

# Identifying Falls Risk Screenings Not Documented with Administrative Codes Using Natural Language Processing

Vivienne J Zhu, MD, MS,<sup>1</sup> Tina D Walker, BSN,RN,<sup>2</sup> Robert W Warren, MD, MS, MPH<sup>2</sup>  
Peggy B Jenny, BSN, RN,<sup>2</sup> Stephane Meystre, MD, PhD,<sup>1</sup> Leslie A Lenert, MD, MS<sup>1</sup>

<sup>1</sup>Biomedical Informatics Center and <sup>2</sup>Information Solutions  
at Medical University of South Carolina, Charleston, South Carolina

## Abstract

*Quality reporting that relies on coded administrative data alone may not completely and accurately depict providers' performance. To assess this concern with a test case, we developed and evaluated a natural language processing (NLP) approach to identify falls risk screenings documented in clinical notes of patients without coded falls risk screening data. Extracting information from 1,558 clinical notes (mainly progress notes) from 144 eligible patients, we generated a lexicon of 38 keywords relevant to falls risk screening, 26 terms for pre-negation, and 35 terms for post-negation. The NLP algorithm identified 62 (out of the 144) patients who falls risk screening documented only in clinical notes and not coded. Manual review confirmed 59 patients as true positives and 77 patients as true negatives. Our NLP approach scored 0.92 for precision, 0.95 for recall, and 0.93 for F-measure. These results support the concept of utilizing NLP to enhance healthcare quality reporting.*

## Background

Falls are a leading cause of injury in the elderly population and a significant morbidity and mortality risk factor. Approximately 35-40% of generally healthy people over 65 fall at least once each year.<sup>1</sup> The risk of falling increases proportionately with age. At 80 years, over 50% of seniors fall annually.<sup>2</sup> Over 20% of falls cause a serious injury such as hip fracture or brain trauma.<sup>3,4</sup> In 2012–2013, 55% of all deaths due to unintentional injury among the elderly came from falling.<sup>5</sup> The estimated direct medical costs for fall injures in the U.S. alone are \$31 billion annually.<sup>6</sup>

Clinical screening of fall risk can significantly prevent falls.<sup>7,8</sup> However, less than half of the elderly patients who fall discuss the fall with their health care providers.<sup>9</sup> In 2001, The American Geriatrics Society and British Geriatrics Society (AGS/BGS) published a clinical practice guideline on falls prevention and management in older persons. The AGS/BGS guideline specifies that, at least annually, providers should ask all elderly patients if they had a fall or had no fall during the prior year.<sup>10</sup> Moreover, falls risk screening is a core quality measurement required by the Centers for the Medicare and Medicaid Services (CMS) payment incentive program under the Affordable Care Act (ACA).<sup>11</sup> Primary Care Providers (PCP) can receive reimbursement for falls risk screening through voluntary participation in the Physician Quality Reporting System (PQRS).<sup>12</sup> In 2015, the Medicare Access and CHIP Reauthorization Act (MACRA) became law, offering value-based alternative payments for reimbursing physician services through a Merit based Incentive Payment System (MIPS).<sup>13</sup> In performance year 2017, falls risk screening remains a high priority measure in MIPS.<sup>14</sup>

Both PQRS and MIPS require reporting the percentage of patients aged 65 years and older who had a risk screening for falls within the prior 12 months. They both use a set of CPT or HCPCS codes to define the denominator and numerator for this quality measurement. However, coded data alone may not completely or accurately reflect clinical activities that providers performed during their practice. Clinical narratives, such as visit notes, admission notes, progress notes, consultation notes, and nursing notes contain important information and are increasingly available in electronic form. However, obtaining structured information from free-text clinical narratives is a major challenge. Natural language processing (NLP) is a technology that uses computer-based linguistics and machine learning approaches to extract and classify information automatically from free-text data.<sup>15</sup> NLP offers an opportunity to identify clinical information from narrative documents and holds the promise of enhancing data availability and quality for diagnostic, patient safety, clinical decision support, and quality performance reporting.<sup>16-19</sup> In this study, we developed an NLP approach to identify documented falls risk screening in clinical notes of patients lacking discretely coded falls risk screening. We also formally evaluated the NLP algorithm performance against a gold standard (i.e. domain expert manual review). We hypothesized that this NLP approach could correctly identify more falls risk screening in electronic health records (EHR) than the quality metrics based on by administrative codes. Confirmation of this hypothesis would support using NLP approaches to enhance current quality reporting systems.

## **Methods**

### **Data resource and NLP software**

The primary study setting was the Medical University of South Carolina (MUSC). MUSC is an academic medical science center with inpatient, outpatient, and emergency facilities serving Charleston, South Carolina, and surrounding areas. MUSC has had the EpicCare EHR system (Epic Systems Corp., Verona, WI) in place for outpatient care since 2012 and for inpatient care since 2014. A Research Data Warehouse (RDW) copies the Epic data warehouse and serves as the data repository for clinical research. This study tests the feasibility of identifying mentions of falls risk screenings in clinical notes using NLP for the defined patient population. Clinical narratives collected during 2015 for MUSC Medicare outpatients are available for analyses, including progress notes and consult notes. This study was approved by the MUSC data access committee. We used commercial NLP software (Linguamatics I2E version 4.4, Cambridge, United Kingdom) licensed by MUSC to index, parse, and query each clinical note for this project. Linguamatics I2E (I2E) applies concept-based indexing techniques to electronically identify key words/phrases from text documents and map them to concepts in the UMLS Metathesaurus.<sup>20</sup> Then, I2E queries retrieve information for reports meeting a user-defined set of criteria through a user-friendly interface to define syntactic and semantic representations.

### **Measurement of falls risk screening**

For performance year 2015, MUSC used screening for future falls risk (CMS193v3) to report the percentage of Medicare patients 65 years of age and older who were screened for falls risk during the measurement period.<sup>21</sup> The numerator is the number of patients who were screened for future falls risk at least once during the measurement period. Two CPT codes, 3288F and 1100F, are required on the claim form to submit this numerator option. The denominator is the number of patients from the initial Medicare outpatient population. Exceptions for the denominator are documented medical reasons not to screen for falls risk, including the patient being non-ambulatory or wheelchair bound. We developed an NLP algorithm and conducted a performance evaluation based on CMS193v3.

### **Development of the lexicon for falls risk screening**

In practice, providers record clinical findings and activities as narratives for each patient encounter, including symptoms, health histories, physical exam results, consultation services, and care plans. In these clinical narratives, NLP can identify mentions of falls risk screening that providers performed during the encounter. A domain expert, a nurse analyst with extensive quality measure experience, manually reviewed 144 patient charts and generated an initial list of terms that were commonly used to represent falls risk screening. These initial terms included “fall,” AMPAC (Activity Measure for Post Acute Care), and ADLs (Activities of Daily Living measures) plus mention of fall. Without domain expert engagement, the NLP informatics team then developed a draft full terminology set based on this initial list. We also generated modifiers (e.g. “evaluation,” “associated symptoms,” “risk screen,” “history of,” etc.). I2E provides default detection of pre- and post-negation, a collection of regular terms that are negative mentions (e.g. “no,” deny,” “negative”). However, the I2E default clinical negation detection is not sufficient to exclude a “false” falls risk screening in clinical notes. In fact, when a patient is identified as “no fall” or “denied falls” in clinical notes, a falls risk screening had likely been conducted during the encounter. Therefore, we considered such negative mentions of falling as true cases of falls risk screening. Another area of ambiguity included words associated with “falling” that might represent lab value decreases (e.g., “hgb fall,” “bone density fall,” “ejection fraction fall”), vital sign reductions (e.g., “weight fall,” “BP fall”), seasons (e.g., “last fall,” “in the fall”), problems (e.g., “fell on ice,” “fall with fracture”), or others (e.g., “fall asleep,” “drain fall,” “fell out”). In context these terms are not relevant to a fall screening and were reconsidered as valid negations for this study. We also included in our negation criteria terms that reliably indicate a non-ambulatory patient (e.g., “wheelchair bound,” “bed bound”) because a non-ambulatory patient is excluded from the denominator for CMS falls risk screening measures. For each term, we utilized the I2E morphologic variants functions and I2E built-in ontology to generate a set of spelling variants, acronyms, and abbreviations; we then queried these terms against clinical notes to extract any relevant lexical representations to form an enhanced and refined list. The domain expert and the NLP informatics team came to consensus agreement, creating the final lexicon. A final lexicon list was imported to the I2E customized macros, which allows for reuse and refining.

### **Development of NLP algorithms to identify falls risk screenings**

We developed a set of NLP queries to identify falls risk screening mentions in clinical notes (progress notes and consults notes), using the following criteria: A) mentions of falls by the lexicon; and B) excluding fall mentions with negations. These NLP queries were designed to capture semantic information, syntactic patterns, and clinical negations in order to translate a documented falls risk screening to structured data elements including: 1) patient MRN;

2) falls risk screening; 3) note create date; 4) note updated date; 5) author type; 6) note ID; and 7) type of clinical note. We used I2E 4.4 to index, query, flag, and count the number of query hits within each type of clinical note. We used all clinical notes for 144 patients to develop an NLP algorithm for each variable, evaluated the results of the NLP queries independently and in combination against the gold standard of expert chart review. These NLP and chart review evaluations were done independently by a domain expert, who was blinded to the patient source. Discrepancies between query results and the manual expert review led the informatics team to iteratively refine and develop the I2E query algorithms until sensitivity and specificity could not be further improved or could not reach predefined thresholds. A multiple I2E query combining these 7 queries was established to produce a structured output table to store information extracted by the 7 queries.

### NLP algorithm performance evaluation

For clinical notes from all 144 patients, we compared the results generated from the I2E multiple query to the results from the gold standard. A domain expert further validated the results for both positive and negative cases of NLP-detected falls risk screening by manual chart review. The average times for falls risk screening case identification by both the I2E multiple query and manual review of clinical notes were reported and compared. Furthermore, we calculated three standard performance measures for the NLP algorithm: the precision, recall, and *F*-measure. Precision (exactness) is the proportion of true positives to the total number of algorithm-identified cases; in contrast, recall (completeness) is the proportion of true positives that are retrieved by algorithms.<sup>22</sup> *F*-measure is a weighted harmonic mean of precision and recall ( $F\text{-measure} = 2 * \text{precision} * \text{recall} / (\text{precision} + \text{recall})$ ); an *F*-measure reaches its best value at 1 and worst at 0. Finally, for all false positives and negatives generated by the NLP algorithms, the reasons for false classification were manually determined and summarized to improve the algorithm.

## Results

### Falls risk screening lexicon

Using the Epic Medical Shared Savings Program (MSSP) abstraction assistant tool, we identified 144 patients who were Medicare outpatient beneficiaries and had no discrete fall screening results recorded during 2015. The I2E multiple query processed 1,588 documents (1,575 progress notes and 13 consult notes) from these 144 unique patients within 0.5 seconds comparing with 10 hours by manual review. The average number of document per patient was 11 (max: 106; minimum: 2). We developed a set of I2E queries to search the initial key words (“fall,” “ADLs,” “AMPC,” etc.) against these 1,588 documents. After iterative evaluations between the keywords hit and the original documents, we developed a lexicon of falls risk screening and negations. The final lexicon presented wide variations (Table 1). A total of 38 terms associated with falls risk screening resulted in 397 hits in 223 documents from 93 unique patients. The leading keywords were “fall” and its morphologic variants (236 hits, 67.4%), “fall risk” (42 hits, 11.7%), and “denies any falls/no falls/Denies falling/no recent falls” (22 hits, 6.2%). Among 26 pre-negations, the leading terms were “fracture fall,” “this fall,” and “in the fall.” Among 35 post-negations, the leading terms were “fall at,” “fall in,” “fall on,” and “fall of [a year].” The most common note in which these instances occurred was a progress note, and the most common author types were physician, physical therapist, and physician assistant (Figure 1).

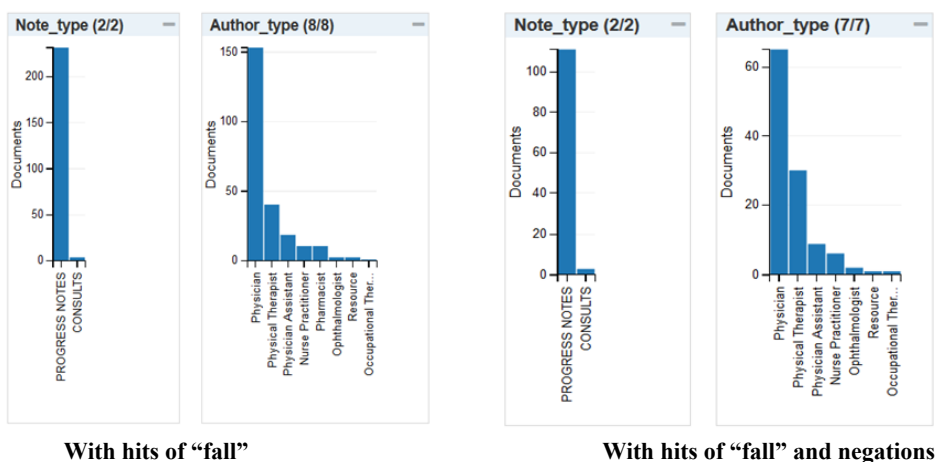


Figure 1: Note type and author type.

**Table 1. Lexicon of falls risk screening and frequency (example)**

Terms relevant to falls risk screening		Pre-negation		Post-negation	
fall	113	fracture fall	13	fall at	40
falls	47	this fall	6	fall in	35
falling	29	in the fall	6	fall on	28
Fall Risk	24	from fall	5	fall + year	20
fell	18	last fall	4	fall asleep	18
fall risk	15	h/o	2	fall fracture	4
Fallen	13	would fall	2	fell out	4
Fall	13	something fall	2	fell down	3
LE AMPAC	12	fractures fall	2	fall off	2
denies any falls	11	tripped fall	2	fall break	2
No falls	8	weight fall	1	fell against	1
AM-PAC Raw Score	4	Tripped fall	1		
fell	3	Indication fall	1		
Denies falling	2	tissue fall	1		
No recent falls	2	fever fall	1		
Denies any fractures or falls	2	Pessaries fall	1		
History of falls	2	level fall	1		
Falls risk scale score	1	every fall	1		
Risks for falls	1	indicated fall	1		
risk of fall	1	ejection fraction	1		
No risk of fall	1				

**NLP algorithm performance**

The multiple I2E query combining seven I2E queries produced a structured output table which extracted a patient MRN, note ID, keywords for falls risk screen, the sentences where keywords hit, note type, author type, and note creation date. It also provided a link to the original document, which the NLP developer and domain expert could review during the development and evaluation phases (Figure 2). Among 144 patients, I2E query identified 93 patients with a likely mention of falls risk screening from 223 documents. Applying both pre- and post-negations, NLP identified 62 patients (40.2%) with falls risk screening from 118 documents. The domain expert’s manual review identified 64 patients who had falls risk screening and 80 patients who did not have a falls risk screening. The I2E algorithm and manual review identified, in common, 59 patients who had falls risk screenings and 77 patients who had no falls risk screening. Three non-ambulatory patients were identified by both NLP and manual review. The I2E query for falls risk screening had a precision of 0.92, recall of 0.95, and an F-measure 0.93 (Figure 3). A major reason for false negatives and false positives was that our NLP approach could not clearly distinguish a fall event as the reason of a clinical visit from a falls risk screening in some clinical notes (Table 2).

EPIC_Falls_144_sample	NoteID	keywords	Note_type	Author_type	create_date	#Docs	Doc#	Doc	#Hits	Hit#	Hit
	137297187	falls	PROGRESS NOTES	Physician Assistant	4/30/2015	1	1	116106	1	1	1... denies any traumatic events or falls, no sense of knee ...
	107184973	Fall AM-PAC Raw Score fall	CONSULTS	Physical Therapist	1/8/2015	1	1	72710	1	1	1... Contact Precautions;IV;Fall Risk;Other (Left ... 1 AM-PAC Raw Score 8 Conversion Table 8 - ... Prevention) risk factors (fall risk, further skin breakdown...); risk factors (fall risk, further skin breakdown...
	107370390	Fall AM-PAC Raw Score	CONSULTS	Occupational Therapist	1/9/2015	1	1	72622	1	1	1 Precautions Contact Precautions;Fall Risk;Other (left ... 1 AM-PAC Raw Score 6 Conversion Table 6 - ...
	107982185	fall	PROGRESS NOTES	Physical Therapist	1/12/2015	1	1	72448	1	1	1... ); risk factors (fall risk, further skin breakdown...
	108623081	AM-PAC Raw Score	PROGRESS NOTES	Physical Therapist	1/14/2015	1	1	72554	1	1	1 1 AM-PAC Raw Score 8 Conversion Table 8 - ...
	133648225	fall	PROGRESS NOTES	Physician	4/16/2015	1	1	70985	1	1	1... to unsteady gait and high fall risk and Patient is bed ...
	117623464	no Risk for Falls Risk for Falls	PROGRESS NOTES	Physician	2/17/2015	1	1	49383	1	1	1 no Risk for Falls?
	123010111	fallen No falls	PROGRESS NOTES	Physician	3/9/2015	1	1	68323	1	1	1 He has not fallen , he is not dragging ... No falls.
	172450902	fallen Fall	PROGRESS NOTES	Physician Assistant	9/4/2015	1	1	77918	1	1	1... the past year have you fallen or had a near fall... you been bothered by: Fall or dizzy when standing - ...
	138081693	denies any pain, falling	PROGRESS NOTES	Physician	5/4/2015	1	1	50083	1	1	1... in her lower extremities but denies any pain, falling or tripping.
	176478541	fallen Fall	PROGRESS NOTES	Physician Assistant	9/21/2015	1	1	54362	1	1	1... the past year have you fallen or had a near fall... you been bothered by: Fall or dizzy when standing - ...

**Figure 2: Example of I2E query output**

Gold standard	NLP algorithm	
	positive	negatives
positives	59 (TP)	5 (FN)
negatives	3 (FP)	77 (TN)

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) = 59 / (59 + 5) = 59 / 64 = 0.92$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) = 59 / (59 + 3) = 59 / 62 = 0.95$$

$$\text{F-measure} = 2 * \text{precision} * \text{recall} / (\text{precision} + \text{recall})$$

$$= 2 * 0.92 * 0.95 / (0.92 + 0.95) = 0.93$$

\*TP=true positive; FP=false positive; FN=false negative; TN=true negative.

**Figure 3. Results of manual review and I2E algorithm**

**Table 2. Detailed information of false positives and false negatives of falls risk screening**

**False Negatives**

No falls but reports some near falls at home  
 She denies any trauma or falls in the last 6 months,  
 One week ago, she tripped and fell on her knees her  
 Denies falls or fractures, urinary stones, heat...

**False Positives**

chest pain related to the fall and movement; but nothing ...  
 the fall - she has no recollection of events

**Discussion**

Using Lingumatics I2E, we developed an NLP lexicon and algorithms to detect mentions of falls risk screening in clinical notes (mostly progress notes) automatically for patients who had no coded falls risk screening reported to PQRS through Epic MSSP. Our NLP algorithms sufficiently extracted and aggregated necessary quality outcomes for falls risk screening and achieved high performance as evaluated with our gold standard. Our study results confirmed that NLP can enhance current quality reporting approaches for falls risk screening. The potential value of NLP enhancements for quality regulatory reporting is extremely high for measures such as annual falls risk screening. Nevertheless, we strongly caution against the use of current NLP technology and algorithms as we have utilized here for official measure reporting. However, automated NLP assessments could be used periodically throughout a quality measure reporting period to help gauge individual and group provider performance and encourage needed discrete documentation. Moreover, some regulatory quality metrics do allow manual abstraction; for example, limited and official chart review could confirm a positive NLP result for falls risk screening within clinical notes that had not been coded. That might well result in significantly better and more accurate provider group performance on the regulatory measure.

Information about falls risk screening is commonly documented in clinical notes, and completely and accurately extracting mentions of falls risk screening from clinical narratives using NLP is an achievable although challenging goal. To the best of our knowledge, this is the first report of NLP-based extraction for quality reporting on falls risk screening. Other studies have focused on falls detection, with variable accuracy. For example, Toyabe reported a high accuracy (F-measure: 0.91) NLP approach to identify inpatient injurious falls from image order entries; he also noted that his NLP algorithm for identifying inpatient falls in progress notes only achieved an F-measure of 0.12 because of the high number of false-positive cases.<sup>23</sup> Elsewhere, Shiner et al. reported about an NLP algorithm to detect falls in progress notes with modest accuracy (specificity: 0.80; sensitivity: 0.44), as evaluated by manual review.<sup>24</sup> While these studies were designed to detect a “fall incident,” our NLP algorithms were designed to detect a “falls risk screening” event from clinical notes, an NLP task posing more challenges. Initially, we used terms related to “fall risk” and “fall screening” to probe a falls risk screening; however, in clinical practice, providers may not use these terms to document falls risk screening even when they have actually performed it. For example, providers may document screening with the following statements: “Patient has no fall in the past year” or “no fall.” Providers may also document clinical information relevant to falls as “there have been some falls, memory loss,” or “fall or dizzy when standing.” Prior types of NLP algorithms would not likely detect these cases nor do typical coding procedures. Therefore, we used the term “fall” as the major search term in order to mimic a common way that providers document a conversation about falls with patients. Before applying negations, a hit of “fall” is much more frequent than hit of “fall risk” (236 vs. 42 mentions in our corpus).

We applied negation detections to classify a mention of “fall” into either a positive case or a negative case of falls risk screening. At the same time, it was important to keep in mind that the term “fall” may have other linguistic meanings that would cause false positives for “falls risk screening” identification. For example, a description of a “fall event” (“fall on wall,” “fall with broken arm”); a season of the year (“last fall,” “fall 2014”); or “fell” as a typo for “feel.” We established a set of pre-negation and post-negation detection rules in order to eliminate false positives; however, these negations also introduced false negatives. For quality reporting, we expect a highly specific (avoiding false positive) NLP algorithm by sacrificing sensitivity. After iterative evaluation and refinement, our final I2E algorithm achieved 0.92 precision and 0.95 recall, which indicated that our NLP approach could effectively identify falls risk screening in clinical notes of patients who had no coded falls risk screening recorded in their EHR.

Our NLP approach accurately identified 59 patients without coded falls risk screening (out of 144) who had a falls risk screening documented in their clinical notes. These patients were previously misclassified as having no falls risk screening performed by their PCPs. This result indicates that quality reporting based on administrative coded data alone may not accurately measure the quality of care providers performed. After years of PQRS implementation, CMS noted that potential data quality issues exist among reporting systems, such as missing data elements and inaccurate or incomplete quality measure calculations. CMS also highlighted that completeness and accuracy of data are critical to the accurate calculation of a quality performance score.<sup>25</sup> Researchers have suggested multimodal approaches rather than relying on a single data resource to capture clinical events.<sup>26, 27</sup> Recently, quality reporting studies, including colonoscopy quality metrics, postoperative complications identification, and advanced adenoma detection, demonstrated that the NLP-based approach offers a powerful alternative to either unreliable administrative data or labor-intensive manual chart reviews.<sup>28-30</sup> In order to reduce non-reporting issues for measuring falls risk screening, one potential strategy is to incorporate NLP-extracted information on quality indicators that exist in the clinical notes as the supplemental data source for current quality reporting systems.

### **Limitations**

Four major limitations are identified for this study. First, the study population (negatives by coded data) only included 144 patients, and we estimate that relatively small number of patients could be identified as cases of falls risk screening; therefore, we did not split the clinical notes into a training and testing set. In order to obtain the NLP performance for the testing set, we plan to include 2014-2016 data and randomly sample a training and testing set to reevaluate our NLP algorithm. Second, we only used clinical notes from patients who had no administrative coded data recorded for falls risk screening to develop our lexicon and NLP algorithms. Indeed, providers may document a falls risk screening differently in clinical notes for patients who had coded falls risk screening. On the other hand, providers may possibly not document a falls risk screening in clinical notes at all because these screening activities are assumed to be reported through a coded form. In fact, MUSC also uses the Morse Fall Scale to assess falls risk for hospitalized patients; and this information is usually not documented in clinical notes. Therefore, our NLP approach would not reliably identify patients who had been assessed with the Morse Fall Scale. A further study focusing on NLP strategies identifying falls risk screening for cases confirmed by coded data may help refine the lexicon and NLP algorithms. Finally, this study was designed for the context of MUSC quality reporting purposes. Our lexicon and NLP algorithms may not be generalizable to other institutions without customization and evaluation. However, the concept of NLP-based approaches to enhance PRQS reporting may well be applicable to other healthcare settings.

### **Conclusions**

Information about falls risk screening can commonly be found in clinical notes of patients lacking such screening recorded by coding systems for quality reporting purposes. For PQRS quality reporting, we developed an NLP approach that automatically and effectively identified and extracted information about falls risk screening found in clinical notes. The results support the concept that using both structured coded data and clinical narratives for quality reporting is superior to the current reporting approach based on administrative coded data alone. Future work will focus on replicating the study findings using a larger sample and developing NLP-based identification strategies for other quality matrix. NLP augmentation of current regulatory quality measure assessment methodologies may result in improved care and more accurate quality reporting.

### **Acknowledgement**

This research was supported in part by an NIH/NCATS grant to the South Carolina Clinical and Translational Research Institute (SCTR) (5UL1TR000062-05). We also would like to thank Paul Sobanski and Ryan Siegle at Epic Systems Corporation.

## References

1. National Council on Aging. Falls Prevention Facts. Available at: <https://www.ncoa.org/news/resources-for-reporters/get-the-facts/falls-prevention-facts/>. Accessed on February 26, 2017.
2. Tinetti ME, Speechley M, Ginter SF. Risk factors for falls among elderly persons living in the community. *N Engl J Med*. 1988 Dec 29;319(26):1701-7.
3. Parkkari J, Kannus P, Palvanen M, Natri A, Vainio J, Aho H, Vuori I, Järvinen M. Majority of hip fractures occur as a result of a fall and impact on the greater trochanter of the femur: a prospective controlled hip fracture study with 206 consecutive patients. *Calcif Tissue Int*. 1999 Sep;65(3):183-7.
4. Sterling DA, O'Connor JA, Bonadies J. Geriatric falls: injury severity is high and disproportionate to mechanism. *J Trauma*. 2001 Jan;50(1):116-9.
5. Centers for Disease Control and Prevention. National Center for Health Statistics. Deaths from Unintentional Injury among Adults Aged 65 and Over: United States, 2000–2013. Available at: <https://www.cdc.gov/nchs/products/databriefs/db199.htm>. Accessed on February 26, 2017.
6. Burns EB, Stevens JA, Lee RL. The direct costs of fatal and non-fatal falls among older adults—United States. *J Safety Res* 2016:58.
7. Cameron ID, Gillespie LD, Robertson MC, Murray GR, Hill KD, Cumming RG, Kerse N. Interventions for preventing falls in older people in care facilities and hospitals. *Cochrane Database Syst Rev*. 2012 Dec 12;12:CD005465.
8. Moyer VA; U.S. Preventive Services Task Force. Prevention of falls in community-dwelling older adults: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2012 Aug 7;157(3):197-204.
9. Stevens JA, Ballesteros MF, Mack KA, Rudd RA, DeCaro E, Adler G. Gender differences in seeking care for falls in the aged Medicare Population. *Am J Prev Med* 2012;43:59–62.
10. American Geriatrics Society, British Geriatrics Society, and American Academy of Orthopaedic Surgeons Panel on Falls Prevention.. Guideline for the prevention of falls in older persons. *J Am Geriatr Soc*. 2001 May;49(5):664-72.
11. Centers for the Medicare and Medicaid Services. CMS Measures Inventory. Available at: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/CMS-Measures-Inventory.html>. Accessed on March 1st, 2017.
12. Centers for Medicare and Medicaid Services. Physician Quality Reporting System. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html>. Accessed on March 1st, 2017.
13. Hirsch JA, Rosenkrantz AB, Ansari SA, Manchikanti L, Nicola GN. MACRA 2.0: are you ready for MIPS? *J Neurointerv Surg*. 2016 Nov 24. pii: neurintsurg-2016-012845.
14. Centers for Medicare and Medicaid services. 2016 Cross-Cutting Measure Set. Available at: [https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2016\\_PQRS-Crosscutting.pdf](https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2016_PQRS-Crosscutting.pdf). Accessed on March 1st, 2017.
15. Liao KP, Cai T, Savova GK, Murphy SN, Karlson EW, Ananthakrishnan AN, Gainer VS, Shaw SY, Xia Z, Szolovits P, Churchill S, Kohane I. Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ*. 2015 Apr 24;350:h1885.
16. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc*. 2005 Jul-Aug;12(4):448-57.
17. Friedman C, Hripcsak G. Natural language processing and its future in medicine. *Acad Med*. 1999 Aug;74(8):890-5.
18. Pons E, Braun LM, Hunink MG, Kors JA. Natural Language Processing in Radiology: A Systematic Review. *Radiology*. 2016 May;279(2):329-43.
19. Demner-Fushman D, Chapman WW, McDonald CJ. What can natural language processing do for clinical decision support? *J Biomed Inform*. 2009 Oct;42(5):760-72.
20. Unified Medical Language System (UMLS). Available at <http://medical-dictionary.thefreedictionary.com/UMLS+Metathesaurus>. Accessed on March 6, 2017.
21. Agency for healthcare research and quality. United States Health Information Knowledgebase Risk Category Assessment: Falls Screening. Available at: <https://ushik.ahrq.gov/ViewItemDetails?&system=mu&itemKey=158801000&enableAsynchronousLoading=true>. Accessed on March 1st, 2017.
22. Stanford NLP group. Evaluation of ranked retrieval results. <http://nlp.stanford.edu/IR-book/html/htmledition/evaluation-of-ranked-retrieval-results-1.html>. Accessed on June 2, 2016.

23. Toyabe S. Detecting inpatient falls by using natural language processing of electronic medical records. *BMC Health Serv Res*. 2012 Dec 5;12:448.
24. Shiner B, Neily J, Mills PD, Watts BV. Identification of Inpatient Falls Using Automated Review of Text-Based Medical Records. *J Patient Saf*. 2016 Jun 22. [Epub ahead of print].
25. Centers for Medicare & Medicaid Services . Request for Information Regarding Implementation of the Merit-Based Incentive Payment System, Promotion of Alternative Payment Models, and Incentive Payments for Participation in Eligible Alternative Payment Models. *Federal Register* Vol. 80, No. 190; 59102-59103.
26. Hill AM, Hoffmann T, Hill K, Oliver D, Beer C, McPhail S, Brauer S, Haines TP. Measuring falls events in acute hospitals-a comparison of three reporting methods to identify missing data in the hospital reporting system. *J Am Geriatr Soc*. 2010 Jul;58(7):1347-52.
27. Olsen S, Neale G, Schwab K, Psaila B, Patel T, Chapman EJ, Vincent C. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. *Qual Saf Health Care*. 2007 Feb;16(1):40-4.
28. Murff H, , FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, Dittus RS, Rosen AK, Elkin PL, Brown SH, Speroff T. Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA*. 2011 Aug 24;306(8):848-55.
29. Raju GS, Lum PJ, Slack RS, Thirumurthi S, Lynch PM, Miller E, Weston BR, Davila ML, Bhutani MS, Shafi MA, Bresalier RS, Dekovich AA, Lee JH, Guha S, Pande M, Blechacz B, Rashid A, Routbort M, Shuttlesworth G, Mishra L, Stroehlein JR, Ross WA. Natural language processing as an alternative to manual reporting of colonoscopy quality metrics. *Gastrointest Endosc*. 2015 Sep;82(3):512-9.
30. Gawron AJ, Thompson WK, Keswani RN, Rasmussen LV, Kho AN. Anatomic and advanced adenoma detection rates as quality metrics determined via natural language processing. *Am J Gastroenterol*. 2014 Dec;109(12):1844-9.