Characterizing Depressive Symptom Fluctuations using Automated Speech Assessment: A Remote Feasibility Study

Danielle DeSouza[1], Melisa Gumus[1,2], Celia Fidalgo[1], William Simpson[1,3], Jessica Robin[1]

(1) Winterlight Labs, Toronto, ON, Canada
(2) Department of Psychology, University of Toronto, Toronto, ON, Canada
(3) Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada

Background

Mood symptoms are highly heterogeneous between individuals and are known to fluctuate over time. Previous research has demonstrated that individuals with major depressive disorder can be separated into distinct classes based on how their symptoms fluctuate [1]. The degree to which mood symptoms fluctuate has also been linked to global functioning and quality of life measures, with high fluctuating individuals having significantly poorer outcomes [2]. One barrier to understanding fluctuating symptoms is that repeated clinical assessments are not always feasible and can be burdensome for patients to complete. There has been a growing interest in the use of automated speech and language assessment to objectively characterize mood symptoms in psychiatric disorders [3]. Speech assessments can be completed remotely, at high-frequency, and are low patient burden. The goal of the current study was to determine the feasibility of collecting daily speech and mood data and to characterize daily speech fluctuations relative to depression symptoms over a one month period.

Methods

Daily assessments of speech and mood were completed remotely for 30 days. Audio recordings in response to an open-ended prompt asking how participants were feeling that day (Figure 1A) were obtained using a speech app and analyzed using signal and natural language processing (NLP) to derive >500 acoustic and linguistic measures (Figure 1B). Daily depression symptoms were separately collected using a 9-item electronic questionnaire based on items from the Patient Health Questionnaire-9, with total daily scores ranging from 0 (no symptoms) to 55 (severe symptoms).

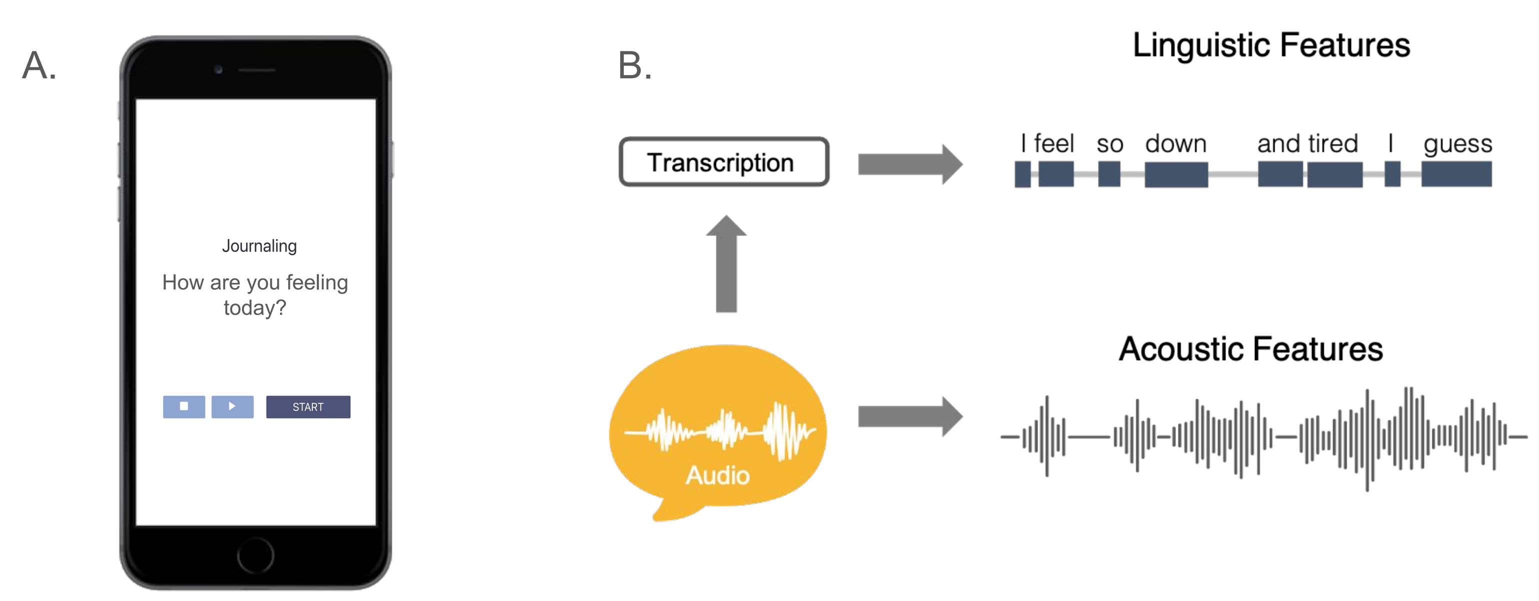

A.
Journaling
How are you feeling today?
START
B.
Linguistic Features
Transcription
I feel so down and tired I guess
Acoustic Features
Audio

Figure 1: A. Screenshot of the open-ended question prompt as displayed in the mobile speech app. **B.** Schematic representation showing how audio samples are used to derive acoustic and linguistic speech features [4]. For linguistic measures, samples are transcribed and NLP methods are applied to assess lexical, syntactic, and content measures. Audio waveforms can be directly analyzed to capture acoustic aspects of speech such as speech rate, pause time, and pitch.

Within-subject variations in daily depression were characterized by plotting mean residual scores (Figure 2). Spearman correlations (r_S) between speech and depression measures were assessed using data pooled across 26 time points (short day assessments) (Figure 3), and repeated measures correlations investigated within-individual associations (Figure 4).

Results

Sixteen individuals (5 M, 11 F, mean age ± SD= 30.8 ± 9.3, mean daily depression score ± SD= 13.1 ± 9.4) participated in the study. Participants provided ~380 audio samples and 400 depression scores (79% and 83% of scheduled assessments, respectively), indicating high study adherence.

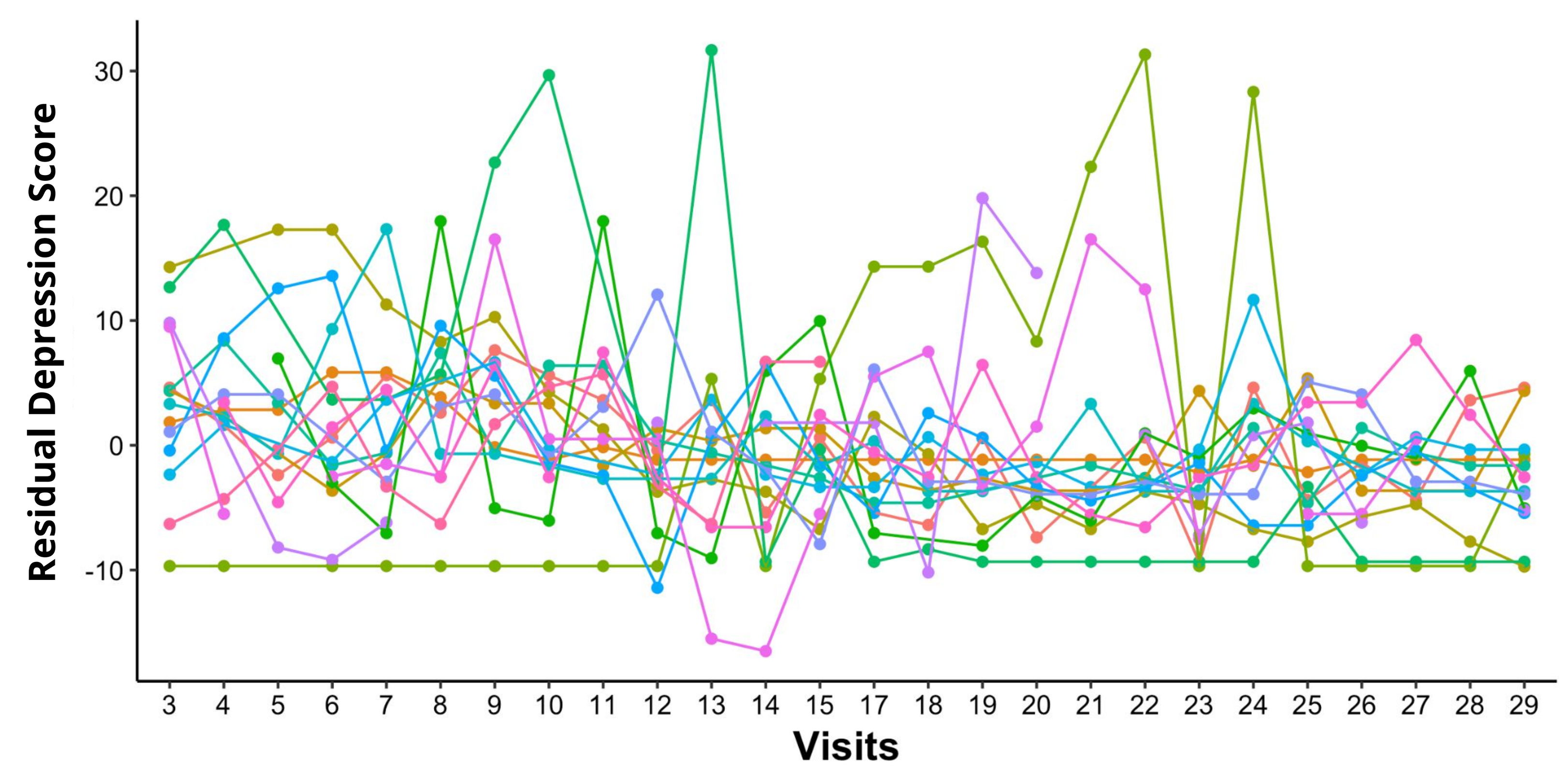

Residual Depression Score
Visits

Figure 2: Residual plots confirmed within- and between-subject variability in daily depression scores. Daily residual depression scores were derived for each participant by subtracting individual means from all time points, from each daily score. Each colour represents one participant.

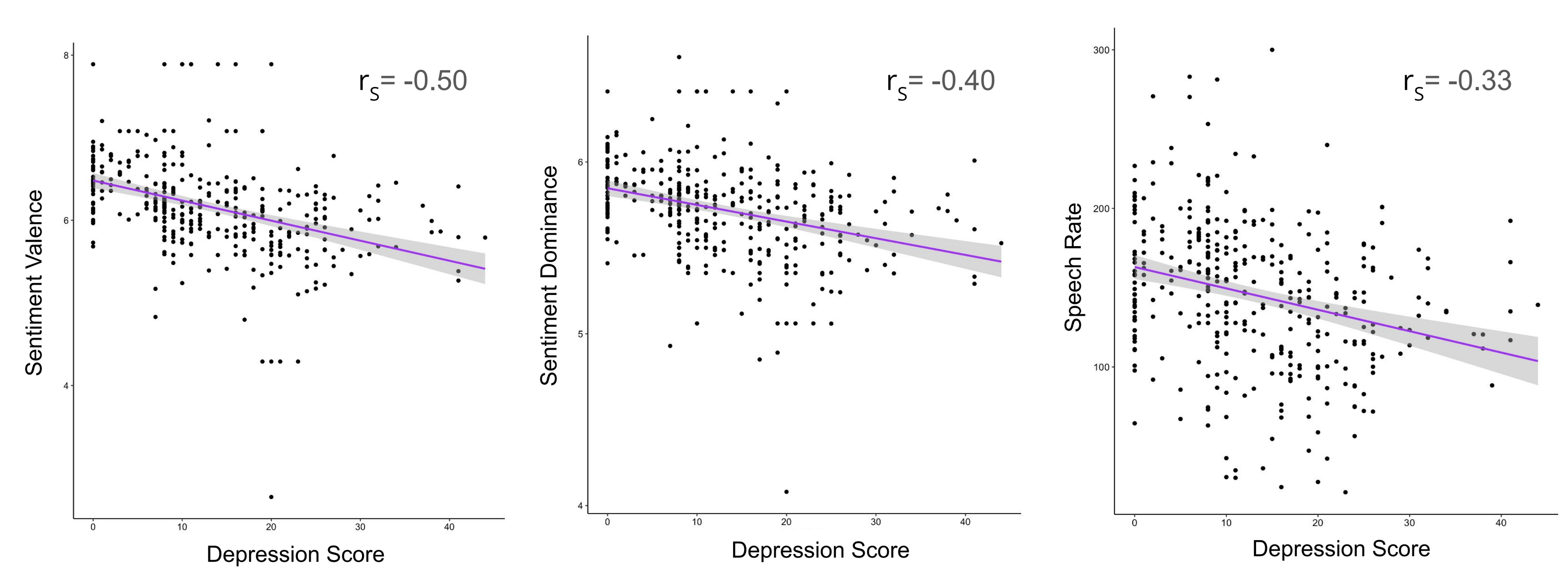

r_S= -0.50
Sentiment Valence
Depression Score
r_S= -0.40
Sentiment Dominance
Depression Score
r_S= -0.33
Speech Rate
Depression Score

Figure 3: Significant negative correlations were observed between speech and depression measures pooled across participants over time, with the strongest correlations being with sentiment valence (left), sentiment dominance (middle), and speech rate (right) (p< 0.0001, Bonferroni corrected).

Acknowledgements

We are thankful to the study participants for taking part in this research. One or more authors report potential conflicts which are described in the program.

Results

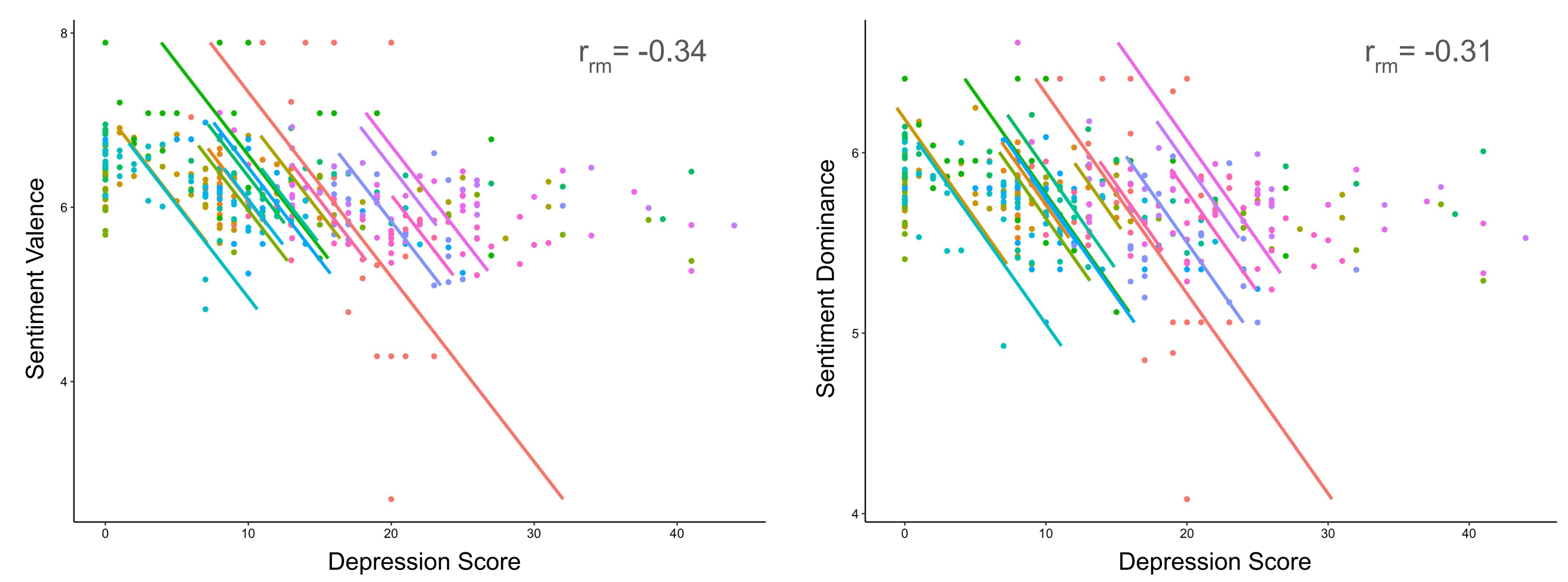

r_{rm}= -0.34
Sentiment Valence
Depression Score
r_{rm}= -0.31
Sentiment Dominance
Depression Score

Figure 4: Significant negative repeated measures correlations (r_{rm}) were observed between depression scores and sentiment valence (r_{rm}= -0.34, 95% CI [-0.43, -0.24], p< 0.001) and sentiment dominance (r_{rm}= -0.31, 95% CI [-0.40, -0.22], p< 0.001). Intra-individual associations with speech rate were not significant.

Discussion

This study supports the feasibility of high-frequency remote speech assessment as a means to understand fluctuating depression symptoms. The results indicate that speech measures could be used to monitor daily fluctuations of depression symptoms and that measures of particular relevance include sentiment valence, sentiment dominance, and speech rate.

Sentiment analysis involves the use of NLP to identify and quantify affective states through linguistic analysis. Sentiment valence informs the use of positive, negative or neutral words, with lower scores indicating the use of more negative words (e.g., unhappy). Sentiment dominance describes feelings of power, with lower scores indicating words that denote feeling less power (e.g., controlled). Lower sentiment valence and dominance were associated with higher depression scores when examining pooled and intra-individual data. Consistent with the literature [3, 4], higher depression was also associated with slower speech rates when examining pooled data, but not intra-individually.

Taken together high-frequency remote speech assessment to monitor depression symptoms may have clinical utility particularly with growing interest in telehealth and digital health interventions. However, future studies on larger clinical samples are warranted to determine specifically how speech may inform symptom fluctuations in the context of symptom remission and/or relapse.

References

[1] Verhoeven et al. Seeing the signs: Using the course of residual depressive symptomatology to predict patterns of relapse and recurrence of major depressive disorder. *Depression & Anxiety* 35, 148-159 (2018).
[2] Vergunst et al. Longitudinal course of symptom severity and fluctuation in patients with treatment-resistant unipolar and bipolar depression. *Psychiatry Research* 207, 143-149 (2013).
[3] Low et al. Automated assessment of psychiatric disorders using speech: A systematic review. Laryngoscope Investigative *Otolaryngology* 5, 96–116 (2020).
[4] DeSouza et al. Natural language processing as an emerging tool to detect late-life depression. *Frontiers in Psychiatry,* doi.org/10.3389/fpsyt.2021.719125.

WINTERLIGHT

Presented at ISCTM 18th Annual Scientific Meeting, 2022

McMaster University