HHS Public Access

Author manuscript

J Am Geriatr Soc. Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

J Am Geriatr Soc. 2020 December ; 68(12): 2965–2967. doi:10.1111/jgs.16831.

Comparing Strategies for Identifying Falls in Older Adult Emergency Department Visits Using EHR Data

Brian W. Patterson, MD MPH[1,2,3], **Gwen Costa Jacobsohn, PhD**[1], **Apoorva P. Maru**[1], **Arjun Venkatesh, MD MBA**[4], **Maureen A. Smith, MD MPH PhD**[2,6,9], **Manish N. Shah, MD MPH**[1,5,6], **Eneida A. Mendonça, MD PhD**[7,8]

[1]BerbeeWalsh Department of Emergency Medicine, University of Wisconsin—Madison School of Medicine and Public Health. Madison, WI, USA.

[2]Health Innovation Program, University of Wisconsin–Madison. Madison, WI, USA.

[3]Department of Industrial and Systems Engineering, Department of Biostatistics and Medical Informatics, University of Wisconsin – Madison

[4]Department of Emergency Medicine and Center for Outcomes Research and Evaluation, Yale University School of Medicine, New Haven, CT. USA.

[5]Department of Medicine, Division of Geriatrics and Gerontology, University of Wisconsin—Madison School of Medicine and Public Health. Madison, WI, USA.

[6]Department of Population Health Sciences, University of Wisconsin—Madison School of Medicine and Public Health. Madison, WI, USA.

[7]Department of Pediatrics and Biostatistics, Indiana University School of Medicine, Indianapolis, IN, USA.

[8]Regenstrief Institute, Indianapolis, IN, USA.

[9]Department of Family Medicine and Community Health, University of Wisconsin School of Medicine and Public Health, Madison, WI, USA

INTRODUCTION:

Emergency Department (ED) visits for falls among older adults are often sentinel events for poor health trajectories, however challenges exist in defining fall-related visits in the ED. Diagnosis codes are a standard method for identifying falls in large administrative datasets. Although a common practice for many conditions, this strategy may miss many patients presenting with falls in the ED. Natural Language Processing (NLP) refers to a set of techniques by which language as written or spoken by humans can be rendered analyzable

Corresponding Author: Brian W Patterson MD MPH, bpatter@medicine.wisc.edu, @BPattersonMD, 800 University Bay Drive, Suite 310, Mail Code 9123, Madison, WI 53705, USA, +1(608) 265-6043.

AUTHORS CONTRIBUTIONS: BP, EM, and MNS conceived the research questions and study design. GCJ and APM developed and operationalized the human coding scheme. BP, EM, APM, and GCJ contributed to the iterative development of the algorithm. MS provided the limited data set containing hospital codes and demographic info and guidance on analysis framework. GCJ conducted all statistical analyses. BP and GCJ interpreted the results and drafted the manuscript. AV, APM, EM, MS, and MNS provided substantial feedback and revised the manuscript. All authors read and approved the final manuscript.

STATEMENTS OF CONFLICTS OF INTEREST: No conflicts of interest to declare

by computation. Our team recently developed and validated a simple, rules-based NLP system that accurately identified falls from the text of ED physician notes.[1] While this performance is encouraging, barriers exist to using this methodology for all investigations of falls. Researchers, administrators, and epidemiologists often face tradeoffs between large datasets which cover populations of interest but do not have granular clinical data, and datasets which are clinically generated which contain more information on fewer patients. Insurance claims, for instance, generate comprehensive data on large populations, making claims-based datasets attractive for studying fall epidemiology. They do not, however, contain provider notes or other clinical information. Furthermore, even when text is available in clinical datasets, application of NLP requires programming expertise for retrospective studies and potentially new information technology infrastructure at the health system level to process large numbers of notes or real time implementation within an electronic health record (EHR).[2,3]

Given these barriers to applying NLP, claims-based strategies will continue to have a role in studying falls. When deciding upon and applying a strategy for identifying falls, it is critical to understand relevant performance characteristics to interpret results or design interventions. The goal of this study was to compare performance characteristics of several fall identification strategies using EHR data from ED visits using manual chart abstraction as a gold standard.

METHODS:

Setting and Population:

We performed a retrospective observational study using EHR data at an academic medical center with ~60,000 ED visits per year. The study was IRB-approved. The current analysis utilizes the same set of 500 ED visits from unique patients aged 65 or older as did the previous NLP validation study.[1] Visits were randomly selected from the period between 12/13/2016 and 4/24/2017.

Measurements:

Seven different strategies were tested:

- Chief complaint (CC): As recorded in the EHR
- ICD codes: Two previously described ICD code-based fall identification strategies were tested.[4,5] These strategies were developed using ICD-9 codes (listed below). We created a crosswalk to migrate our ICD-10 data the more limited ICD-9 codes, then applied the ICD-9 fall definitions.
    - **a.** *Restrictive* – using only codes that *specifically* mention fall as a mechanism of injury (E880, E881, E884, E885, E888).[6,7]
    - **b.** *Broad* – to maximize sensitivity, including the above mechanism of injury codes as well as codes for injuries possibly sustained as the result of a fall (aforementioned E-codes OR ICD-9 codes 800–848, 850–854, and 920–924)[8–11]

- Combined approaches: As prior studies determined there to be subsets of falls that could only be identified using either hospital diagnosis coding OR chief complaint (not both),[6] it was important to assess the effectiveness of combining identification strategies. These two strategies combined chief complaint with either the Restrictive or Broad coding strategies (Restrictive/CC and Broad/CC respectively).
- NLP: Details of this approach presented in a separate publication.[1]
- Manual Abstraction (gold standard): Manual abstraction of ED notes performed as part of the initial validation study by two trained nonclinical reviewers demonstrating an inter-rater reliability of Kappa = 0.96.[1]

Analysis:

Each method was compared to manual abstraction, using Stata® 15 (StataCorp, College Station, TX) and MedCalc® (MedCalc Software, Ostend, Belgium) to calculate statistics. Accuracy, specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratios, and F1 scores were calculated for each comparison. F1 score refers to the harmonic mean of positive predictive value and sensitivity, and is often used as a measure of accuracy in the field of document retrieval. 95% confidence intervals (CI) were used to determine significant differences between fall identification strategies.

RESULTS:

Of the 500 ED visits, 494 were able to be matched with abstracted hospital administrative data. Human coders (manual abstraction) determined that 119 of the 494 ED provider notes (24.1%) explicitly mentioned a fall as a contributing reason for that visit. All 5 code-based strategies identified fewer falls, ranging from 72 (14.6%) identified using chief complaint to 116 falls (23.5%) identified using the Broad/CC strategy. Accuracy, sensitivity, specificity, PPV, NPV, and F1 score are presented in Table 1.

DISCUSSION:

Compared to manual chart review, NLP was the most accurate fall identification strategy, followed by the combination of a restrictive ICD code-based definition with chief complaint (Restrictive/CC strategy). The ICD code and chief complaint alone strategies performed more poorly. Among ICD code definitions, the broad definition had the same overall accuracy as the restrictive strategy, with the broad definition improving on the restrictive strategy's sensitivity at the cost of a significant loss in precision. Strategies combining ICD codes with chief complaint performed significantly better than those using only chief complaint or an ICD-based strategy, but still more poorly than the NLP approach with regard to both overall accuracy and F1 score. As demonstrated by the likelihood ratios and specificities, however, the chief complaint, restrictive coding, and combination of these strategies (Restrictive/CC) all had superior performance in the task of conclusively identifying (ruling in) a fall visit.

This study was performed at a single center, potentially limiting generalizability to other settings where coding and/or chief complaint may be subject to different recording procedures, or falls prevalence may differ. As with any study aimed at determining the performance characteristics of a testing methodology, this work is potentially subject to various biases of diagnostic accuracy. Incorporation and partial or differential verification biases are very unlikely based on our study design. Our choice to use human review of ED provider notes as a gold standard, however, may overestimate sensitivity of all methodologies if a significant number of falls were not mentioned by providers in their notes.[12]

Overall, our results suggest that analytic tools such as NLP should be rapidly adopted to improve the identification and appropriate treatment of falls in the ED. When NLP is unavailable, code-based definitions enhanced with chief complaint data where possible may have acceptable performance depending on the use case. When using these strategies however, their limitations should be recognized, especially the potential for coding based strategies alone to underestimate the true burden of fall visits.

ACKNOWLEDGEMENTS:

SPONSOR'S ROLE: Funding Sources/Disclosures: This project was supported by grant numbers K08HS024558 (Patterson) from the Agency for Healthcare Research and Quality, and 1K24AG054560 (Shah) from the National Institutes of Health (Shah). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality nor the National Institutes of Health.

REFERENCES:

1. Patterson BW, Jacobsohn GC, Shah MN, et al. Development and validation of a pragmatic natural language processing approach to identifying falls in older adults in the emergency department. BMC Med Inform Decis Mak. 2019;19(1):138. [PubMed: 31331322]
2. Velupillai S, Suominen H, Liakata M, et al. Using clinical Natural Language Processing for health outcomes research: Overview and actionable suggestions for future advances. Journal of biomedical informatics. 2018;88.
3. Kaggal VC, Elayavilli RK, Mehrabi S, et al. Toward a Learning Health-care System – Knowledge Delivery at the Point of Care Empowered by Big Data and NLP. Biomed Inform Insights. 2016;8(Suppl 1):13–22. [PubMed: 27385912]
4. Hoffman GJ, Hays RD, Shapiro M, Wallace SP, Ettner SL. Claims-Based Identification Methods and the Cost of Fall-related Injuries among U.S. Older Adults. Med Care. 2016;54(7):664–671. [PubMed: 27057747]
5. Kim SB, Zingmond DS, Keeler EB, et al. Development of an algorithm to identify fall-related injuries and costs in Medicare data. Inj Epidemiol. 2016;3:1. [PubMed: 27747538]
6. Tinetti ME, Baker DI, King M, et al. Effect of dissemination of evidence in reducing injuries from falls. N Engl J Med. 2008;359(3):252–261. [PubMed: 18635430]
7. Hoffman GJ, Hays RD, Shapiro MF, Wallace SP, Ettner SL. The Costs of Fall-Related Injuries among Older Adults: Annual Per-Faller, Service Component, and Patient Out-of-Pocket Costs. Health Serv Res. 2017;52(5):1794–1816. [PubMed: 27581952]
8. Patterson BW, Smith MA, Repplinger MD, et al. Using Chief Complaint in Addition to Diagnosis Codes to Identify Falls in the Emergency Department. J Am Geriatr Soc. 2017.
9. Bohl AA, Phelan EA, Fishman PA, Harris JR. How are the costs of care for medical falls distributed? The costs of medical falls by component of cost, timing, and injury severity. Gerontologist. 2012;52(5):664–675. [PubMed: 22403161]

10. Bohl AA, Fishman PA, Ciol MA, Williams B, Logerfo J, Phelan EA. A longitudinal analysis of total 3-year healthcare costs for older adults who experience a fall requiring medical care. J Am Geriatr Soc. 2010;58(5):853–860. [PubMed: 20406310]
11. Roudsari BS, Ebel BE, Corso PS, Molinari NA, Koepsell TD. The acute medical care costs of fall-related injuries among the U.S. older adults. Injury. 2005;36(11):1316–1322. [PubMed: 16214476]
12. Kohn MA, Carpenter CR, Newman TB. Understanding the direction of bias in studies of diagnostic test accuracy. Academic emergency medicine : Official journal of the Society for Academic Emergency Medicine. 2013;20(11).

Table 1:

Performance Characteristics of Fall Identification Strategies vs Manual Abstraction, with 95%CI

	Chief Complaint	Restrictive	Broad	Restrictive/CC	Broad/CC	NLP
Accuracy	90.1 (87.1–92.6)	92.7 (90.1–94.8)	92.1 (89.4–94.3)	95.3 (93.1–97.0)	93.3 (90.8–95.4)	97.1 (95.0–98.3)
Sensitivity	59.7 (50.3–68.6)	70.6 (61.5–78.6)	79.0 (70.6–85.9)	82.4 (74.3–88.7)	84.9 (77.2–90.8)	95.8 (90.5–98.6)
Specificity	99.7 (98.5–100.0)	99.7 (98.5–100.0)	96.3 (93.8–97.9)	99.5 (98.1–99.9)	96.0 (93.5–97.7)	97.3 (95.2–98.7)
PPV	98.6 (90.9–99.8)	98.8 (92.2–99.8)	87.0 (79.9–91.9)	98.0 (92.5–99.5)	87.1 (80.3–91.8)	91.9 (86.1–95.5)
NPV	88.6 (86.2–90.7)	91.4 (89.0–93.4)	93.5 (91.1–95.3)	94.7 (92.3–96.3)	95.2 (92.9–96.8)	98.7 (96.9–99.4)
LR +	224 (31.4–1590)	265 (37.3–1880)	21.1 (12.6–35.7)	154 (38.7–617)	21.2 (12.9–35.0)	35.9 (19.5–66)
LR −	0.40 (0.33–0.50)	0.29 (0.22–0.39)	0.22 (0.15–0.31)	0.18 (0.12–0.26)	0.16 (0.10–0.24)	0.04 (0.02–0.10)
F1 score	74.4	82.4	82.8	89.5	86.0	93.8

CC: Chief Complaint, PPV: Positive Predictive Value, NPV: Negative Predictive Value, LR+: Likelihood Ratio (Positive), LR−: Likelihood Ratio (Negative)