Building and evaluating a speech biomarker from 2-seconds sustained phonation to detect and monitor Parkinson's disease

Johannes Tröger [1], Louisa Schwed [1], Ebru Baykara[1], Nicklas Linz [1], Juan Rafael Orozco-Arroyave[2]

[1]ki elements, Germany [2]GITA Lab, Universidad de Antioquia, Medellín, Colombia

International Parkinson and Movement Disorder Society

Objective

To evaluate the feasibility of a speech-based algorithm geared towards frequent remote measuring of Parkinson's Disease (PD) severity with minimal patient burden and minimal technical requirements.

Background

Monitoring of disease progression and treatment response at low-burden and low-cost is a recognised target in clinical research on PD. Thanks to the progress of speech technology, speech is among the most promising digital biomarkers as it can be collected with low patient burden over natural interactions with the patient. Besides, it enables remote evaluations with minimal tech requirements at a low cost.

Method

The speech biomarker was developed and validated using audio recordings of a sustained phonation task from the PC-GITA data set ((1), see Table 1).
Clinicians evaluated patients with the UPDRS Motor examination part. Both patients and age-/ sex-matched healthy controls performed speech tasks, including sustained phonation of five different vowels. We focused on the most commonly used sustained phonation task (vowel /a/), repeated three times per participant.

	Building dataset		Validation dataset	
	HC	PD	HC	PD
N participants	50 (25 F)	50 (25 F)	20 (9 F)	20 (11 F)
N voice samples	150	150	59	59
Age	61.2 ± 11.3 (61.8 ± 7.8)	62.2 ± 11.2 (61.8 ± 7.8)	60.67 ± 9.97 (64.18 ± 10.27)	57.91 ± 17.69 (65.33 ± 7.94)
UPDRS-III	-	37.76 ± 14.02 (37.56 ± 22.09)	-	33.82 ± 28.01 (47.78 ± 11.11)

Table 1: Demographics. HC: Healthy controls, PD: Parkinson's Disease, UPDRS: Unified Parkinson's Disease Rating Scale.

We extracted acoustic features using our proprietary speech extraction library and selected features based on face validity for measuring motor function as well as other descriptive statistics.
Features are then aggregated across all three assessments, normalized and combined into a single score (speech-based motor biomarker - SB-M). For validation, we followed the Digital Medicine Society (DiMe) V3 framework.

Results

For analytical validation we establish construct validity by demonstrating a significant correlation between the biomarker score and an established measure of speech in PD, UPDRS 3.1, (Spearman-Rank Correlation, r = 0.33, p < 0.05, Cohen's d = 0.72).

For clinical validation, we assessed differences in biomarker scores between HC and PD groups and their correlation with MDS-UPDRS-III Motor Examination score. We observed a significant group difference (Kruskal Wallis, $x^2(1) = 6.74$, $p < 0.01$, Cohen's d = 0.84), along with a significant correlation (Spearman-Rank Correlation with UPDRS-total, r = 0.41, p < 0.05, Cohen's d = 0.89); see Figure 1.

Conclusion

Results suggest that it is feasible to measure PD severity with minimal patient burden, almost unobtrusively and remotely from only a couple of seconds of speech from a vocal phonation task.

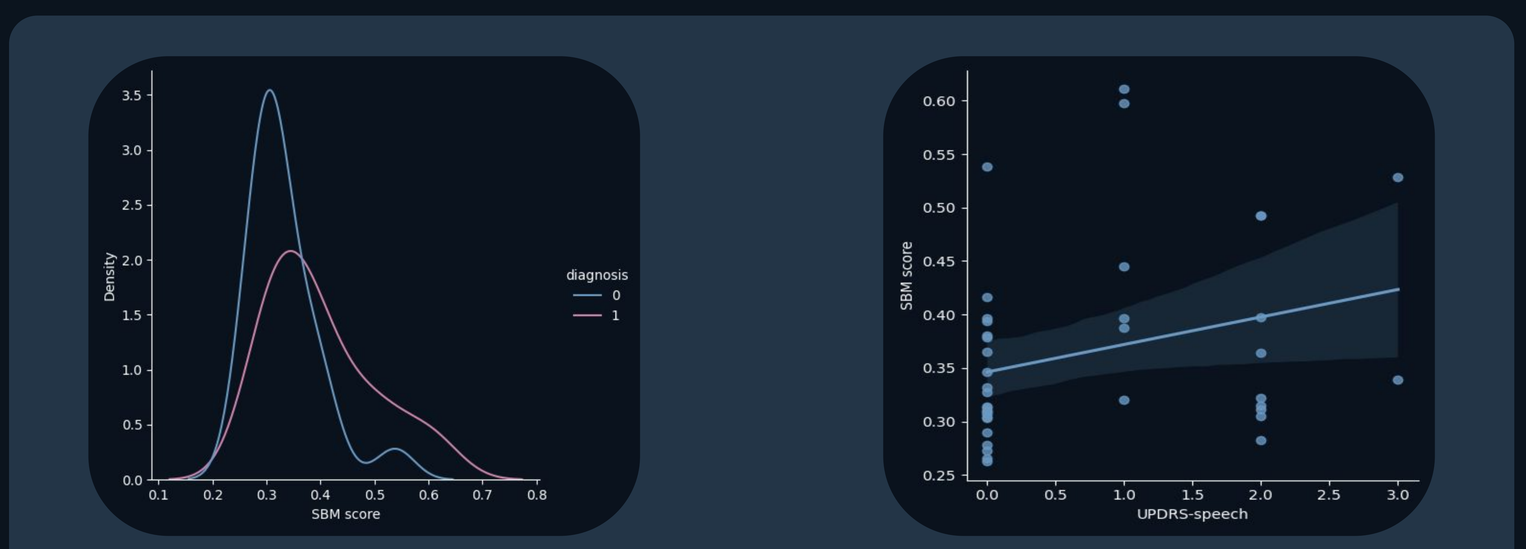

Figure 1 (a): Score distribution; diagnosis codes: 0 = HC, 1 = PD. Spearman-Rank Correlation with UPDRS-speech (r = 0.33, p < 0.05, Cohen's d = 0.72)

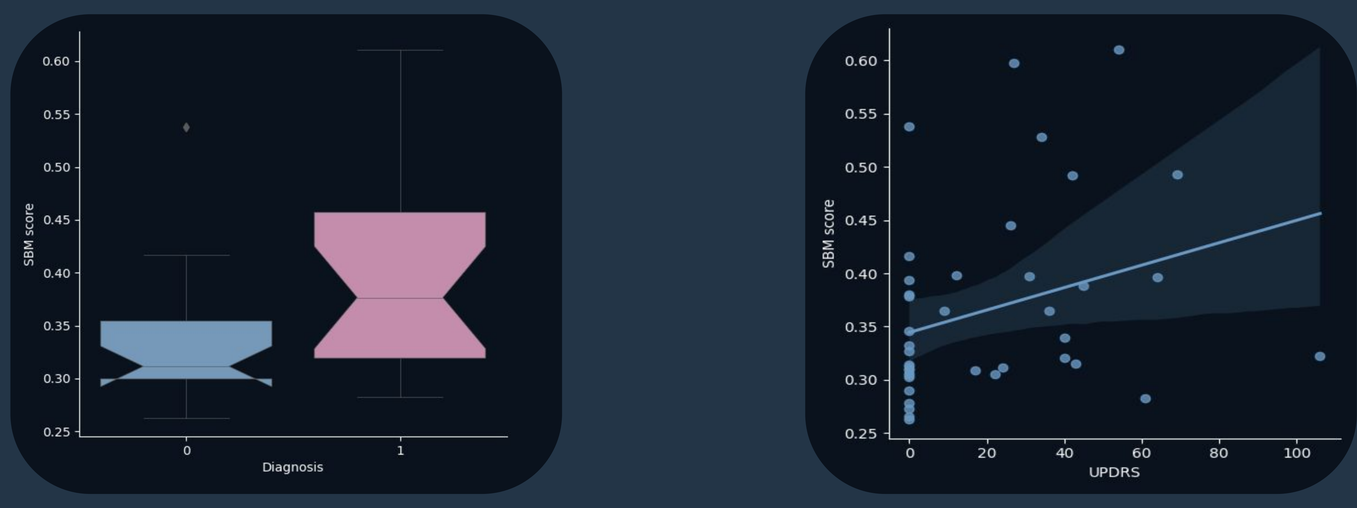

Figure 1 (b): Kruskal Wallis Group difference (x2 (1) = 6.74; p < 0.01, Cohen's d = 0.84); diagnosis codes: 0 = HC, 1 = PD. Spearman-Rank Correlation with UPDRS-total (r = 0.41, p < 0.05, Cohen's d: 0.89)

1. *Orozco-Arroyave, J. R., Arias-Londoño, J. D., Vargas-Bonilla, J. F., Gonzalez-Rátiva, M. C., & Nöth, E. (2014, May). New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease. In LREC (pp. 342-347).*