We sincerely thank the reviewers for their feedback. We have developed a deep learning NLP pipeline to detect signs of cognitive impairment from unstructured clinician notes in electronic health records (EHR). The reviewers noted several key strengths including the relevance to healthcare (as dementia is often underdiagnosed, and EHR may not accurately capture the diagnosis as a structured diagnosis code [ICD code]), as well as the novelty of our approach. The many methodological innovations as described by Reviewer C are: (1) combination of rule-based filters (regex on keywords to identify candidate sentences) and NLP methods applied on sentences to further improve model performance; (2) adjudication at sentence-level; (3) a novel dataset with manual annotation by clinician experts; and (4) use of genomic data and subgroup analysis based on APOE alleles. There were also several criticisms, which we address below.

Bias in study data: Reviewers A and B raised concerns that constructing our study sample using the 18 pre-selected keywords may lead to biases in the study. To clarify, our 18 keywords were based on careful literature review of established methods [1-3] for identifying patients with dementia using EHR. We collaborated with our memory-division physicians and ensured that these keywords comprehensively capture evidence of cognitive impairment, and that it would be exceedingly rare to describe cognitive impairment (or the lack thereof) in EHR notes without using one of these keywords. In our revised camera-ready manuscript, we will elaborate on the source of these keywords and include the references mentioned above.
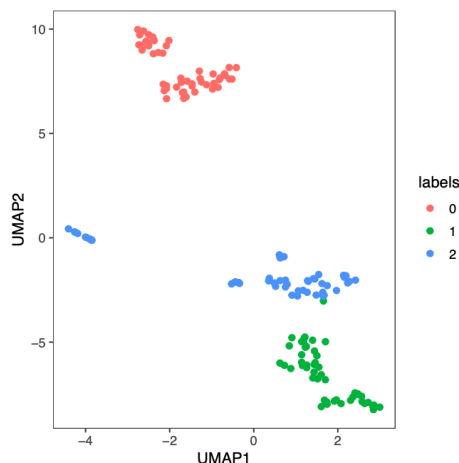

Figure 1

Generalization of the sequence-level model: Reviewer A noted thar our annotations were primarily derived from regular expressions (always-patterns). We thank the reviewer for their insightful comment. Per their suggestion, we tested the model's generalizability using a small dataset of manually annotated sequences (n=150) which did not match an always-pattern. The ClinicalBERT embeddings were able to able accurately discriminate between all three classes (0: no cognitive impairment; 1: neither; 2: cognitive impairment); see Figure 1.

Patient level annotation: Reviewers A, B, and C all noted the lack of patient-level gold-standard labels. We agree that this is one of the main limitations of our study. Per the suggestion of Reviewer B, we performed several manual chart reviews of patients who were classified as cognitively impaired from our ClinicalBERT model but did not have a dementia-related ICD code or medication. We identified many subtle examples where a non-specialist care provider notes cognitive impairment. For example, a physical therapist note states, "XX y.o. male who presents to PT with primary impairments in balance, cognition," and a geriatrician reports, "Risk of delirium is moderate due to age, acute care setting, and underlying dementia." Although our next step is to construct a carefully adjudicated patient-level annotated dataset, these two illustrative excerpts suggest that the ClinicalBERT model may indeed correctly classify patients with cognitive impairment who do not have structured data in EHR. In our updated manuscript, we will provide metrics on a set of patients (n=300) with clinically adjudicated labels.

Other issues: Reviewer B noted presentation issues with Table 1 and Figure 1. We thank the reviewer and will addresses these issues in the revised manuscript. We will also add details on the TF-IDF and ClinicalBERT setup, as suggested by Reviewer D, and will elaborate on future plans to evaluate our models at other institutions and healthcare settings.

REFERENCES

1. Gilmore-Bykovskyi, A.L., et al., Unstructured clinical documentation reflecting cognitive and behavioral dysfunction: toward an EHR-based phenotype for cognitive impairment. J Am Med Inform Assoc, 2018. 25(9): p. 1206-1212.

2. Reuben, D.B., et al., An Automated Approach to Identifying Patients with Dementia Using Electronic Medical Records. J Am Geriatr Soc, 2017. 65(3): p. 658-659.

3. Amra, S., et al., Derivation and validation of the automated search algorithms to identify cognitive impairment and dementia in electronic health records. J Crit Care, 2017. 37: p. 202-205.