

Robustness and generalizability of a speechbased composite score for measuring disease progression in AD

Michael Spilka¹, Mengdan Xu¹, Jessica Robin¹, & William Simpson^{1,2}

¹Winterlight Labs (a division of Cambridge Cognition), Toronto, ON, Canada; ²Department of Psychiatry & Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada

Background

- Natural language processing (NLP) and computational linguistics can help objectively quantify language impairment and provide novel measures of disease progression in AD.
- We previously developed a novel speech-based composite score by analyzing patient verbal responses from an excerpt of the Clinical Dementia Rating (CDR) interview (mean audio duration of 3.5 min).
- This composite score was sensitive to longitudinal language change and significantly associated with clinical endpoints in a sample of AD participants [1].

Results

 The speech-based composite showed significant longitudinal change over time (B = 1.78, t = 4.09, p < .001), which paralleled longitudinal trajectories of both the ADAS-Cog-11 and object content scores (Fig. 1).

CAMBRIDGE

COGNITION

WINTERLIGHT

Contact: bill@winterlightlabs.com

- The magnitude of change from baseline to endpoint was similar for the composite (d = 0.37) and object content score (d = 0.46) but smaller than for the ADAS-Cog-11 (d = 0.81) (Fig. 2).
- The composite score had moderate correlations with the ADAS-Cog-II (r = 0.48, p < .001) and object content score (r = -0.41, p < .001). ADAS-Cog change score
- When compared to the original development study (composite score obtained from CDR interview speech), the current composite (from the Picture Description task) had a 50% smaller magnitude of longitudinal change (Fig. 2).
- The current analysis examined the generalizability of this composite to speech obtained from a brief *Picture Description task* in a separate sample of AD participants.

Methods

Participants:

- Current sample (Picture Description speec English-speaking individuals with mild-to-AD who were randomized into the placebo phase 2/3 clinical trial.
- Reference sample (CDR speech): 101 Englis speaking individuals with mild cognitive in reatment and placebo arms).
 Reference sample (CDR speech): 101 Englis speaking individuals with mild cognitive in reatment and placebo arms).

Assessments (completed at baseline, and at $\frac{v}{g}$ 6-month, and 12-month follow-up):

- Clinical: ADAS-Cog-11
- Speech: two app-based Picture Description tasks (1.5 min mean audio duration per task)

- 5/9 component features had similar trajectories in both samples (Fig. 3).
- Time Adding the object content score to the composite led to greater change from baseline but this was comparable to the magnitude of change for the object content score alone (Fig. 2).

Figure 1: Composite score mean change from baseline at 6- and 12-months



<u>Quantitative speech features</u> (extracted for each participant from transcribed speech recordings using signal processing and NLP tools):

- Speech-based composite score: 9 speech and language features
- Picture Description performance: object content score [2]





<u>Analyses:</u> The longitudinal trajectory and magnitude of change for the composite score were compared to those of the clinical study endpoints (e.g., ADAS-Cog) and the object content score.

Figure 3: Component features mean change from baseline in each sample



Conclusions

Our previously published, novel speech composite score was

Results were compared against our previously published performance of the composite derived from CDR interview speech [1].

We further examined associations among speechbased scores and clinical endpoints, and whether combining speech-based scores increased sensitivity.

Participant characteristics

	Current sample	Reference sample
Baseline N	148	101
Age – mean (SD)	70.0 (6.6)	69.3 (7.0)
Sex – n (%)		
Female	83 (56.1)	58 (57.4)
Male	65 (43.9)	43 (42.6)
ADAS-Cog-11 – mean (SD)	24.2 (8.1)	17.9 (5.7)

broadly generalizable to a shorter, more structured speech task.

- The composite remained sensitive to longitudinal change in early AD was and associated with clinical endpoints.
- Analysis of individual components showed similar trajectories for 5/9 features between datasets, highlighting the broader importance of these features for tracking progression in AD.
- Results further validate digital speech assessments as efficient, low-burden, and patient-relevant measures, which are sensitive to disease progression in AD.

References

- I. Robin J., Xu M., Balagopalan A., Novikova J., Kahn L., Oday A., Hejrati M., Hashemifar S., Negahdar M., Simpson W., & Teng E. (2023). Automated detection of progressive speech changes in early Alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring.
- 2. Robin J., Xu M., Detke, M., & Simpson W. (2022). Validation of an objective, speech-based object content score for measuring disease progression in AD. *Clinical Trials on Alzheimer's Disease Conference*.