



2023 ANNUAL CONGRESS

SCHIZOPHRENIA INTERNATIONAL
RESEARCH SOCIETY

Toronto

THE WESTIN HARBOUR CASTLE
TORONTO, CANADA
11-15 MAY 2023

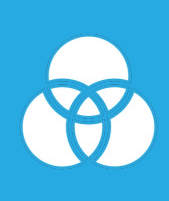
OVERCOMING GLOBAL ADVERSITY: MIND AND BODY MATTER

SCHIZOPHRENIARESEARCHSOCIETY.ORG

Multimodal speech and facial digital assessment to assess negative symptoms

Measuring for Expressive Negative Symptoms

Jean-Pierre Lindenmayer, Anzalee Khan, Vikram Ramanarayanan,
David Suendermann-Oeft, David Pautler, Benedicto Parker & Mohan Parak

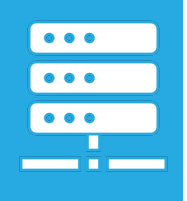


Negative Symptoms Distinction

Primary and Secondary Expressive Negative Symptoms

Type	Description
Primary Expressive Negative Symptoms	Primary expressive negative symptoms are intrinsic to the underlying pathophysiology of schizophrenia.
Secondary Expressive Negative Symptoms	Can be due to positive symptoms, depression, pharmacological effects of the antipsychotic treatment causing extrapyramidal (EPS) side effects. These EPS are also referred to as pseudo-parkinsonism symptoms as they share their clinical presentation with Parkinson's disease (PD)

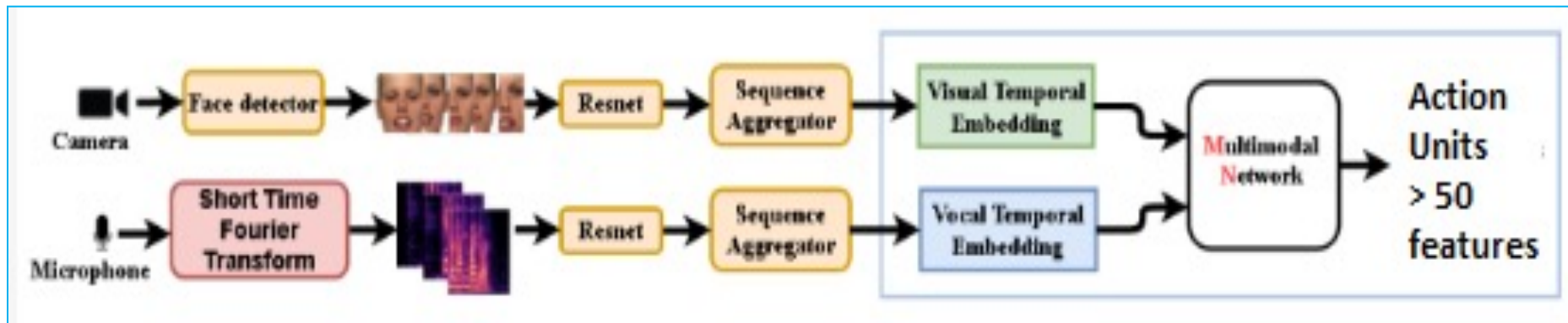




Facial Expressions and Acoustic Features

Expressive Negative Symptoms

- In recent years, progress has been made in audiovisual data processing
- Advances in this technology could play a pivotal role in supporting automated methods of collecting objective adjunctive patient data to inform negative symptom identification and monitoring
- Speech analysis: Has demonstrated that participants with schizophrenia tend to exhibit less total time talking, reduced speech rate, and higher pause duration than healthy participants
- Facial feature analysis: Alterations in facial expressivity are hallmarks of blunted affect in schizophrenia and includes flat or inappropriate affect in individuals with schizophrenia

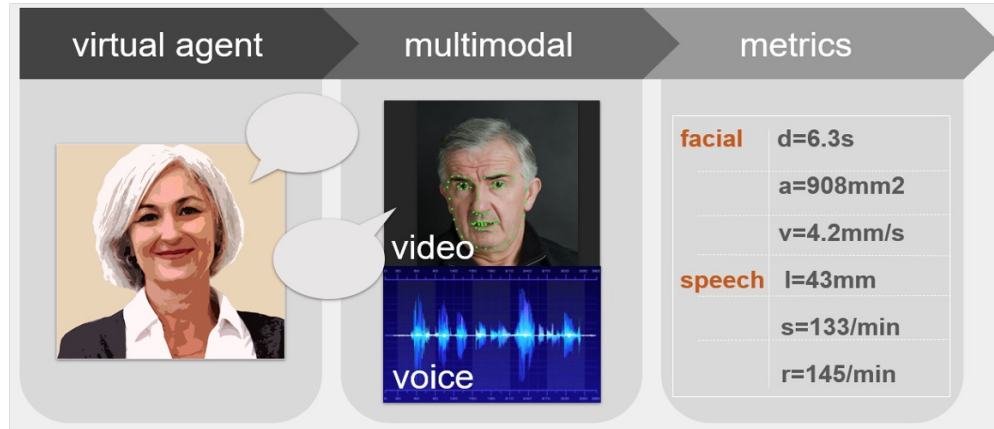


The multi-modal facial and vocal systems using Deep Neural Network (DNN) approach. Video frames and audio spectral segments get independent embeddings that produce separate features and can be fused by a multi-modal network

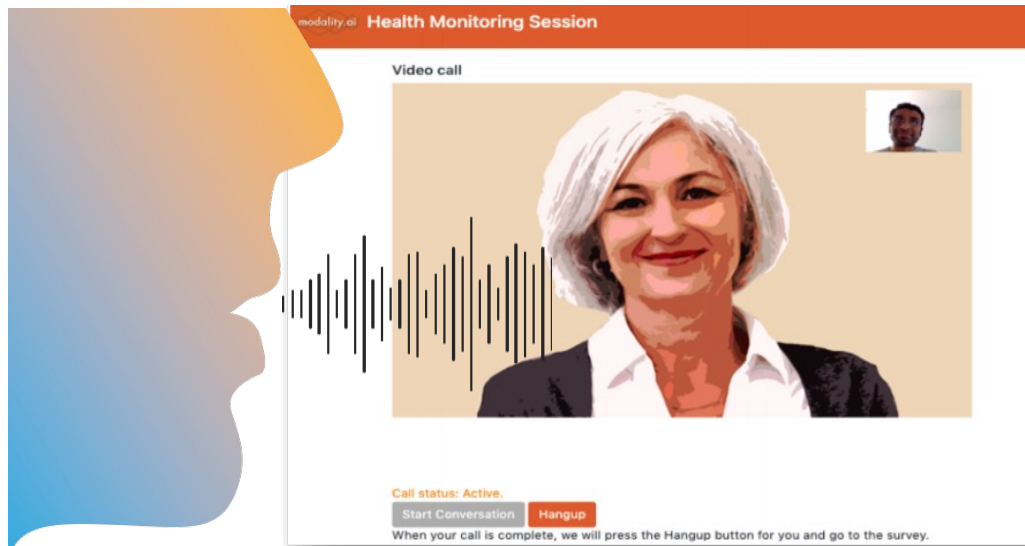


How does the multimodal system work?

For Expressive Negative Symptoms Measurement



Computer-based negative symptom measure: Participants interact with an avatar that provides a series of emotionally-ambiguous, valence-neutral tasks including a series of reading aloud tasks composed of sentences and a passage; an eyebrow raising task, an image description task, and a free speech task related to a topic of interest from the list provided.



Speech and Facial Action Units include:

Phonation

Cepstral Peak Prominence (level of noise in vocal signal, measures dysphonia)

Speech Intelligibility (SIT), Duration, and Rate (with and without pauses)

Articulation Rate and Loudness

DDK - also known as syllable alternating motion rate (AMR), assesses repetitive movements of oral articulators

Internal Silence (pauses)

Syllable Rate and Count

Lip Aperture

Mouth Surface Area

Jaw Velocity and acceleration

Lower Lip Velocity and Acceleration

Eye Opening and Eyebrow vertical position

Head Tilt





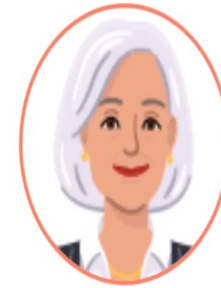
Video Presentation

Facial and Acoustic Expressions

- Transdiagnostic phenomenon of reduced facial and vocal features in:
 1. Parkinson's disease
 2. Schizophrenia
 3. Depression
 4. Apathy in dementia
- The multimodal software could be used to differentiate these illness specific "negative symptoms"
- Using the multimodal software for assessment of expressive negative symptoms is a step in this direction.

7

Talk with Tina



This is Tina. She will be guiding you through your session.



When the "Start Conversation" button below turns orange, you can use it to start your conversation.

If your call does not go through, or you hear only a short tone, please end the call and try again after a few minutes.





AIMS of our Study

Expressive Negative Symptoms

Aim 1

To investigate whether facial and voice features from the AI-enabled vocal and facial program could discriminate between individuals with schizophrenia and negative symptoms (SZ) and healthy controls (HC)

Aim 2

Investigate whether negative symptoms can be meaningfully measured using AI-enabled vocal and facial analysis software by comparing speech metrics (e.g., prosody, rate, intelligibility, pausing duration etc.) and video metrics (e.g., specific facial and head movements) and compared to clinician-rated psychometric assessments measuring expressive negative symptoms





Design and Methods

Expressive Negative Symptoms

Inpatients with schizophrenia (n = 42) and demographically matched healthy controls (n = 44) completed a brief 8-minute assessment using the multimodal software at Time 1 and Time 2 (4-7 days apart)

Inclusion for SZ: Inpatients with diagnosis of schizophrenia, age 18 - 65, English speaking, WRAT-IV Reading Score $\geq$ 8th grade, Negative symptoms as evidenced by score of $\geq$ 18 on PANSS NSF

Exclusion for SZ: A score >3 on the sum of the first eight items of the Simpson-Angus Scale (SAS)

Inclusion for HC: Individuals with no prior history of mental illness, age 18 - 65, English speaking.

SZ Group

Visit 1

Positive and Negative Symptoms Scale
Clinical Global Impression - Severity
Brief Negative Symptoms Scale
Calgary Depression Scale
EPS Scales (BARS, AIMS, SAS)
Multimodal Software



4 - 7 days

Visit 2

Positive and Negative Symptoms Scale
Clinical Global Impression – Improve...
Brief Negative Symptoms Scale
Calgary Depression Scale
EPS Scales (BARS, AIMS, SAS)
Multimodal Software

HC Group

Visit 1

Multimodal Software



4 - 7 days

Visit 2

Multimodal Software

All assessments were completed by trained raters > 20 year experience





Data Analysis

Expressive Negative Symptoms

- Random Forest (RF) classifiers were built to discriminate between classes (Breiman, 2001).
 - All classifications were binary (e.g., SZ versus HC). Data for Time 1 and Time 2 during the study period were included in the analyses. We evaluated RF models on the resulting data set through a five-fold participant-based cross-validation.
- Each unit from the multimodal software were compared to negative symptom severity overall using the PANSS Negative Symptoms Factor Score (PANSS NSFS) and using Pearson's correlation
- Digital measurements that demonstrated significance in relation to specific subscales or items on the BNSS and PANSS were then further explored in relation to the specific symptoms that derive from those subscales, correcting for multiple comparisons using a Benjamini-Hochberg adjusted p-value (Li & Barber, 2019) $((i/20)*0.2)$ where $i = \text{rank of } p\text{-value}$





Results

SZ and HC Demographics

Characteristics	Schizophrenia		Healthy Controls	
	Mean	SD	Mean	SD
Age (in years)	39.96	11.60	42.18	12.23
	n	%	n	%
Sex				
Male	31	73.81	28	63.63
Female	11	26.19	16	36.37
Race				
Black	28	66.67	25	56.82
White	14	33.33	19	43.18
Asian	0	0	0	0
Other	0	0	2	6.45
Ethnicity				
Hispanic	20	47.62	20	45.45
Non-Hispanic	22	52.38	23	52.27
Not reported	0	0	1	2.27

PANSS	n	Mean	SD
Positive Subscale	42	19.22	4.56
Negative Subscale	42	25.46	3.12
General Psychopathology	42	39.11	5.66
PANSS Total	42	85.12	10.45
Marder Positive Symptom	42	21.45	4.01
Marder Negative Symptom	42	25.46	2.66
Marder Disorganized Symptom	42	20.12	3.99
Marder Hostility Symptom	42	7.90	2.37
Marder Anxiety Symptom	42	6.12	2.34

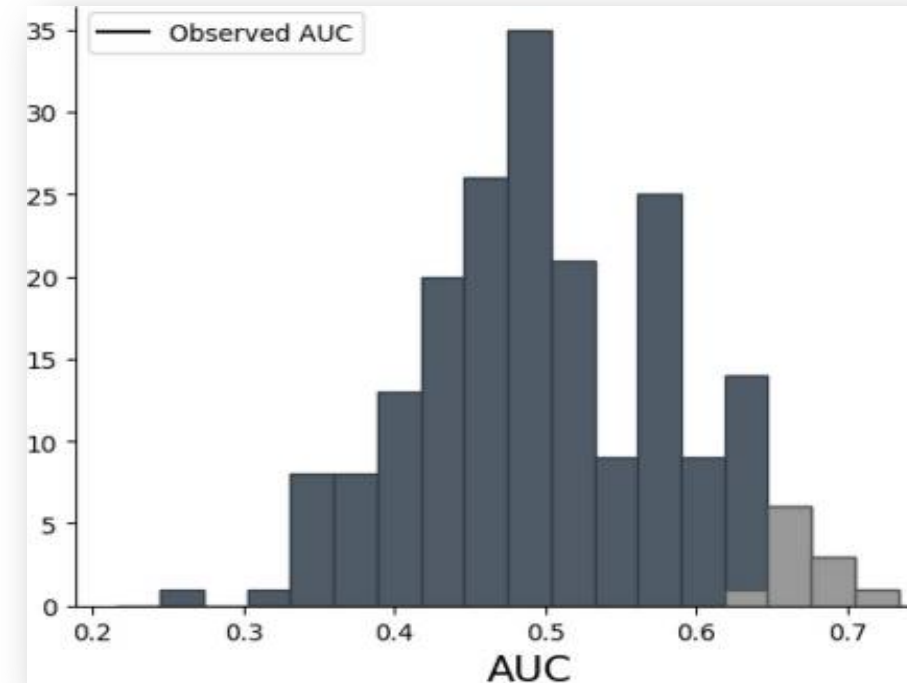




Results

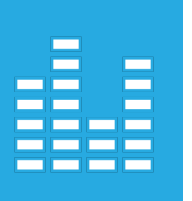
SZ compared to HC

- The sensitivity and specificity for classifying patients with SZ versus HC was **0.78 (SD 0.10)** and **0.56 (SD 0.21)**, respectively and with an **AUC of 0.77 (SD 0.13)**.
- The figure presents the generated null-distribution of AUC scores from permuted class labels as generated from the randomized permutation model. The lighter area shows the critical level for a one-tail test with a significance level of 0.05
- The observed AUC for participants: SZ versus HC ($p < 0.001$)



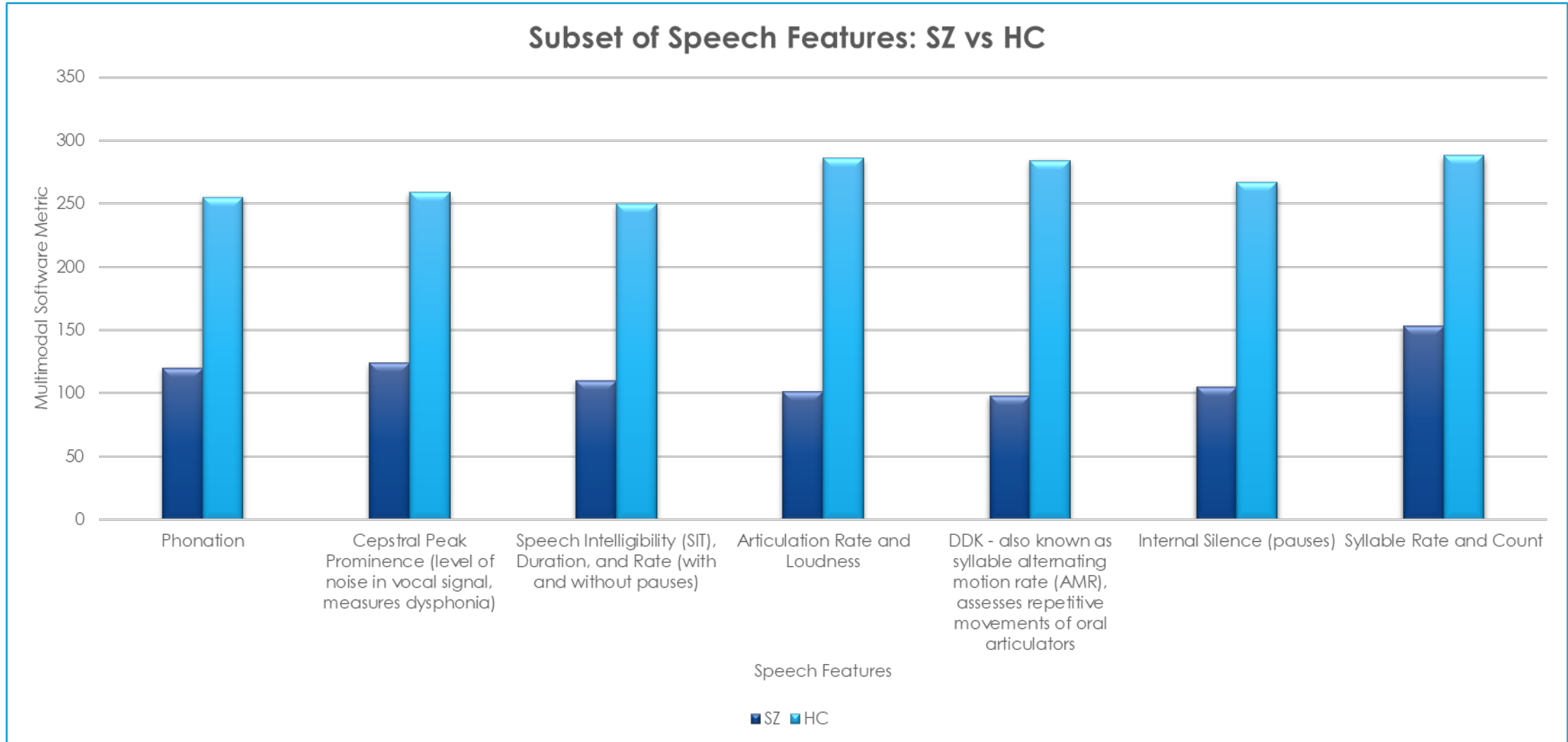
Binary Classifier	Model Type	Accuracy (SD)	F1 Score (SD)	Sensitivity	Specificity	AUC
SZ vs HC	Random Forest Model	0.73 (0.10)	0.80 (0.05)	0.78 (0.10)	0.56 (0.21)	0.77 (0.13)





Results

Expressive Negative Symptoms: Acoustic Features



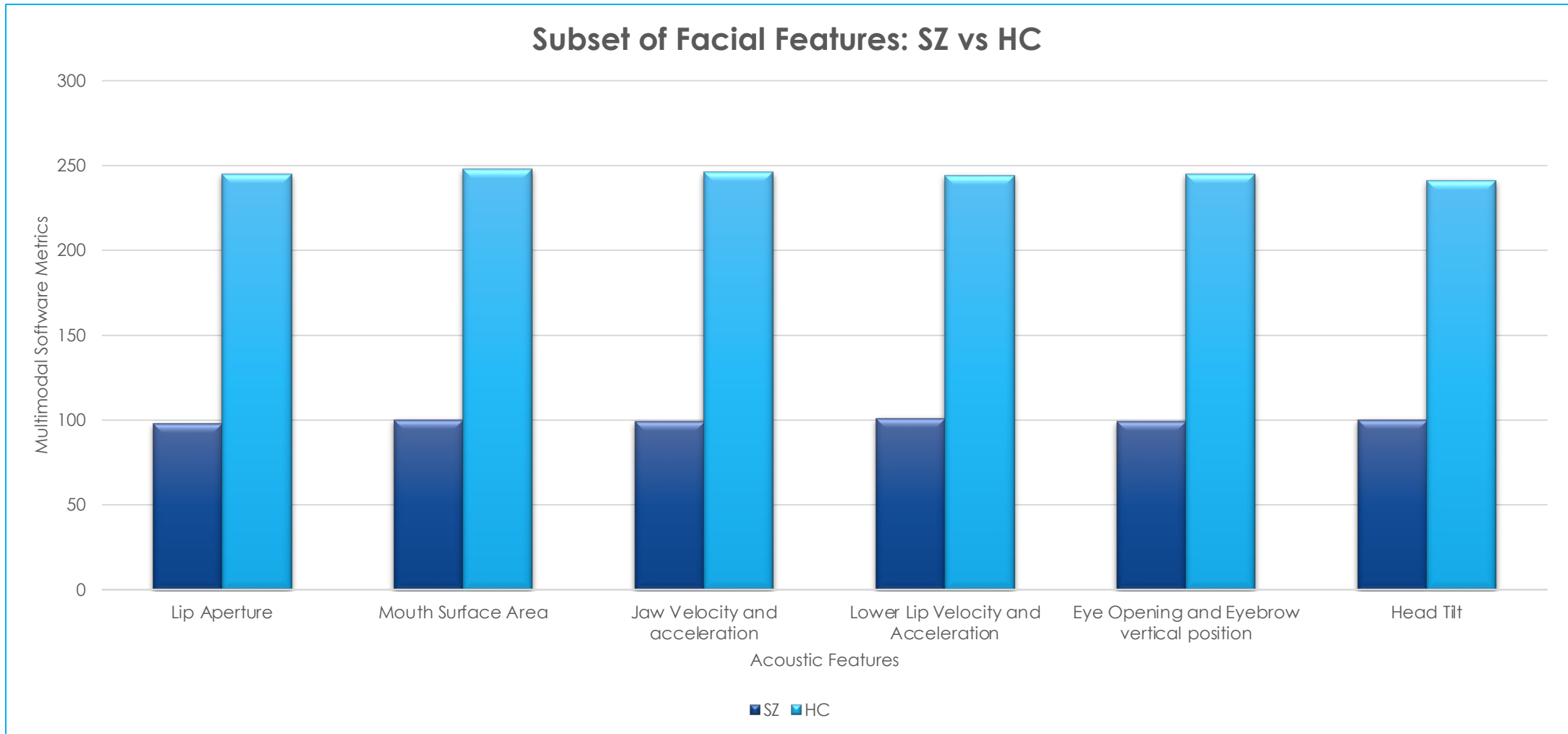
Lower values represent deficits in speech features





Results

Expressive Negative Symptoms: Facial Features



Lower values represent deficits in movement of facial features



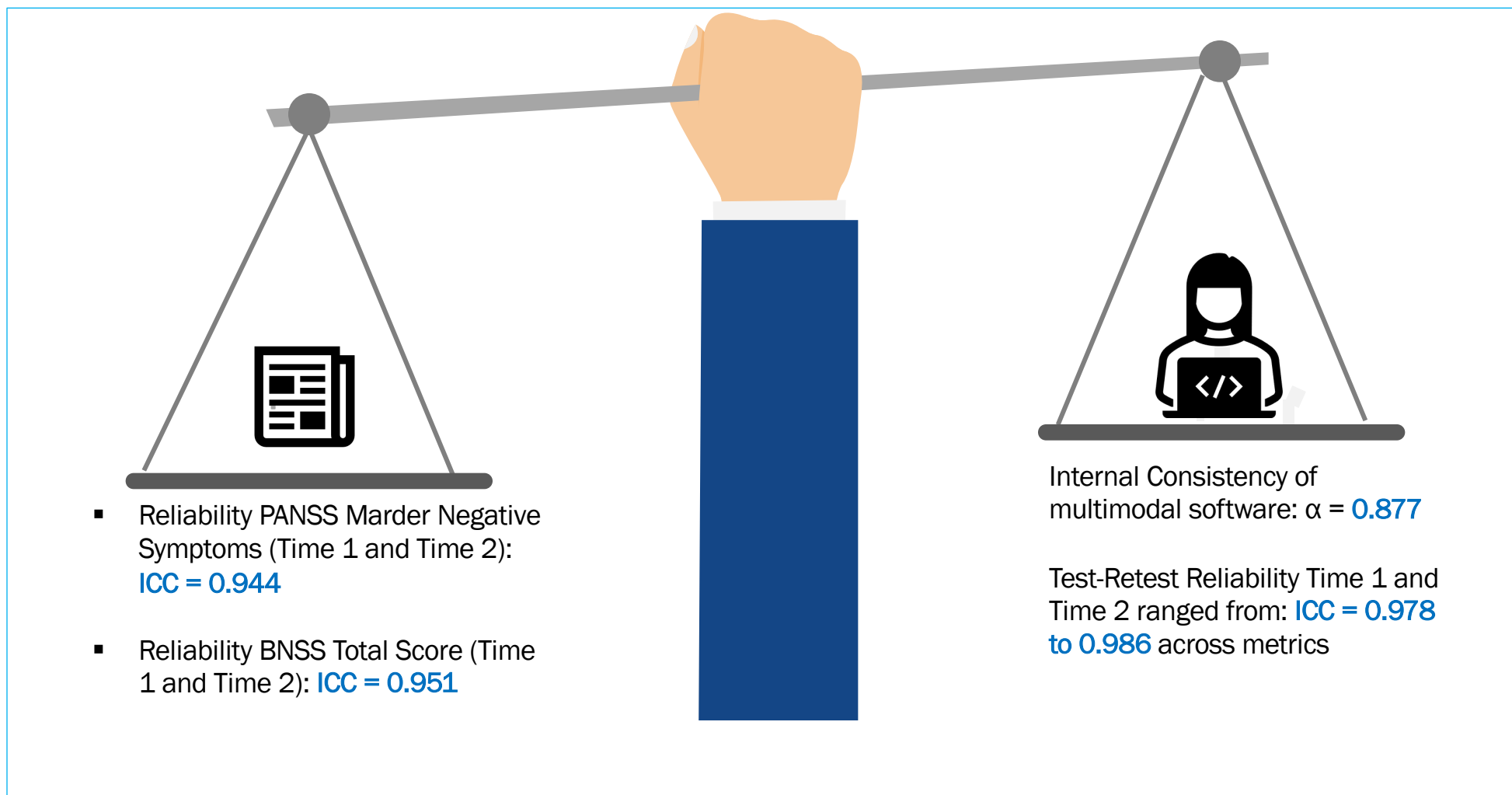


Comparisons	Significant Findings
Acoustic Features with PANSS Negative Factor Score	- Articulation rate (words/minute) ($r = -0.62$; adjusted $p = .03$) - Average fundamental frequency ($r = -0.61$; adjusted $p = 0.04$) - Percent pause time ($r = 0.60$; adjusted $p = .02$) - Speaking rate (words/minute) ($r = -0.64$; adjusted $p = .05$) Increased negative symptom severity was reflected in decreased Articulation rate, decreased Speaking rate, and increased Percent pause time.
Acoustic Features with BNSS and PANSS Items	Significant negative correlations with multiple vocal measures (i.e., Articulation rate (words/minute), Average fundamental frequency, and Speaking rate (words/minute)) and: - Quantity of speech ($r = -0.63$; $p = .02$) - Spontaneous elaboration ($r = -0.62$; $p = .02$) - PANSS Motor Retardation ($r = -0.53$; $p = .04$)
Facial Features with BNSS and PANSS Items	- PANSS Blunted Affect ($r = -0.49$; $p = .04$) - PANSS Social Withdrawal ($r = -0.63$; $p = .005$) - PANSS Motor Retardation ($r = -0.45$; $p = .06$) BNSS Facial expression and BNSS Expressive Gestures and average number of eye blinks in blinks per second for both eyes, Mean and maximum speed of the jaw (center, or tip of the chin) and average and maximum lip aperture/opening.



Results

Reliability: Multimodal software and Clinician-Administered Scales





Conclusions

Multimodal Software for Expressive Negative Symptoms

- Utilizing an AI-based program that is also being studied in PD can help better identify which metrics are reliable predictors of expressive negative symptoms
- Preliminary data shows a large separation between SZ and HC and good reliability and internal consistency of the NEMSI digital assessment instrument
- Test-retest shows greater reliability across all metrics compared to clinician-administered scales
- Although the multimodal digital program cannot capture all aspects of negative symptoms, these tools can serve as a supplement to clinician assessments and may provide more precise measurement of symptoms
- The differentiation between primary expressive negative symptoms and secondary expressive negative symptoms may allow for better understanding of the underlying pathophysiology and for treatment of negative symptoms





Future Steps

Multimodal Software for Expressive Negative Symptoms

- **Continued enrollment to increase the sample size to develop more robust analysis models and to potentially develop factor structures that may be able to identify composite scores for multimodal software that map on to key negative symptom constructs:**
 - Decreased thought and speech productivity
 - Speech with little or no change to their tone
 - Little or no change in their facial expression
- **Examine sensitivity to change**
 - Difficult due to lack of efficacious treatment
- **Compare scores for facial and acoustic metrics between primary negative symptoms and secondary expressive symptoms due to EPS**
 - Disentangling the primary from the secondary sources of negative symptoms is important for a number of reasons, including implementing appropriate treatment strategies



Multimodal Software Modality AI

Team at Manhattan Psychiatric Center



David Suendermann-
Oeft
Chief Executive Officer
PhD, EE



David Pautler
Chief Technology
Officer
PhD, CS



Ira Shoulson
Chief Medical Officer
MD



David Fox
Chief Business Officer
Serial Entrepreneur



Vikram
Ramanarayanan
Chief Science Officer
PhD, EE

