**Identifying Falls from the Electronic Health Record**

**Annotation Guidelines**

Version 1.6

04/07/2020

**1. Background**

Unstructured electronic health record (EHR) clinical notes contain a wealth of information on the diagnosis, treatment, and prognosis of aging persons. In particular, the concept of “geriatric syndromes” is evolving over time, and a wide range of geriatric syndromes have been identified, including: delirium, falls, incontinence, cognitive impairment, vision or hearing loss, pain, dizziness, sarcopenia, and self-neglect.1-6 However, for this proposal, we will adopt the more restrictive definition of geriatric syndromes proposed by Inouye and colleagues.1 Specifically, geriatric syndromes must occur primarily in older persons, must result from an accumulation of impairments to multiple systems, and must be associated with substantial morbidity and poor clinical outcomes.1

Under this definition, falls may be considered a key geriatric syndrome. Falls have been identified using only billing codes nearly 90% of the time from Veterans Health Administration ambulatory records.7 However, in preliminary research, we found that a significant number of falls were missed when only codes were used to identify the presence of falls among an aging population.

Therefore, in this study, we will mine EHRs using NLP techniques to identify the presence of falls among persons 65 years of age or older who are also participating in the Mayo Clinic Biobank. Our goal is to develop, benchmark, and evaluate NLP algorithms to identify persons with a history of falls and to identify new falls that occur during our study time frame.

Our study population consists of persons who were participants in the Mayo Clinic Biobank, and who were 65 years of age or older at the time of consent (n=22,772). We then identified persons who had an ICD-9 or ICD-10 code for “accidental fall” after the index date (enrollment in the Mayo Clinic Biobank). We used the REP code sets for “falls” (<http://repweb.mayo.edu/Diagnosis-Code-Sets>).

For initial annotation, we will focus on persons with an ICD-9 or 10 code for “accidental fall”.

Fall cohort definition: 30 days before and after the first diagnosed fall index event.

**2. Annotation tool**

The annotation tool for this project is the Multi-document Annotation Environment (MAE), a Java-based natural language annotation software package. MAE is a non-web-based annotation tool. All annotation tasks will be defined in a document type definition (DTD) file. Due to its lightweight feature, the software can be easily shared and updated across multiple sites without configuration and testing.

Download: <https://github.com/keighrim/mae-annotation>

EHR notes for an age-stratified sample of 30 persons with a code for “fall” will be loaded into the MAE tool. All notes will be loaded beginning with the patient’s index date (date of first fall diagnosis) through 2 years after the index date or 12/31/2018, whichever is first.

**3. Instructions**

Please use this document for annotation to make sure we are consistently collecting the same information in the same way.

Annotators will be given access to a sample of medical records in the MAE tool. For each medical record, annotators will identify and highlight keywords pertaining to falls. Then, annotators will assign attributes to these keywords. Finally, annotators will assign a definition of delirium that seems most appropriate for the patient. All this must be done using the software MAE.

**Annotation Flow:**

Guideline Development Phase: Each annotator will annotate **30** clinical notes. Then, inter-annotator agreement (IAA) will be calculated and the guidelines will be revised. This process will continue until a high agreement is reached.

Annotation Phase: Approximately XXX clinical notes (10% sample) will be annotated. After the annotations are completed, we will compute IAA, resolve disagreements, and clarify the guidelines. Each document will be independently annotated by 2 annotators.

For each medical record, annotators will identify and highlight relevant keywords or phrases whenever they appear. After highlighting the keywords, the annotators will assign definitions to the keywords. In addition, annotators will assign attributes to the relevant keywords or phrases, as appropriate. The annotation will end when all medical records of that patient for a 2-year period are reviewed.

Please record the time spent in annotating each document. Please take your time and try your best to annotate all concepts and properties.

Adjudication: The final gold standard annotations will be created by combining the individual experts’ annotations followed by adjudication of the mismatches. The first 20 clinical notes will be added to the gold standard but excluded from the final IAA computation.

**4. Guideline**

Annotators will have four responsibilities for each medical record. Annotators will:

1. Highlight key words related to falls
2. Search for expressions that are not highlighted, but that are conceptually similar to a keyword;

(3) Assign attributes to each highlighted expression;

(4) Assign a definition of fall that seems most appropriate for the patient.

**4.1 Definition**

Annotators will assign a definition of whether a fall is present or absent. There are multiple definitions of falls, but we will use the all-inclusive definition proposed by the American Nurses Association. National database for nursing quality indicators: Guidelines for data collection and submission on quarterly indicators, version 5.0 2005. Specifically, they propose the definition of:

“An unplanned descent to the floor (or extension of the floor, e.g., trash can or other equipment) with or without injury. All types of falls are included, whether they result from physiological reasons or environmental reasons.”

**4.1 Keywords and attributes**

Once you highlight a keyword or phrase, you will be prompted to fill out a set of attributes about the highlighted text. Specifically, most keywords or phrases have multiple possible attributes, and each attribute has categories that must be selected. Keywords or phrases are listed by lowercase letter. Attributes are indented from keywords or phrases.

**1. Fall concept:**

**Direct mention of fall: fall/fell, tripped, slipped, slid**

**Direct indication of fall: syncope**

Certainty:

* Negated: the problem does not exist in the patient
  + Patient does not have SYMPTOM [infection].
  + No history of infection
  + His infection resolved
* Possible: patient may have a problem, but there is uncertainty expresed in the note. Possible takes precedence over negated, so terms like “probably not” or “unlikely” categorize problems as being possible just as “probably” and “likely” do
  + This is very likely to be an asthma exacerbation.
  + Doctors suspect an infection of the lungs.
  + Questionable / small chance of pneumonia.
  + Pneumonia is possible / probable
  + Suspicion of pneumonia
  + We are unable to determine whether she has leukemia.
  + It is possible / likely / thought / unlikely that she has pneumonia
  + We suspect this is not pneumonia
  + This is probably not cancer
  + Pneumonia unlikely
* Hypothetical: medical problems that the note asserts the patient may develop
  + If you experience wheezing or shortness of breath
* Confirmed
  + Patient has infection
  + He has history of infection

Status:

* Present
* Follow up
* History

Experiencer:

* Patient
* Family member
* Other

Exclusion: choose Yes, if you find a piece of information that does not fit into one of the categories mentioned, but could be useful information. Please elaborate in the comments if needed. i.e. creatinine fell, occurred in the fall

* Yes
* No

**2. Date:** first mention of fall after index date (date of enrollment into Biobank; month/day/year)

* Include “today”, “yesterday”, actual dates

**3. Consequence of falls**

* Injuries: fractures, bruise, hematomas, etc
* Or LACK of injuries: “fall without injury”

**4. Fall temporal expression**

**5. Location**

**6. Indirect mention of falls**

i.e seizure, syncope/fainting, narcolepsy

Certainty:

* Negated: the problem does not exist in the patient
  + Patient does not have SYMPTOM [infection].
  + No history of infection
  + His infection resolved
* Possible: patient may have a problem, but there is uncertainty expresed in the note. Possible takes precedence over negated, so terms like “probably not” or “unlikely” categorize problems as being possible just as “probably” and “likely” do
  + This is very likely to be an asthma exacerbation.
  + Doctors suspect an infection of the lungs.
  + Questionable / small chance of pneumonia.
  + Pneumonia is possible / probable
  + Suspicion of pneumonia
  + We are unable to determine whether she has leukemia.
  + It is possible / likely / thought / unlikely that she has pneumonia
  + We suspect this is not pneumonia
  + This is probably not cancer
  + Pneumonia unlikely
* Hypothetical: medical problems that the note asserts the patient may develop
  + If you experience wheezing or shortness of breath
* Confirmed
  + Patient has infection
  + He has history of infection

**Note:** Annotators should also highlight variations of keywords. For example, the keyword “fall” has variations such as “fell”, <others?>.

**Comments**: Please feel free to comment anything that could be useful information.

## Fall Semantic Analysis

T: True fall

* To drop or descend under the force of gravity, as to a lower place through loss or lack of support.
* To come or drop down suddenly to a lower position

N: True fall but

* Negated
* Hypothetical
* Uncertain/possible

H: Historical fall

R: Fall risk

P: Fall prevention

E: Exclusion: e.g. fall asleep

* To become less or lower; become of a lower level, degree, amount, quality, value, number, etc.; decline:
  + The temperature **fell ten degrees**. Stock prices fell to a new low for the year.
* To subside or abate.
  + To sink or **fall to the bottom**
* Extend downward; hang down:
  + Her hair **falls to her shoulders**.
* To become lowered or directed downward, as the eyes:
  + My **eyes fell** before his steady gaze.
* To become lower in pitch or volume:
  + Her **voice fell**, and she looked about in confusion.
* To succumb to temptation or sin, especially to become unchaste or to lose one's innocence.
  + **fall from grace**
* To lose status, dignity, position, character, etc.
  + **fall afoul**
* To succumb to attack:
  + The city **fell to the enemy**.
* To be overthrown, as a government.
  + **can fall** if they lose control of too much territory
* To drop down wounded or dead, especially to be slain:
  + to **fall in battle**.
* To pass into some physical, mental, or emotional condition:
  + to **fall asleep**; to fall in love.
* To envelop or come as if by dropping, as stillness or night.
* To issue forth:
  + Witty **remarks fall easily** from his lips.
* To come by lot or chance:
  + The chore **fell to him**.
* To come by chance into a particular position:
  + to **fall among thieves**.
* To come to pass, occur, or become at a certain time:
  + **Christmas falls** on a Monday this year. The rent falls due the first of every month.
* To have its proper place:
  + The **accent falls** on the last syllable.
* To come by right:
  + The inheritance **fell to the only living relative**.
* To be naturally divisible (usually followed by into):
  + The story **fell into two distinct parts**.
* To lose animation; appear disappointed, as the face:
  + His **face fell** when he heard the bad news.
* To slope or extend in a downward direction:
  + The **field falls** gently to the river.
* To be directed, as light, sight, etc., on something:
  + His **eyes fell upon** the note on the desk.
* To collapse, as through weakness, damage, poor construction, or the like; topple or sink:
  + The **old tower fell** under its own weight. The cake fell when he slammed the oven door.
* (of an animal, especially a lamb) to be born:
  + **Two lambs fell** yesterday.

L: Lack of context: e.g. only single keyword fall or fell was mentioned

References

1. Inouye SK, Studenski S, Tinetti ME, et al. Geriatric syndromes: clinical, research, and policy implications of a core geriatric concept. *J Am Geriatr Soc* 2007;55(5):780-91. doi: 10.1111/j.1532-5415.2007.01156.x [published Online First: 2007/05/12]

2. Carlson C, Merel SE, Yukawa M. Geriatric syndromes and geriatric assessment for the generalist. *The Medical clinics of North America* 2015;99(2):263-79. doi: 10.1016/j.mcna.2014.11.003 [published Online First: 2015/02/24]

3. Vetrano DL, Foebel AD, Marengoni A, et al. Chronic diseases and geriatric syndromes: The different weight of comorbidity. *Eur J Intern Med* 2016;27:62-7. doi: 10.1016/j.ejim.2015.10.025 [published Online First: 2015/12/09]

4. Koroukian SM, Schiltz N, Warner DF, et al. Combinations of Chronic Conditions, Functional Limitations, and Geriatric Syndromes that Predict Health Outcomes. *J Gen Intern Med* 2016;31(6):630-7. doi: 10.1007/s11606-016-3590-9 [published Online First: 2016/02/24]

5. Cruz-Jentoft AJ, Landi F, Topinkova E, et al. Understanding sarcopenia as a geriatric syndrome. *Curr Opin Clin Nutr Metab Care* 2010;13(1):1-7. doi: 10.1097/MCO.0b013e328333c1c1 [published Online First: 2009/11/17]

6. Pavlou MP, Lachs MS. Could self-neglect in older adults be a geriatric syndrome? *J Am Geriatr Soc* 2006;54(5):831-42. doi: 10.1111/j.1532-5415.2006.00661.x [published Online First: 2006/05/16]

7. Luther SL, McCart JA, Berndt DJ, et al. Improving identification of fall-related injuries in ambulatory care using statistical text mining. *Am J Public Health* 2015;105(6):1168-73. doi: 10.2105/AJPH.2014.302440 [published Online First: 2015/04/17]