**Title:** AI-based Voice biomarkers to detect Mild Cognitive Impairment and Alzheimer's Disease biomarker positivity.

**Authors:** Ihab Hajjar, Brandon Henley, Zhiyi Yang, Merve Apalak. Ramcharan Chandrashekar, Malik Obideen, Felicia Goldstein, Albert Montillo

**Introduction:** Alzheimer's disease is underdiagnosed by ~40%, in part due to the time intensive cognitive tests that can take several hours to conduct and require interpretation by a skilled clinician. A non-invasive, fast and efficient screening tool, which identifies subjects at high risk for MCI and a candidate for more extensive testing, would better utilize healthcare resources. Brain biomarkers, such as the CSF phosphorylated tau to amyloid-beta 42 ratio (pTAR), are one such route to ascertain the risk of MCI. Such biomarkers enjoy improved objectivity, but require an invasive lumbar puncture. Accumulating evidence suggests that audio recordings can predict cognitive performance; however, there is comparatively much less evidence that audio recordings can predict biomarker status, which would enable efficient patient screening without lumbar puncture invasiveness.

**Methods:** We analyzed data from the Brain Stress Hypertension and Aging Program (B-SHARP: n=487, mean age=65.1, women=60%, African American=44%, 62% MCI positive). All participants underwent an identical structured voice recording protocol including 1-minute free speech of responding to the study personnel conversation (FS), a 1-minute fluency task (VF), and a 2-minute picture description (PD). The audio processing pipeline steps include audio normalization, transcription using Automatic Speech Recognition, audio wave representation, Natural Language Processing and Understanding (NLP/NLU), feature engineering and selection. Both acoustic and linguistic features obtained were then introduced into supervised classification algorithms using a feedforward neural network (FNN) with stratified K-fold cross-validation (CV) to predict CSF pTAR positivity status by dichotomizing the tau-mediated neuronal damage to amyloid plaque accumulation ratio. AI model performance metrics included model Area Under the Curve (AUC), sensitivity, specificity, and model accuracy and recall. Each task was evaluated separately.

**Results:** Clinical performance based on AUC from each of the three tasks ranged from 80.1% to 75.5% for AD positivity across held out test data not used to train the model. The FS model achieved the highest test AUC (80.1%), followed by VF (79.8%) and PD (75.5%). Sensitivity and specificity varied across tasks, highlighting differences in discriminative power between modalities.

**Conclusion:** Our approach incorporates state-of-the art AI approaches and allows for rapid patient screening through the prediction of pTAR biomarker positivity with significant accuracy using a ~5-minute protocol. Further development of these approaches may provide additional accuracy improvements and precision for disease progression tracking and enable patient screening to reduce the under diagnosis of AD.