Audio and video digital measurements observed in naturalistic settings predict PANSS Marder Factor scores in schizophrenia

Sarah M. Kark[1], Deshana Desai[1], Tejendra. R. Patel[2], William P. Horan[2], Soumya A. Chaturvedi[2], Aaron J. Masino[1,3], Richard H. Christie[1]

[1]AiCure, New York, NY, United States
[2]Bristol Myers Squibb, Princeton, NJ, United States
[3]The School of Computing, Clemson University, Clemson, SC, United States

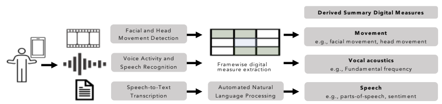
The Methodological Issue Being Addressed

Determine predictive capability of audio-video digital measurements collected during a remote picture description task to predict schizophrenia symptom severity, utilizing all available longitudinal clinical trial data points.

Introduction

- Standard regression analysis assumes independence between observations, however longitudinal datasets violate this basic assumption with repeated within-subject measures.
- Generalized estimating equations (GEE) are to designed to leverage longitudinal datasets by modeling within-subject correlations separately, resulting in parameter estimates that can indicate the direction and strength of relationships between digital measurement predictors and PANSS clinical scores based on all available datapoints.
- Here, we leverage over 1,000 longitudinal datapoints to determine relationships with audio-video digital measures and PANSS factor scores.

Fig 1. Data Pipeline: The AiCure app records participant audio and video during a picture description task. Computational pipeline automatically extracts framewise and summary-level multi-modal measures (vocal, facial, speech, and movement digital measurements).



Study Design and Methods

Patients:

- 227 clinically stable outpatients with schizophrenia aged 18-65 (154 males, 73 females) participating in open-label clinical trial
- Patients completed **remote digital assessments** (Fig. 2) and the **Positive and Negative Syndrome Scale (PANSS)** up to 11 times across a 48-week period (Table 1).

Digital Picture Description Task:

- **Stimuli:** 1 pleasant image and 1 unpleasant image, interleaved with and 2 neutral pictures (Fig. 2)
- Emotional digital measurements were scaled by neutral digital measurements (e.g, Negative / Neutral; and see Fig. 3) and neutral image digital measurements were entered as raw values
- We herein refer to digital measures collected under specific task contexts as "digital traits"

Figure 2. Remote smartphone picture description task.

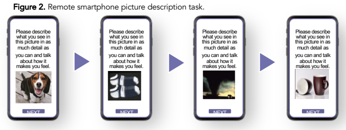
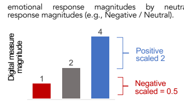


Table 1. Number of available observations per week for entry into GEE analysis.

Week	Sample Size
1	177
2	30
4	172
8	149
12	133
16	117
20	116
24	114
32	81
40	49
48	35
TOTAL OBSERVATIONS	1073

Figure 3. Scaling digital measures. Sample digital measurement data to demonstrate scaling emotional response magnitudes by neutral response magnitudes (e.g., Negative / Neutral).



Digital measurements and Analysis:

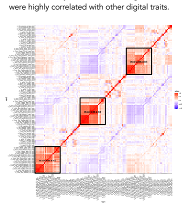
Feature reduction approach reduced the number of features for analysis by 60%:

- **204 digital traits** (68 digitals traits per emotional valence category)
  - Removed **80 traits** with > 10% missingness
  - Removed **44 traits** with high redundancy (r>0.8) (84% were movement-related digital traits [Fig. 4])
- **80 digital traits** entered the final GEE analysis (51% acoustic, 36% movement, 13% speech)

Generalized Estimating Equations (GEE) Approach:

- GEE implemented using *repolr* in R using an auto-regressive correlation structure (ar1)
- **6 PANSS sub-scale scores** re-coded into 3 ordered severity groups (low: $\leq Q_1$, mid: Q_1-Q_3, high: $\geq Q_3$)
  - 5 PANSS "Marder" Factors score (with Negative Symptoms split by Expressive and Experiential [Fig. 5])
- **Conducted 480 GEE comparisons** (80 digital traits x 6 PANSS scores)
- Applied false discovery rate (FDR) correction set to 5% using q-values

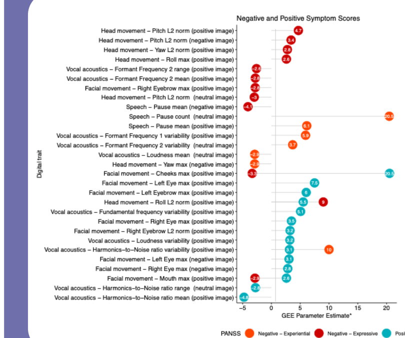
Figure 4. Feature reduction strategy. Cross-correlation analysis identified 44 digital traits that were highly correlated with other digital traits.



Results

- 48 GEE associations survived FDR correction, with most comparisons associated with Negative Symptoms (40%) or Positive Symptoms (29%) (Fig. 5)
- Negative symptoms were associated with reduced emotional facial movement, psychomotor slowing, and reductions in expressive vocal prosody measures.
- Conversely, positive symptoms were associated with increased emotional facial, head movement, and increased vocal prosody.

Figure 5. GEE Results for Negative and Positive Symptoms for comparisons that survived FDR correction. Note: *Positive parameter estimate denotes positive correlation, the sign has been changed for easier interpretation. Variability denotes standard deviation across the video. L2 norm is calculation of distance.



PANSS Score | **Patterns associated with greater symptom severity:**

Negative Symptoms Expressive (Blunted Affect + Poor Rapport + Lack on spontaneity + Motor retardation)

- **Head movement** shows overall psychomotor slowing for neutral images but relatively increased emotion-related head movement
- **Blunted facial movement** toward emotionally pleasant images (e.g., anhedonia-like pattern)
  - Reversed pattern with greater **positive symptoms**
- **Shorter speech pauses** during negative picture descriptions consistent with emotional reactivity

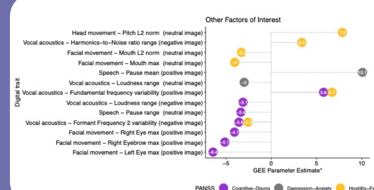
Negative Symptoms Experiential (Emotional Withdrawal + Passive Social Withdrawal + Active Social Avoidance)

- **Quieter speech and more frequent pauses** to neutral images consistent with reduced prosody

Positive Symptoms (Delusion + Hallucinatory Behavior + Grandiosity + Suspiciousness + Stereotyped thinking + Somatic concern + Unusual thought content + Lack of insight)

- **Elevated facial movement** toward emotional images
  - Reversed pattern with greater **Negative-Expressive symptoms**
- **Head movement** shows increased emotional reactivity (similar to high **Negative-Expressive**)
- **Increased acoustic variability** during positive pictures is consistent with increased vocal prosody

Figure 6. GEE Results for other Marder factor comparisons that survived FDR correction. Note: *Positive parameter estimate denotes positive correlation, the sign has been changed for easier interpretation. Variability denotes standard deviation across the video. L2 norm is calculation of distance.



PANSS Score | **Patterns associated with greater symptom severity:**

Hostility – Excitement (Hostility + Excited + Uncooperativeness + Poor impulse control)

Depression-Anxiety (Anxiety + Guilt + Tension + Depression)

Cognitive-Disorganization (Conceptual disorganization + Difficulty with abstract thinking + Mannerisms + Disorientation + Poor attention + Disturbance of volition + Preoccupation)

- **Overall greater head movement** and **reduced facial movement** toward neutral images might suggest gross motor agitation and flat affect in the absence of an emotionally salient stimulus
- **Longer pauses** toward positive images and **constricted range of vocal loudness** toward neutral images consistent with reduced responsivity (anhedonia-like pattern)
- **Many patterns shared with other factors** limits interpretation

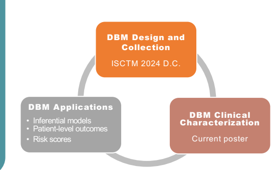
Conclusions

- Audio-video digital measurements collected in naturalistic setting predict PANSS symptom severity scores, demonstrating the utility of brief remote assessments to predict sub-scale scores in the context of a clinical trial
- Some digital traits showed opposing patterns between negative and positive symptom scores, consistent with the orthogonal nature of the factor scores and further suggesting separable behavioral patterns in audio-video digital traits
- GEE goes beyond standard cross-sectional analyses by leveraging all longitudinal datapoints to determine digital measurements that predict clinical scores

Next steps:

1. **Use GEE models to predict PANSS scores in an independent dataset**
   - Accuracy of the models to predict symptoms on an independent dataset
2. **Determine the optimal amount of longitudinal data needed**
   a. Test model predictive capability at various timepoints
   b. Apply GEE approach to shorter trial with fewer datapoints
3. **Determine specificity of negative and positive symptoms**
   a. Test links between digital traits and specific items (blunted affect, motor retardation, grandiosity, unusual thought content)

Figure 7. Iterative approach to digital biomarker (DBM) and applications development.



Disclosures: At the time of the study, SMK, DD, AJM, and RHC were employed by AiCure and held stock options in AiCure. TP, WH, and SC are employees of Bristol Myers Squibb.