PPV was 0.861. On the test set, the model's AUROC was 0.859. At a sensitivity of 0.274, the specificity was 0.986 and the PPV was 0.92 (►Table 3).

Discussion

Clinical surveillance of drugs is important for better understanding the causes of adverse events and preventing harm to future patients, especially since a majority of events are preventable.[1] This study shows progress toward an accurate automated means of identifying bleeding ADEs. The resulting CART and Random Forest models were both very similar in terms of overall performance as measured by AUROC. Both algorithms also had high PPVs and specificity on the test set (►Table 3) suggesting that either would signal true ADEs much more frequently than false ADEs. The CART model's

Table 2 Ten feature selection strategies tested on the training set and the resulting bleeding ADE classification performance using CART and Random Forest algorithms.

Dataset	CART				Random Forest			
	Average precision (PPV)	Average recall (sensitivity)	Average F-measure	Area under the ROC curve	Average precision	Average recall (sensitivity)	Average F-measure	Area under the ROC curve
No feature selection	0.745	0.825	0.783	0.873	0.799	0.257	0.389	0.834
Drug features only	0.674	0.825	0.742	0.837	0.793	0.488	0.604	0.879
CFS-Union-1 (10%)	0.72	0.787	0.752	0.87	**0.861**	**0.556**	**0.676**	**0.917**
CFS-Union-2 (30%)	0.735	0.791	0.762	0.871	0.86	0.532	0.658	0.91
CFS-Union-3 (60%)	0.734	0.786	0.759	0.875	0.833	0.543	0.657	0.907
CFS-Intersection-1 (10%)	0.735	0.793	0.763	0.87	0.847	0.537	0.658	0.91
CFS-Intersection-2 (30%)	0.735	0.793	0.763	0.871	0.85	0.541	0.661	0.908
IG-Union-1 (top 300)	**0.752**	**0.827**	**0.788**	**0.882**	0.838	0.509	0.633	0.906
IG-Union-2 (top 1000)	0.754	0.835	0.793	0.88	0.845	0.473	0.606	0.893
IG-Intersection-1 (top 300)	0.75	0.825	0.785	0.879	0.843	0.51	0.636	0.901

Abbreviations: ADE, adverse drug event; CART, Classification and Regression Tree; PPV, positive predictive value.
Note: All results are from 10-fold cross-validation. F-measure, recall, and precision are averaged across all folds using a probability of 50% as the threshold for predicting a bleeding ADE. The cells in bold indicate the best performing feature selection strategies for each Machine learning algorithm.
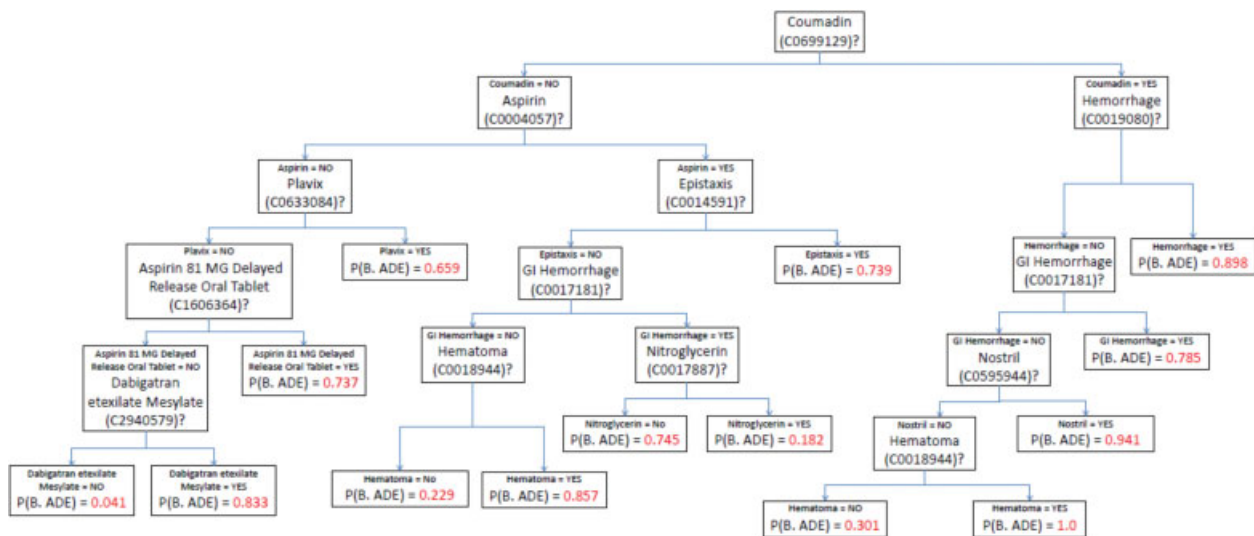
Fig. 3 The final CART model for classifying emergency department notes that mention the occurrence of a bleeding adverse drug reaction for patients ≥ 65 years of age who had an ICD-9 code for bleeding (see ►**Appendix**), the presence of hemoglobin value ≤ 8 g/dL, or were transfused > 2 units of packed red blood cells in the emergency department. Concept identifiers from the Unified Medical Language System are shown in parentheses. P(B. ADE) = the probability that a note mentioning the concepts in the decision path also mentions a bleeding adverse drug event.

sensitivity on the test set at the 50% probability threshold was more than double the Random Forest model's (0.679 vs. 0.274) but with more than six times the noise as calculated by 1—specificity (0.092 vs. 0.014).

To our knowledge, there is currently no bleeding ADE trigger tool that uses data text mined from clinical notes. The widely used Global Trigger Tool for measuring adverse events[25] is a tool that triggers for a wide range of inpatient ADEs including bleeding events. Its specificity and sensitivity have been found to be as high as 0.94 and 0.49, respectively.[26] In comparison, the CART model has much better sensitivity but lower specificity, while the Random Forest has better specificity but much lower sensitivity (►**Table 3**). Another trigger tool called the National Electronic Injury Surveillance System–Cooperative Adverse Drug Event Surveillance was evaluated for use in EDs and found to have a sensitivity of 0.33 to 0.45 and a PPV of 0.92 (specificity was not evaluated).[27] However, the tool was noted as being particularly poor at detecting bleeding ADEs. Moreover, a recent prospective cohort study found that both of these trigger tools had sensitivities of < 0.20 for identifying ED ADEs.[28]

Based on these facts, both models reported in this article identify bleeding ADEs well enough to be potentially useful as a trigger tool for quality improvement and safety surveillance

activities. The CART model is particularly promising because of its greater transparency and sensitivity, and also because it could be implemented much the same as typical trigger tools. The process would involve first querying an electronic health record for patients with similar characteristics to those used to build the model (see section "Patient Population"). A simple computer program could then scan the text of ED reports written for the patients that the query identifies. The program would look for text mentions of concepts present in the CART model (►**Supplementary Table**, available in the online version). It could then apply the decision rules shown in ►**Fig. 3** to determine if the concept co-occurrences found in the ED reports suggest a bleeding ADE.

Implementing the Random Forest model as a trigger tool would require a more sophisticated approach. A computer program would use the NOBLE tool to create a vector of UMLS concept identifiers for all concepts mentioned in the ED notes of patients selected in the same way as above (see section "Patient Population"). The program would then process the concept vectors by loading the WEKA library and applying the Random Forest model that is available upon request from the authors. The program would output a classification (1 or 0) indicating whether or not the ED note suggested a bleeding ADE.

Table 3 Performance of the best performing models for CART and Random Forest on the test set

Algorithm and feature selection strategy	Average precision (PPV)	Average recall (sensitivity)	Average specificity	Average F-measure	Area under the ROC curve
CART with IG-Union-1 (top 300)	0.814	0.679	0.908	0.74	0.827
Random Forest with CFS-Union-1 (10%)	0.92	0.274	0.986	0.422	0.859

Abbreviations: CART, Classification and Regression Tree; PPV, positive predictive value.
Note: All results are from 10-fold cross-validation. F-measure, recall, and precision are averaged across all folds using a probability of 50% as the threshold for predicting a bleeding ADE.

The current study is not without limitations. The test and training sets were annotated for bleeding ADEs by only a single pharmacist. It is possible that the pharmacist missed ADEs or made some misclassifications. The current study used a completely separate test set pulled at a different time period for evaluation model performance. This is good practice to ensure that models are not over-fit to the training dataset. However, due to resource limitations, the test set was much smaller in size than the training set. It also had relatively small counts of certain adverse events (see ►Table 1). While model performance remained relatively stable when ran on the test set, it would be desirable to replicate the results on a larger test set pulled from another institution.

In building the training and test sets, we adapted a report selection criteria reported to have a high sensitivity.[5] However, it is possible that some patients with actual bleeds might have been missed due to miscoding of the ED report. Another potential limitation is that the patient selection method used ICD-9 codes which are no longer in use. We do not expect this to be a major limitation because trusted mappings exist that could be used to map between the ICD-9 codes used in this study and ICD-10 codes. Also, ICD-9 codes were only part of the criteria used for patient selection. However, future work should test if using ICD-10 codes in patient selection has any significant effect on the performance of the models. Finally, a potential limitation of the CART model is it depends on limited criteria that include only a subset of prescribed anticoagulants (Coumadin, aspirin, clopidogrel, and dabigatran). Thus, the model does not account for bleeding ADEs associated with adoption of other anticoagulants.

Conclusion

The increased prevalence of electronic medical records has resulted in significant amounts of electronic data, creating the potential to leverage a more technology-driven approach for bleeding ADE detection. In this study, we showed the value of text mining for automatically identifying bleeding ADEs in the ED. Two text mining models were created that accurately identify bleeding ADEs using the presence or absence of certain clinical concepts in ED notes taken on older adult patients. The new CART model should be easy to implement as a trigger tool. Future work will test if the results can be replicated using data from a different ED site where ICD-10 codes are used in the patient selection process.

Clinical Relevance Statement

Drug-induced bleeding has been the focus of recent national attention due to an association with preventable serious adverse events. Tools are needed to help accurately identify suspected bleeding adverse drug events (ADEs) for evaluation of causation and to seek opportunities for future prevention. This manuscript discusses two text mining models that accurately identify bleeding ADEs using the presence or absence of certain clinical concepts in the emergency department notes taken on older adult patients.

Multiple Choice Question

Which of the following is a reason why the text mining of clinical notes is a promising approach for adverse drug event (ADE) identification:

A. Voluntary reporting by clinical staff overestimates the actual rate of ADE.
B. Manual medical record review underestimates the actual rate of ADE.
C. Text mining might capture more true events than voluntary reporting by focusing manual medical record review on likely ADE cases.
D. Text mining does not in general require significant technical overhead.

Correct Answer: The correct answer is **C**. Text mining might capture more true events than voluntary reporting by focusing manual medical record review on likely ADE cases.

Protection of Human and Animal Subjects
The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects, and was reviewed by the University of Pittsburgh Institutional Review Board.

Conflict of Interest
None.

References

1 US Department of Health and Human Services. National action plan for adverse drug event prevention. 2014. Available at: https://health.gov/hcq/pdfs/ade-action-plan-508c.pdf
2 Dormann H, Sonst A, Müller F, et al. Adverse drug events in older patients admitted as an emergency: the role of potentially inappropriate medication in elderly people (PRISCUS). Dtsch Arztebl Int 2013;110(13):213–219
3 Pretorius RW, Gataric G, Swedlund SK, Miller JR. Reducing the risk of adverse drug events in older adults. Am Fam Physician 2013; 87(05):331–336
4 Institute for Safe Medication Practices. Safe medication practices high alert medications in acute care settings. 2017. Available at: http://www.ismp.org/Tools/institutionalhighAlert.asp. Accessed July 28, 2014
5 Moriarty JP, Daniels PR, Manning DM, et al. Going beyond administrative data: retrospective evaluation of an algorithm using the electronic health record to help identify bleeding events among hospitalized medical patients on warfarin. Am J Med Qual 2017; 32(04):391–396
6 Cullen DJ, Bates DW, Small SD, Cooper JB, Nemeskal AR, Leape LL. The incident reporting system does not detect adverse drug events: a problem for quality improvement. Jt Comm J Qual Improv 1995;21(10):541–548
7 Kane-Gill SL, Devlin JW. Adverse drug event reporting in intensive care units: a survey of current practices. Ann Pharmacother 2006; 40(7-8):1267–1273

8 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(03):305–314

9 LePendu P, Iyer SV, Bauer-Mehren A, et al. Pharmacovigilance using clinical notes. Clin Pharmacol Ther 2013;93(06):547–555

10 Warrer P, Hansen EH, Juhl-Jensen L, Aagaard L. Using text-mining techniques in electronic patient records to identify ADRs from medicine use. Br J Clin Pharmacol 2012;73(05):674–684

11 Wiley LK, Moretz JD, Denny JC, Peterson JF, Bush WS. Phenotyping Adverse Drug Reactions: Statin-Related Myotoxicity. AMIA Jt Summits Transl Sci Proc 2015;2015:466–470

12 Kane-Gill SL, MacLasco AM, Saul MI, et al. Use of text searching for trigger words in medical records to identify adverse drug reactions within an intensive care unit discharge summary. Appl Clin Inform 2016;7(03):660–671

13 Rochefort CM, Buckeridge DL, Forster AJ. Accuracy of using automated methods for detecting adverse events from electronic health record data: a research protocol. Implement Sci 2015;10:5

14 Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Belmont, CA: Wadsworth International Group; 1984

15 Tseytlin E, Mitchell K, Legowski E, Corrigan J, Chavan G, Jacobson RS. NOBLE – Flexible concept recognition for large-scale biomedical natural language processing. BMC Bioinformatics. January 14, 2016:17 [cited October 6, 2016]. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4712516/

16 Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res 2004;32(Database issue):D267–D270

17 RxNorm Technical Documentation. Available at: http://www.nlm.nih.gov/research/umls/rxnorm/docs/2015/rxnorm_doco_-full_2015-1.html. Accessed October 10, 2016

18 The International Health Terminology Standards Development Organisation. SNOMED-CT. 2016. Available at: http://www.ihtsdo.org/snomed-ct

19 International Conference on Harmonisation. MedDRA [Internet]. Welcome to MedDRA. 2016 [cited October 10, 2016]. Available at: http://www.meddra.org/. Accessed October 10, 2016

20 National Library of Medicine. Medical Subject Headings - Home Page. Welcome to Medical Subject Headings! 2016. Available from: https://www.nlm.nih.gov/mesh/. Accessed October 10, 2016

21 Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 2001;34(05):301–310

22 Hall MA. Correlation-based Feature Subset Selection for Machine Learning. Hamilton, NZ: University of Waikato; 1998

23 WEKA. WEKA documentation - CfsSubsetEval. Class CfsSubsetEval. 2016. Available at: http://weka.sourceforge.net/doc.stable/weka/attributeSelection/CfsSubsetEval.html. Accessed October 10, 2016

24 WEKA. WEKA documentation - InfoGainAttributeEval. Class InfoGainAttributeEval. 2016. Available at: http://weka.sourceforge.net/doc.dev/weka/attributeSelection/InfoGainAttributeEval.html. Accessed October 10, 2016

25 Griffin F, Resar R. IHI Global Trigger Tool for Measuring Adverse Events. Second Edition. Cambridge, MA: Institute for Healthcare Improvement; 2009 (IHI Innovation Series white paper)

26 Sharek PJ, Parry G, Goldmann D, et al. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. Health Serv Res 2011;46(02):654–678

27 Centers for Disease Control and Prevention (CDC). Assessing the National Electronic Injury Surveillance System-Cooperative Adverse Drug Event Surveillance project–six sites, United States, January 1-June 15, 2004. MMWR Morb Mortal Wkly Rep 2005;54(15):380–383

28 Karpov A, Parcero C, Mok CPY, et al. Performance of trigger tools in identifying adverse drug events in emergency department patients: a validation study. Br J Clin Pharmacol 2016;82(04):1048–1057

Appendix ICD-9 codes used to identify bleeding events

Code	Description
280.0	Anemia secondary to blood loss (chronic); blood loss anemia NOS
285.1	Acute posthemorrhagic anemia
287.4	Secondary thrombocytopenia
287.9	Unspecified hemorrhagic conditions
430.0	Subarachnoid hemorrhage
431.0	Intracerebral hemorrhage
432.0	Nontraumatic extradural hemorrhage
432.1	Subdural hemorrhage
432.9	Unspecified intracranial hemorrhage
456.0	Esophageal varices with bleeding
459.0	Hemorrhage, unspecified
530.7	Gastroesophageal laceration-hemorrhage syndrome
530.82	Esophageal hemorrhage
531.0	Gastric ulcer, acute with hemorrhage
531.2	Gastric ulcer, acute with hemorrhage and perforation

(*Continued*)

Appendix (*Continued*)

Code	Description
531.4	Gastric ulcer, chronic or unspecified with hemorrhage
531.6	Gastric ulcer, chronic or unspecified with hemorrhage and perforation
532.0	Duodenal ulcer, acute with hemorrhage
532.2	Duodenal ulcer, acute with hemorrhage and perforation
532.4	Duodenal ulcer, chronic or unspecified with hemorrhage
532.6	Duodenal ulcer, chronic or unspecified with hemorrhage and perforation
533.0	Peptic ulcer, acute with hemorrhage
533.2	Peptic ulcer, acute with hemorrhage and perforation
533.4	Peptic ulcer, chronic or unspecified with hemorrhage
533.6	Peptic ulcer, chronic or unspecified with hemorrhage and perforation
534.0	Gastrojejunal ulcer, acute with hemorrhage
534.2	Gastrojejunal ulcer, acute with hemorrhage and perforation
534.4	Gastrojejunal ulcer, chronic or unspecified with hemorrhage
534.6	Gastrojejunal ulcer, chronic or unspecified with hemorrhage and perforation
557.0	Acute vascular insufficiency of intestine
562.02	Diverticulosis of small intestine with hemorrhage
562.03	Diverticulitis of small intestine with hemorrhage
562.12	Diverticulosis of colon with hemorrhage
562.13	Diverticulitis of colon with hemorrhage
578.0	Gastrointestinal hemorrhage
596.7	Hemorrhage into bladder wall
599.7	Hematuria
784.7	Epistaxis
784.8	Hemorrhage from throat
786.3	Hemoptysis
852.0	Subarachnoid, subdural, and extradural hemorrhage, following injury
853.0	Other and unspecified intracranial hemorrhage following injury
860.0	Traumatic pneumothorax and hemothorax
864.01	Injury to liver, hematoma
864.11	Injury to liver, open wound into cavity, hematoma
865.01	Injury to spleen, hematoma
865.11	Injury to spleen, open wound into cavity, hematoma
866.01	Injury to kidney, hematoma
866.11	Injury to kidney, open wound into cavity, hematoma
900.0–904.0	Injury to blood vessels
964.0	Poisoning by agents primarily affecting blood constituents

Appendix (*Continued*)