**INTRODUCTION**

- Eligibility criteria in clinical trials play a pivotal role in patient recruitment and in safety and treatment evaluation, resulting in improved patient care.
- Natural language processing (NLP) techniques offer the potential to enhance the efficiency of clinical trial studies by automatically extracting eligibility criteria. However, NLP approaches face challenges in capturing fine-grained criteria within a given text and may lack applicability across various disease areas.
- Recently, large language models such as ChatGPT and GPT-4 have gained traction in the open NLP domain as well as biomedical and clinical domains.²,³

**OBJECTIVE**

Our aim is to develop a system that automatically extracts eligibility criteria, emphasizes contextual attributes, and can handle diverse diseases utilizing a cutting-edge large language model.

**AUTO-CRITERIA: SYSTEM OVERVIEW AND COMPONENTS**

1. **System Overview**
   - Pre-processing
     - Split the raw criteria text for each trial document into two parts: Inclusion and Exclusion, then split each of these parts into smaller chunks and run AutoCriteria on each chunk.
   - Knowledge ingestion
     - Identified the ontology of key criteria entities and attributes for each disease with the help of knowledge experts. This knowledge is also leveraged in prompt modeling.
   - Prompt modeling
     - Created two separate comprehensive prompts, one for Inclusion and another for Exclusion.
     - Each prompt consists of three main components: 1) a general instruction, 2) the Inclusion or Exclusion criteria text, and 3) the query that asks about criteria attributes (phrases), their corresponding values, modifier information, entity types, and source sentences.
   - Post-processing
     - Processed the GPT-4 responses to handle any inconsistencies in the model output.
   - Interim evaluation
     - Evaluated the prompts manually, and iteratively calibrated them using expert feedback for every disease.

2. **Prompt Modeling**

**DISCUSSION**

- AutoCriteria system not only identifies key criteria but also extracts contextual details including negation/temporal information and classifies entity types.
- Our system is highly adaptable to various diseases, including Cancers, Autoimmune diseases, Alzheimer’s, NAS, and other rare diseases, highlighting its generalizability.

**CONCLUSION**

- We have developed a generalizable GPT-based system that can identify granular eligibility criteria information from clinical trial documents across a variety of disease domains without requiring manual annotations.
- Our adaptable prompts allow easy customization for new diseases without manual training data, facilitating large-scale criteria analysis.
- With enhanced granularity, improved accuracy and generalizability, AutoCriteria has the potential to significantly streamline the clinical trial initiation and conduct process, ultimately reducing time requirements.

**REFERENCES**


**CONTACT**

Please contact the first authors via email: Surabhi Datta: sdatta@imohealth.com, Kyeryoung Lee: klee@imohealth.com