Speech-Based Biomarkers for Evaluating Tau PET Status in Preclinical Alzheimer's Disease: Insights from the RETAIN Study

Federico Parisi, PhD[1]; Dzmitry A. Kaliukhovich, PhD[1]; Macarena Garcia Valdecasas Colell, PhD[1]; Charmaine Demanuele, PhD[1]; Nicklas Linz, PhD[2]; Johannes Tröger, PhD[2]; Janna Herrmann, MS[2]; Stephanie Slania, PhD[1], Ziad S. Saad, PhD[1]; Gayle Wittenberg, PhD[1]; Arthur Simen, MD, PhD[1]; Fiona Elwood, PhD[1]; Tricia Thornton-Wells, PhD[1]; Lennert Steukers, DVM, PhD[1]; Anahita Kyani, PhD[1]

1. Johnson & Johnson Innovative Medicine, 2. ki:elements GmbH

Background

RETAIN (NCT06544616) is a Phase 2b trial evaluating the efficacy, safety, and immunogenicity of a phosphorylated tau-targeted active immunotherapy (JNJ-64042056), developed with AC Immune SA, in **preclinical/asymptomatic Alzheimer's disease** (AD). The study's pre-screening funnel employs an exploratory speech-based digital assessment to assess subtle cognitive impairment (ki:elements GmbH) and a blood-based p217+tau assay (Johnson & Johnson; Quanterix platform, LucentAD) to enrich for increased probability of meeting tau PET eligibility. During screening, individuals complete in-clinic cognitive assessments, MRI, and tau PET imaging to confirm that individuals were cognitively unimpaired and had elevated brain tau levels. Although tau PET imaging is the *in vivo* gold standard for assessing tau pathology in AD, it entails high patient burden and cost.

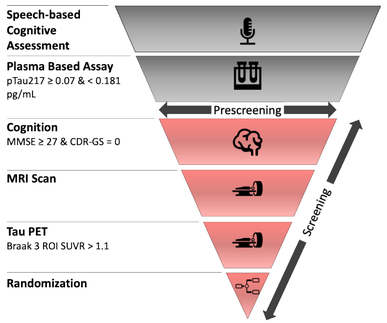

Figure 1. RETAIN (pre-)screening funnel. Speech-based cognitive assessment was exploratory and not used as eligibility criteria.

Objective

Evaluate whether speech-derived metrics are predictive of tau PET positivity, in pTau217+ cognitively unimpaired individuals.

Methods

- Analyzed cohort profile:
  - **Cognitively Unimpaired**: CDR global score of 0 and MMSE ≥ 27;
  - **Plasma pTau217 positive**: 0.07 ≤ pTau217 ≤ 0.181 pg/mL;
- All participants completed a speech-based cognitive assessment using the Mili platform developed by ki:elements in the pre-screening period (Figure 2).
- The speech-based cognitive assessment yielded 114 features designed to characterize the cognitive performance of each individual.
- Demographic variables, including sex, age and education were used as covariates.
- **Tau PET positivity** was defined as: Tau PET Braak 3 ROI SUVR > 1.1.
- The approach used to train the binary classification model using the speech-derived features and to evaluate its performance in predicting tau PET positivity is shown in Figure 2. A baseline model using only demographic variables was also trained and evaluated using the same data splits.

Results

- Preliminary data (September 2025) included a total of 136 US-based, English-speaking asymptomatic pTau217-positive participants (Table 1).
- Classification performance on the training sets using a compact set of speech-derived features (4–10 features) achieved a mean ROC AUC above 73%, with peak performance of 73.9% at K = 9 (Figure 3A).
- Similar performance was observed on the test sets, with models using 4 and 9 speech-derived features achieving mean ROC AUCs of 74.1% (SD = 6.9%) and 74.8% (SD = 6.8%), respectively, and outperforming the baseline model (mean ROC AUC = 57.8%, SD = 9.3%) by approximately 17% (Figure 3B).

Table 1. Participant characteristics.

Characteristic	Value
Female	101 (74.3%)
Age (Mean ± SD)	69.5 ± 16.5
Education (Mean ± SD)	15.2 ± 2.6

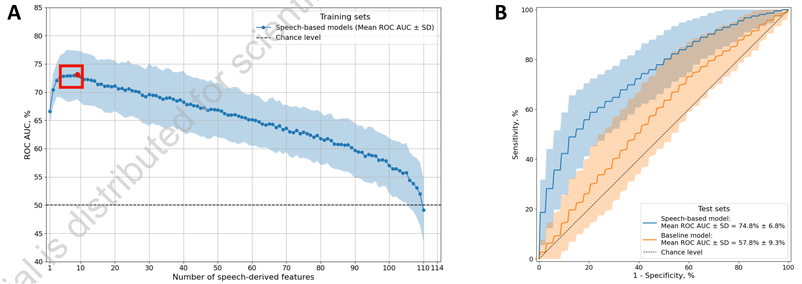

Figure 3. (A) Mean ROC AUC±SD as a function of the number of speech-derived features obtained on the 100 training sets. (B) Mean ROC AUC±SD for the baseline and best (K = 9) speech-based models obtained on the 100 test sets.

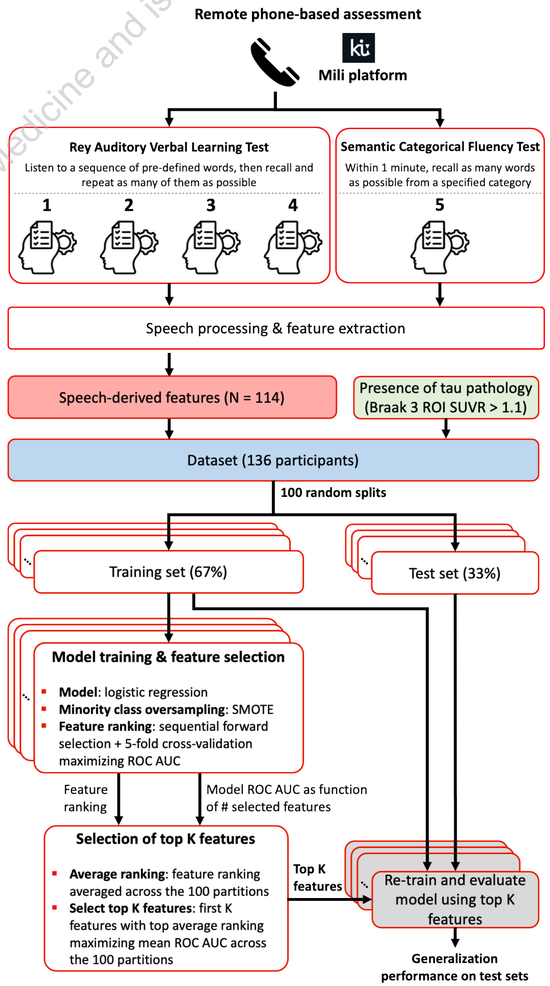

Figure 2. Remote cognitive phone-based assessment and model training/evaluation pipeline.

Conclusions

Key takeaway:

Our findings indicate that a small set (4–10) of speech-derived metrics can predict tau PET positivity in pTau217+ cognitively unimpaired individuals, achieving ROC AUCs above 74%, compared to 58% using demographics alone.

Utility:

The scalability, minimal burden, and cost efficiency of speech-derived metrics enable their seamless integration into multimodal screening strategies, with the potential to reduce recruitment duration and lower trial costs.

Limitations:

Further validation is warranted given the small sample size and class imbalance. Furthermore, the prescreening and screening criteria limit the validity of these findings to the specific characteristics of the analyzed population.

References

1. Tröger, J., Schwed, L.,. (2022). Validation of the Remote Automated ki:e Speech Biomarker for Cognition in Mild Cognitive Impairment: Verification and Validation following DiME V3 Framework. Digital Biomarkers, 6(3), 107–116.
2. Wilson, D., Copeland, K (2025). Alzheimer's Pathology Detection Made Simple: A Blood Test for p-Tau 217 to Aid in the Diagnostic Evaluation for Alzheimer's Disease. Lucent Diagnostics, 1-5.