BACKGROUND

• Cognitive impairment is under-detected during hospitalization
• The electronic health record (EHR) contains detailed clinical information including structured (e.g. ICD codes) and unstructured (e.g. narrative text) data that may improve methods for detection and diagnosis
• High-risk and underrepresented groups are more likely to experience hospitalization but less likely to receive diagnoses, limiting the utility of structured data (i.e. ICD codes, medications) in EHRs
• Understanding the availability and characteristics of unstructured data describing cognition is necessary to advance text mining applications capable of improving detection of potential cognitive impairment

OBJECTIVE

• Identify and characterize unstructured data reflecting symptoms of cognitive impairment in the acute-care EHRs of patients with dementia

METHODS

• Retrospective cohort study, medical record review
• N=343 non-hospice Medicare beneficiaries with primary diagnosis of hip fracture or stroke and documented dementia as identified by Medicare claims techniques demonstrating good sensitivity (85%) and specificity (89%)
• Clinician reviewers identified and classified unstructured data using standardized criteria; narrative text was evaluated for key terminology
• Blinded, random reliability assessment of 10% of EHRs; Cohen’s Kappa Coefficient 0.82 (0.02 SD; 95% CI 0.78-0.87)

RESULTS

• Findings suggest that inpatient clinicians use specific terminology to describe symptoms of cognitive impairment
• Relevant terminology was prevalent in acute care EHRs of dementia patients and predominantly described symptoms in general/vague terms such as “poor mental status.” Most terminology was not consistent with gold-standard clinical assessment or diagnostic criteria, which may reflect lack of clinician training

NEXT STEPS

• Findings provide a preliminary understanding of the potential for unstructured data to complement existing automated EHR screening approaches
• Future work will focus on validation of an EHR-based phenotype model incorporating structured and unstructured data elements within a well-characterized ADRD cohort