Multimodal AI Based Facial and Acoustic Biomarkers of Negative Symptoms in Schizophrenia

Anzalee Khan1;2, Jean-Pierre Lindenmayer1;2,3; Saqib Bashir1;2; Sebastian Prokop1;2;4, Beverly Insel1;2; Brianna Fitapelli1;2;5, Krishnapriya Bodicherla1;2;3,4, Mohan Parak1;2, Benedicto Parker1;2; Christian Yavorsky6; Hardik Kothare7; David Pautler7; David Suendermann-Oeft7; Vikram Ramanarayanan8,9


METHODOLOGICAL QUESTION

- Can negative symptoms in schizophrenia be meaningfully measured using an AI-enabled audiovisual dialog system? If reliability and validity are adequate, results can be used to contact-free, non-invasive, cost-effective assessment and monitoring of negative symptoms.
- Many individuals with schizophrenia present with negative symptoms including abnormalities in vocal expression, such as altered vocal production (i.e., alogia, reduced speech) and intonation-emphasis (i.e., blunted affect; effective liability). This is reflected in communication via coupled mechanisms: vocal articulation, facial gesturing and dialogue content.
- One barrier in understanding and measuring vocal abnormalities in negative symptoms is a reliance on clinician-based rating scales - these scales can be subjective, insensitive to change in treatment, require extensive training, lack regional and cultural adaptability, and have abstruse operational definitions.
- Speech behaviors and facial movements can inform clinicians about negative symptoms and include monotonous and monosyllabic speech, few gestures, pausing, speech rates, and speed of movement of certain facial areas. Facial and speech changes in negative symptom patients are difficult to track and quantify with conventional techniques. A rising number of conversational agents (or chatbots) are equipped with artificial intelligence (AI) architecture which are non-invasive and can capture facial and speech movements efficiently.

AIMS

To investigate whether negative symptoms can be meaningfully measured using AI-enabled audiovisual dialog system called Neurological and Mental Health Screening Instrument (NEMSI assessment) by comparing speech metrics (e.g., prosody, rate, intelligibility, pause duration etc.) and video metrics (e.g., specific facial and head movements) to clinician-rated psychometric assessments for negative symptoms.

METHOD

- Experimental Approach: At the first visit, the following instruments are administered: sociodemographic and clinical questionnaire, PANSS, BNSS, CDS, CGI-S, AIMS, SAS, BARS and NEMSI assessment. The second visit occurs within a one-week period and is done by the same clinician to assess test retest reliability and intra-rater reliability. The second visit includes the same instruments in addition to the CGI-I (severity of illness, improvement, and degree of change). Healthy controls only performed the NEMSI assessment.
- Patient Eligibility: Inpatients with diagnosis of schizophrenia, age 18-60, English speaking, WRAT IV Reading score ≥ 8th grade. Negative symptoms as evidenced by score of ≥ 18 on PANSS Marder Negative Symptom Factor.
- Healthy Control Eligibility: Individuals with no prior history of mental illness, age 18-60, English speaking.

Analysis:
1. Reliability (Intra Class Correlation Coefficient, ICC) 2. Validity: concurrent, convergent, divergent and discriminative using Pearson r correlation of NEMSI assessment speech and facial metrics to the BNNS, PANSS Marder Negative factor and the CDS5. Internal Consistency of NEMSI using Cronbach Alpha a. 4. Comparison of HC and schizophrenia using ANOVA and descriptive statistics

RESULTS: RELIABILITY AND VALIDITY

- Reliability NEMSI AI (Time 1 and Time 2): ICC = 0.982
- Reliability PANSS Marder Negative Symptoms (Time 1 and Time 2): ICC = 0.953
- Reliability BNSS Total Score (Time 1 and Time 2): ICC = 0.956
- Validity (Pearson r) of NEMSI with 1. BNSS Total Score = 0.801, 2. BNSS Alogia = 0.812, 3. BNSS Avolition = 0.844
- Internal Consistency (Cronbach a) NEMSI: 0.867

CONCLUSIONS

- Speech and facial AI technology could aid in negative symptoms assessments
- NEMSI variables articulation rate and average mouth surface have strong correlations to negative symptoms.
- The NEMSI assessment showed good to excellent reliability, validity, and internal consistency.