Investigating Conversational Speech Latency As a Digital Biomarker of Schizophrenia: A Comparison of Manual and Automated Approaches

Submitter Hardik Kothare

Affiliation Modality.AI, Inc.

SUBMISSION DETAILS

I agree to provide poster pdf for attendee download. Yes

I have used the poster abstract template to develop my abstract. Yes

Methodological Issue Being Addressed We investigate two methodological issues regarding conversational speech latency, defined as the time duration between prompt completion and patient response. First, whether speech latency discriminates between patients with schizophrenia and healthy controls in picture description tasks administered remotely via a multimodal dialog platform. Second, we assess whether automated machine learning methods can reliably quantify this metric to enable scalable implementation.

Introduction We hypothesized that speech latency – a potential negative symptom of schizophrenia – would differentiate patients from controls as a remotely-assessable digital biomarker. We investigated whether automated computation of the measure using OpenAI's Whisper speech recognition software could still discriminate between cohorts while replacing time-intensive manual annotation.

Methods We analyzed 402 picture description sessions from 206 participants (117 healthy controls, 89 patients with schizophrenia) collected via the Modality platform using two methods. First, we manually annotated speech latency as the interval between prompt completion and speech onset. We then evaluated how well it discriminated between schizophrenia and healthy controls using Mann-Whitney U tests, estimating effect sizes using Glass's Delta. Second, we also developed an automated speech latency measure, using OpenAI's Whisper software to transcribe audio, detect prompt endings, identify speech onset, and assess analytical and clinical validity at cohort discrimination vis-a-vis manual annotations.

Results Patients exhibited prolonged speech latency (hand-annotated) vis-a-vis controls (3.68s vs. 2.92s; Mann-Whitney $p<0.01$, Glass's Delta $=0.55$), establishing speech latency as a digital biomarker. Automated computation maintained discriminative power (5.48s vs. 3.67s; $p<0.000001$, Glass's Delta=0.78) with moderate accuracy against this manual ground truth (MAE=2.08s, Spearman $\rho=0.47$).

Conclusion Automated speech latency significantly distinguishes schizophrenia patients from controls in remote virtual conversational assessments . These findings enable integration of speech latency into digital platforms for schizophrenia monitoring.

Co-Authors

Cathy Zhang[1], **Hardik Kothare[1],** Michael Neumann[1], Beverly Insel[2], Anzalee Khan[2], Jean-Pierre Lindenmayer[2], Vikram Ramanarayanan[1]

[1] Modality.AI, Inc.

[2] Nathan Kline Institute

Keywords

Keywords
schizophrenia
speech latency
digital biomarker
multimodal dialog platform

Guidelines I have read and understand the Poster Guidelines

Disclosures Authors HK, MN and VR are full-time salaried employees of Modality.AI, Inc. Authors CZ, HK, MN and VR may hold ownership interest in Modality.AI, Inc.