DOI: 10.1002/alz.041034

BIOMARKERS

POSTER PRESENTATIONS

Alzheimer's & Dementia®
THE JOURNAL OF THE ALZHEIMER'S ASSOCIATION

## Biomarkers (non-neuroimaging) / Differential diagnosis

# Scalable diagnostic screening of mild cognitive impairment using AI dialogue agent

**Fengyi Tang**[1] | Ikechukwu Uchendu[1] | Fei Wang[2] | Hiroko H Dodge[3] | Jiayu Zhou[1]

[1] Michigan State University, East Lansing, MI, USA

[2] Weill Cornell Medical School, New York, NY, USA

[3] Oregon Center for Aging & Technology (ORCATECH), Portland, OR, USA

**Correspondence**
Fengyi Tang, Michigan State University, East Lansing, MI, USA.
Email: tangfeng@msu.edu

## Abstract

**Background:** The search for early biomarkers of mild cognitive impairment (MCI) has been central to Alzheimer's Disease (AD) and the dementia research community in recent years. While there exist in-vivo biomarkers (e.g., beta-amyloid and tau) that can serve as indicators of pathological progression toward AD, biomarker screenings are prohibitively expensive to scale if widely used among pre-symptomatic individuals in the outpatient setting. Behavior and social markers such as language, speech, and conversational behaviors reflect cognitive changes that may precede physical changes and offer a much more cost-effective option for preclinical MCI detection, especially if they can be extracted from a non-clinical setting.

**Method:** We developed a prototype AI conversational agent that conducts screening conversations with participants. Specifically, this AI agent must learn to ask the right sequence of questions to distinguishing the conversational characteristics of the participants with MCI from those with normal cognition. Using transcribed data obtained from recorded conversational interactions between participants and trained interviewers generated in a recently completed clinical trial, and applying supervised learning models to these data, we developed a novel reinforcement learning (RL) pipeline and a dialogue simulation environment to train an efficient dialogue agent to explore a range of semi-structured questions. We train and validate our AI dialogue agent based on transcribed data from a randomized controlled behavioral intervention study, where we use the transcribed data from 41 subjects (14 MCI, 27 NL). Each subject has an average of 35 turns of dialogue on average.

**Result:** The results show that while using only a few turns of conversation, our framework can significantly outperform state-of-the-art supervised learning approaches used in a past study. An AI agent of 30 turns of dialogue achieves over 0.853 Area Under the Receiver Operating Characteristic Curves (AUC) and 0.809 AUC with 20 turns, as compared to 0.811 AUC with the full dialogue turns.

**Conclusion:** Our dialogue-based AI agent presents a step toward using AI to extend clinical care beyond the classical hospital and clinical settings, where we find that AI-generated dialogues produce more predictive linguistic markers.

 © 2020 the Alzheimer's Association 

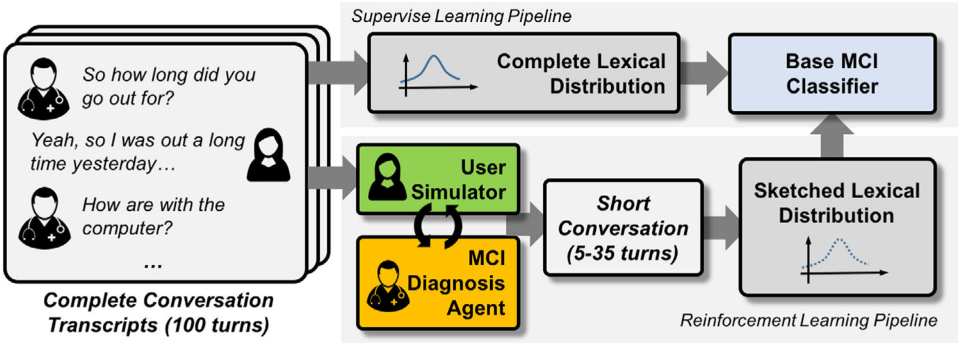


**FIGURE 1**

**TABLE 1**

**Table 1. Classification of MCI based on complete transcript vs. simulated conversations**

| *Model* | *AUC* | *F1-Score* | *Sensitivity* | *Specificity* |
|---|---|---|---|---|
| *SVM w/ LIWC* | *0.712 (0.612–0.811)* | *0.631 (0.500 – 0.761)* | *0.680 (0.476 – 0.886)* | *0.744 (0.563 – 0.922)* |
| *Supervised DL w/ LIWC* | *0.689 (0.560–0.818)* | *0.182 (0.055 – 0.370)* | *0.300 (0.010 – 0.758)* | *0.767 (0.364 – 0.970)* |
| *SVM w/ SKP* | *0.797 (0.719–0.879)* | *0.719 (0.591 – 0.846)* | *0.654 (0.473 – 0.835)* | *0.939 (0.855 – 1.0)* |
| *Supervised DL w/ SKP* | *0.811 (0.715–0.907)* | *0.642 (0.469 – 0.813)* | *0.600 (0.366 – 0.833)* | *0.911 (0.838 – 0.984)* |
| *RL (T=5)* | *0.633 (0.535–0.703)* | *0.486 (0.288 – 0.680)* | *0.459 (0.280 – 0.630)* | *0.811 (0.661 – 0.936)* |
| *RL (T=10)* | *0.741 (0.631–0.852)* | *0.590 (0.352 – 0.829)* | *0.560 (0.309 – 0.811)* | *0.922 (0.823 – 0.969)* |
| *RL (T=15)* | *0.721 (0.618–0.827)* | *0.595 (0.399 – 0.790)* | *0.50 (0.327 – 0.713)* | *0.922 (0.856 – 0.987)* |
| *RL (T=20)* | *0.809 (0.706–0.914)* | *0.726 (0.551 – 0.901)* | *0.620 (0.413 – 0.827)* | *0.988 (0.953 – 1.0)* |
| *RL (T=30)* | *0.853 (0.796–0.914)* | *0.801 (0.733 – 0.880)* | *0.818 (0.678 – 0.958)* | *0.898 (0.828 – 0.969)* |
| *RL(T=35)* | *0.859 (0.787–0.952)* | *0.808 (0.735 – 0.883)* | *0.818 (0.677 – 0.958)* | *0.911 (0.839 – 1.0)* |
| *Difference* | *0.0616 (-0.049–0.172)* | *0.089 (-0.078 – 0.259)* | *0.163 (-0.083 – 0.410)* | *-0.040 (-0.130 – 0.050)* |