INSIGHTS FROM NARRATIVES OF FALL-RELATED INJURIES: CAUSE VS. EFFECT

Analysis by Wendy Phillips

Introduction

Falls are the number-one cause of injuries, both fatal and non-fatal, in older adults (individuals aged ≥65 years)[1]. In 2014, 28.7% of older adults in the United States reported having fallen at least once in the previous year. In this analysis, I use the contents of clinical text to determine 1) how consequences differ when striking the floor versus striking an object in a falling event and 2) which objects most commonly cause falls versus which are most often struck during a fall.

Methods

Data set: Data from the National Electronic Injury Surveillance System (NEISS) were examined. The data table consisted of 115,128 rows and 22 columns. Each row described a falling incident and associated metadata. The primary focus of the analysis was the clinical narrative column. **Figure 1** shows the distribution of the number of characters in the narratives. Notably, 75% of the entries had less than 148 characters, reflecting the limited detail available in most entries. Given the frequent use of abbreviations, symbols, and non-standard spacing, a function was applied to standardize the narratives.

Model training: After investigating the clinical narratives in the dataset, I chose to build a novel Natural Language Processing (NLP) model to extract explicit types of information from the narratives. Specifically, I used the spaCy (v6.1) toolset to build a model that would identify spans of text that belonged to any of the six categories described in **Table 1**. I chose those specific six categories based on the frequency with which they occurred in narratives and their potential usefulness. I semi-manually labeled narratives to develop a model. I began by manually labeling 400 randomly chosen narratives and using those for building an initial model. I manually corrected spans predicted using that model and performed several more iterations of the training-predicting-correcting cycles. My final labeled set contained 1635 unique clinical narrative entries. Of those, 423 entries had no spans deserving a label. The entire labeled set contained a total of 1843 labeled spans. These were randomly split in a ~2:1 ratio to create training and test sets.

Statistical tests: The *scipy.stats.contingency odds_ratio* test was used to determine if the probability of certain outcomes were statistically significant between groups.

Results

Model performance: The final model had these performance values: F1 = 0.84, recall = 0.79 and precision = 0.90. Examples of actual spans predicted using the constructed model are shown in **Figure 2**.

Striking an object vs. striking the floor: A patient who struck the floor had 1.9x (95% confidence interval of 1.8, 2.0) greater odds of being admitted/hospitalized than one who had struck an object (**Figure 3**, left). In terms of diagnoses, a patient who struck the floor had 2.4x (95% C.I. = 2.3, 2.5) greater odds of having a fracture and 0.34x (95% C.I. = 0.31, 0.36) the odds of having a laceration than a patient who struck an object (**Figure 3**, right).

Objects instigating falls vs. objects hit during a fall: The types of objects that instigated falls differed from those that were struck during a fall (**Figure 4)**. Notably, dogs and cats were among the 14 most frequent objects in spans tagged as being caused by objects (OBJ, Table 1). This is perhaps in contrast to other scientific findings that having pets can increase life expectancy[2]!

Falling down stairs or having to wait for help: While many spans related to these two categories were identified, they await further analysis.

Conclusion

The application of the model herein developed yielded important insights into the causes and effects of falls in older adults. Even a fairly limited analysis of spans identified by the model reveals information possibly very useful in helping prevent and limit injury from falling events.

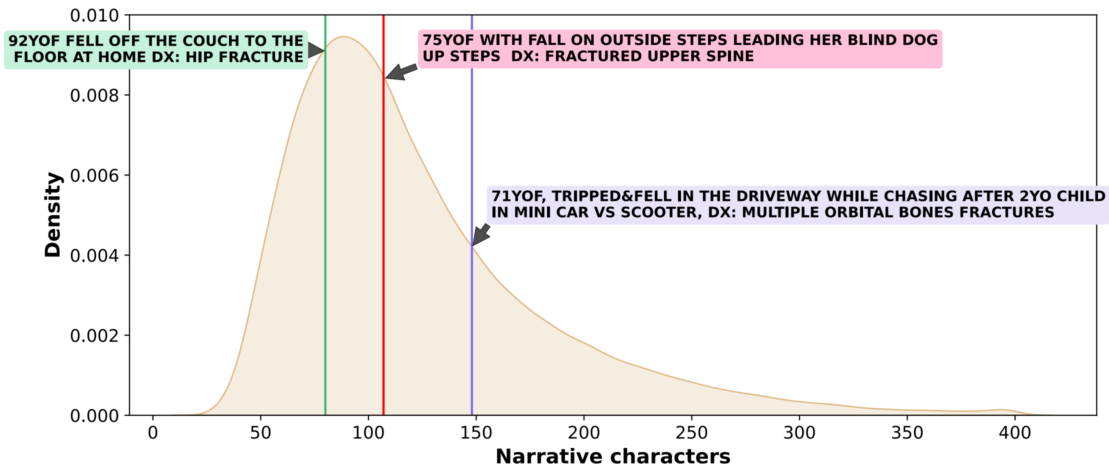

Figure 1. The distribution of the length of characters in the clinical narratives. Vertical lines indicate the 25th, 50th, and 75th percentiles of the distribution. The text strings shown in the boxes are examples of real narrative entries with the given narrative character lengths. This reveals the very limited details available in most narratives available.

Table 1. Categories used to classify spans of text

Category	Description	Example
Struck object (SO)	Striking a body part on or falling on an object	hit head on tub
Struck floor (SF)	Specific statement of body part contacting floor	landing onto head
Struck unknown (SU)	A body part struck something, not stated what	struck head
Caused by object (OBJ)	Tripping over something or something giving way	tripped over the cord
Involved stairs (STR)	Event happened on stairs	fall down 6 steps
Found by other (WT)	Was found by someone else	found by family

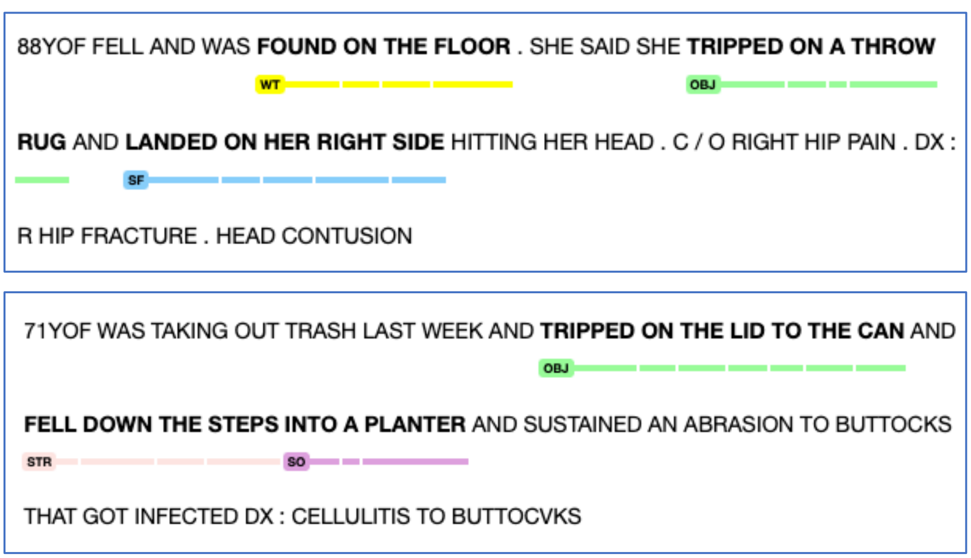

Figure 2. Examples of text spans and categories (from real narratives) predicted by the trained model. All span categories are shown within these two examples.

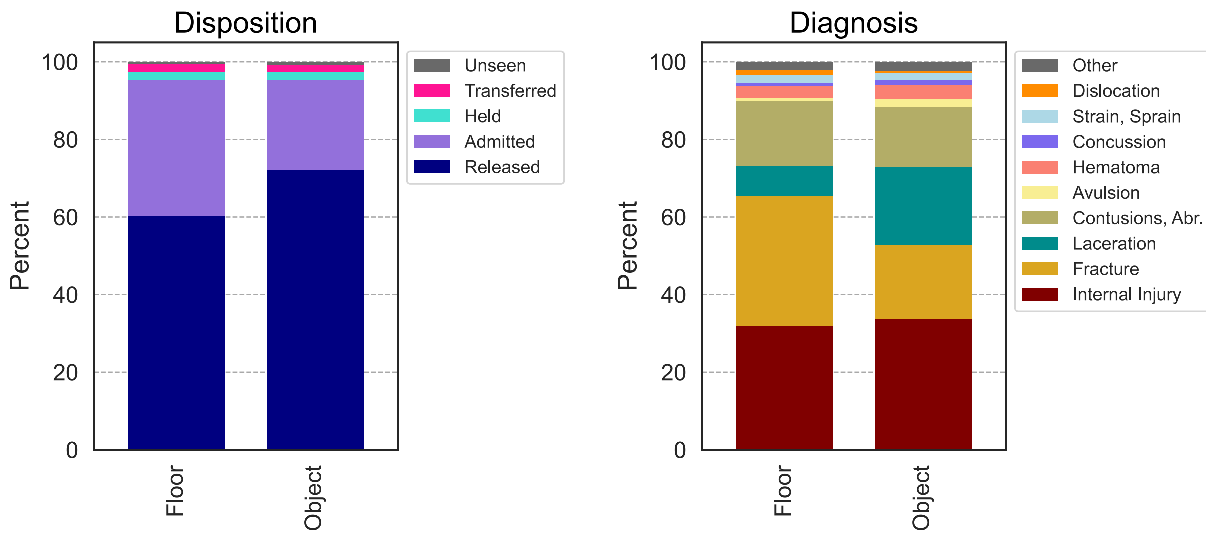

Figure 3. Outcomes differ between patients who struck the floor and those who struck an object. Left: Patients who struck the floor had 1.9x the odds of being admitted/hospitalized than those who struck an object. Right: Patients who struck the floor had 2.4x the odds of having fractures and 0.34x the odds of having lacerations than those who struck an object.

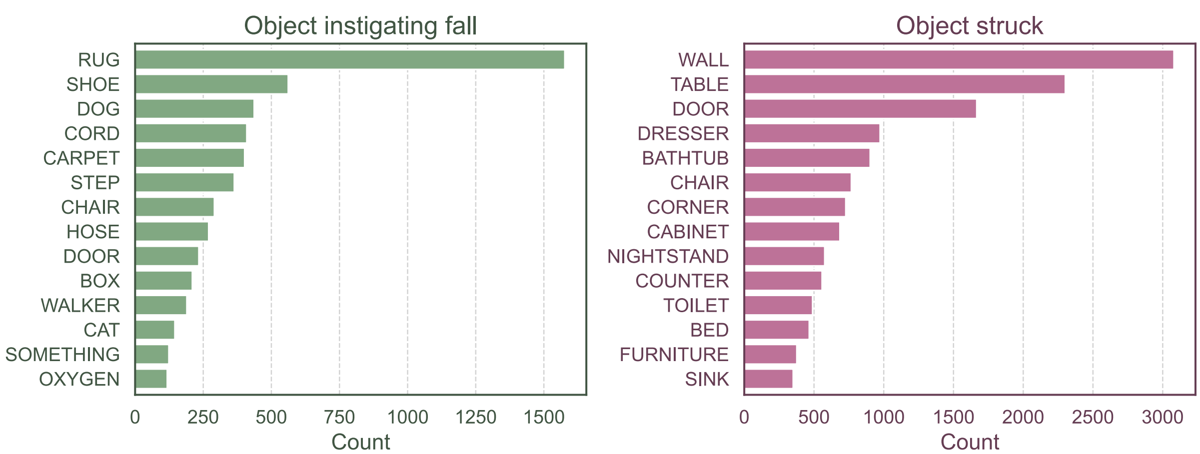

Figure 4. The types of objects that were involved in instigating falls (left) differed from those that were struck during a falling event (right). Only two objects, door and chair, were shared between the fourteen most frequent object words counted in spans of labeled text.

REFERENCES

1, Bergen G, Stevens MR, Burns ER. Falls and Fall Injuries Among Adults Aged ≥65 Years — United States, 2014. MMWR Morb Mortal Wkly Rep 2016;65:993–998.
DOI: http://dx.doi.org/10.15585/mmwr.mm6537a2

2. https://doi.org/10.1161/CIRCOUTCOMES.118.005342