# Developing Data-Driven Clinical Signatures for People Who Experience Hallucinations

U01MH135901

MPI Dror Ben-Zeev, Trevor Cohen

NIH

LSU
LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON

# Hallucination Assessment through Longitudinal Observation (HALO)

U01MH135901

MPI Dror Ben-Zeev, Trevor Cohen

NIH

LSU
LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON

**Dror Ben-Zeev**
PI
UW

**Trevor Cohen**
PI
UW

**Alex Cohen**
Co-I
LSU

**Serguei Pakhomov**
Co-I
UM

**Oliver Bear Don't Walk**
Co-I
UW

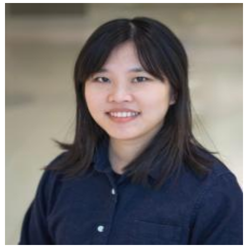
**Justin Tauscher**
Co-I/Project Director
UW

**Benjamin Buck**
Co-I
UNC

**Patrick Heagerty**
Co-I
UW

**Alexa Beaulieu**
Research Manager
UW

**Anna Larsen**
Co-I
UW

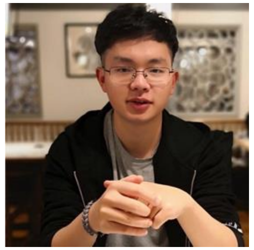


**Changye Li**
Post-Doc
UW

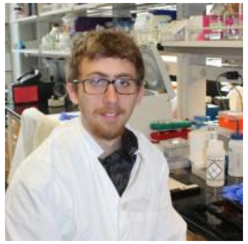
**Ella DeVries**
Research Coordinator
UW

**Gillian Sparks**
Research Coordinator
UW

**Michael Michalowski**
Co-I
UM



**Feng Chen**
Research Assistant
UW



**Jacob Solinsky**
Research Assistant
UM

**Maya Stemmer**
Co-I
UW

**Patrick Wedgeworth**
Co-I
UW

**Weizhe Xu**
Research Assistant
UW

**Meliha Yetsigen**
Co-I
UW

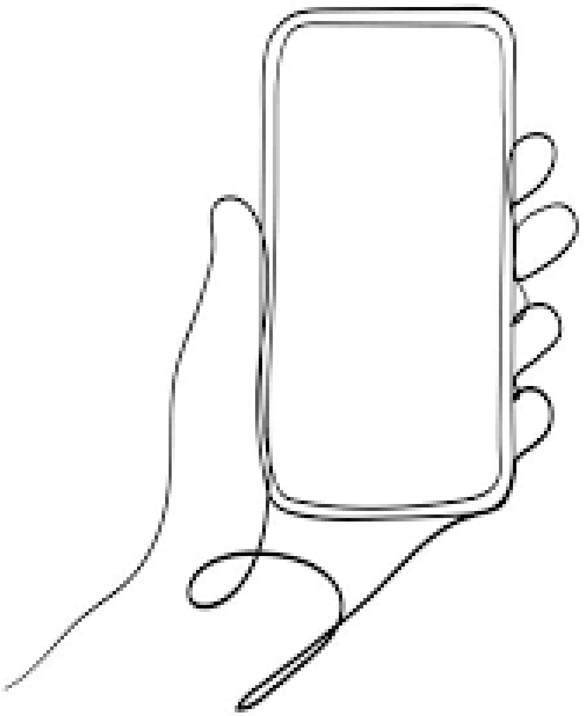
UNIVERSITY of WASHINGTON

UNIVERSITY OF MINNESOTA

LSU
LOUISIANA STATE UNIVERSITY

NIH National Institute of Mental Health

# Hallucinations and poor outcomes do not always coincide



## Hallucinations

- Prevalent in many mental disorders
- Also occur in 10% of the general population
- Don't always portend serious negative outcomes
- Early intervention can improve outcomes

## Smartphones

- Ubiquitous
- Scalable behavioral measurement platforms

## Automated speech and language analysis

- Substrate for actionable clinical signatures

NIH

LSU
LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON

# Scalable predictive models to inform clinical decisions

**Our objective is to leverage novel data sets and cutting-edge computational methods to generate clinical signatures that can be used to inform scalable and timely risk detection and clinical decision making.**

Specifically, we aim to:

1) **Derive data-driven clinical signatures** from mobile behavioral tasks to predict individual differences in severe negative outcomes among people experiencing hallucinations;
2) **Identify and mitigate inaccuracies** in modelling across groups;
3) Examine whether **adding smartphone-captured behavioral data** to information that is typically available in the clinical record **improves model clinical utility**; and
4) Produce **machine learning-ready data structures** that adhere to FAIR (Findable, Accessible, Interoperable, Reusable) principles.

NIH    LSU LOUISIANA STATE UNIVERSITY    UNIVERSITY OF MINNESOTA    W UNIVERSITY of WASHINGTON

# Computable phenotypes in the wild

- Prospective Cohort Study

- Online recruitment of **2000 individuals** 18 years or older in the USA who experience hallucinations for one year of data collection.

- Data collected from the subjects directly via data collection system downloaded on their device.

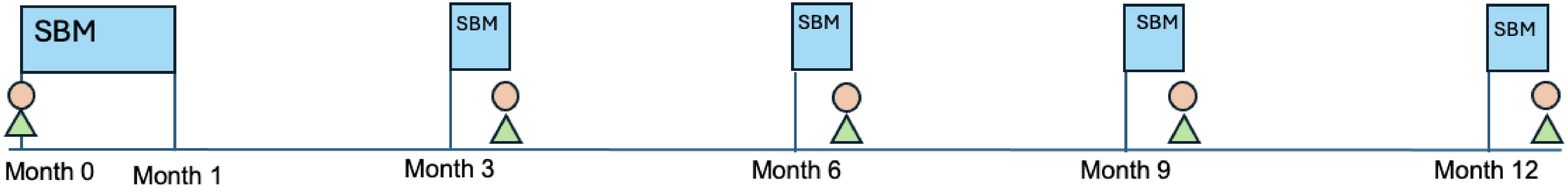
NIH

LSU
LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON

# Infrequent outcomes – intermittent extended data capture




SBM –
- 30 days at baseline (3x per week)
- 7 day "bursts" at 3, 6, 9, & 12 months (3x per week)

Clinical records –
- 0, 3, 6, 9, & 12 months

Clinical outcomes –
- 0, 3, 6, 9, & 12 months

**Smartphone Behavioral Measurement (SBM)**
- Ecological Momentary Assessment (EMA)
- Audio Diaries
- Verbal Recall Task

**Clinical Records**
- Clinical History
- Psychiatric Symptoms
- Functional Impairment

**Clinical Outcomes**
- Relevant ED visits
- Hospitalizations
- Suicide attempts / behavior

# E.g. “clinical records data”

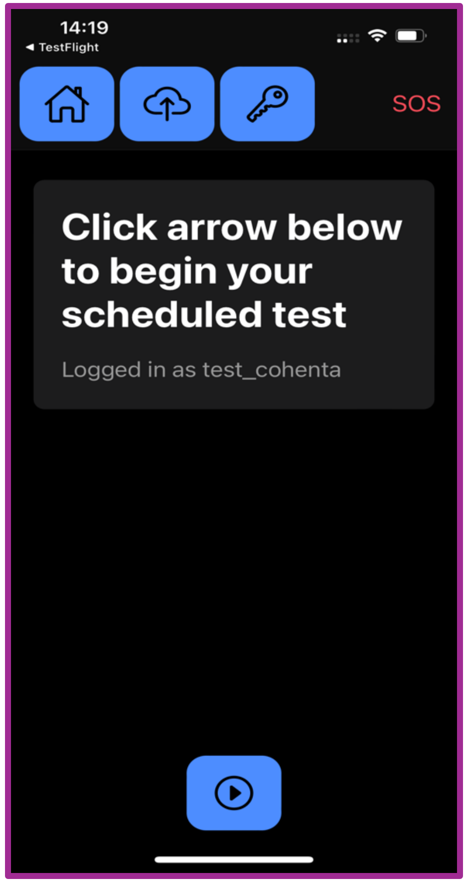
**History:** Clinical, service use, demographics
**Depression:** Patient Health Questionnaire, 9-item Version (PHQ-9)
**Anxiety:** Generalized Anxiety Disorder Questionnaire, 7-item version (GAD-7)
**Symptoms (General):** DSM-5 Cross-Cutting Symptom Measure (DSM-5-CCM)
**Functional Impairment:** WHO Disability Assessment Schedule (WHODAS)

REDCap®
Research Electronic Data Capture

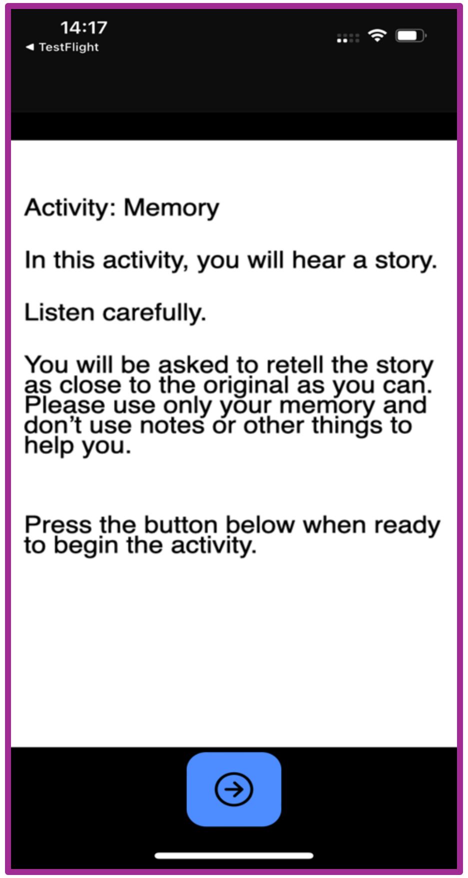
PSYCHIATRIC SYMPTOMS (DSM5 Cross Cutting Symptom Measure)

| | | | | | | |
|---|---|---|---|---|---|---|
| **Psychiatric Symptoms** | | | | | | |
| **DSM5 Cross Cutting Symptom Measure** | | | | | | |
| *“During the past two (2) weeks, how much (or how often) have you been bothered by the following problems?”* | | | | | | |
| 0 = Not at all, 1 = Less than a day or two, 2 = Several days, 3 = More than half the days, 4 = Nearly every day | | | | | | |
| 1 | Little interest or pleasure in doing things? | 0 | 1 | 2 | 3 | 4 |
| 2 | Feeling down, depressed, or hopeless? | 0 | 1 | 2 | 3 | 4 |
| 3 | Feeling more irritated, grouchy, or angry than usual? | 0 | 1 | 2 | 3 | 4 |
| 4 | Sleeping less than usual, but still have a lot of energy? | 0 | 1 | 2 | 3 | 4 |
| 5 | Starting lots more projects than usual or doing more risky things than usual? | 0 | 1 | 2 | 3 | 4 |
| 6 | Feeling nervous, anxious, frightened, worried, or on edge? | 0 | 1 | 2 | 3 | 4 |
| 7 | Feeling panic or being frightened? | 0 | 1 | 2 | 3 | 4 |
| 8 | Avoiding situations that make you anxious? | 0 | 1 | 2 | 3 | 4 |
| 9 | Unexplained aches and pains (e.g., head, back, joints, abdomen, legs)? | 0 | 1 | 2 | 3 | 4 |
| 10 | Feeling that your illnesses are not being taken seriously enough? | 0 | 1 | 2 | 3 | 4 |
| 11 | Thoughts of actually hurting yourself? | 0 | 1 | 2 | 3 | 4 |
| 12 | Hearing things other people couldn’t hear, such as voices even when no one was around? | 0 | 1 | 2 | 3 | 4 |
| 13 | Feeling that someone could hear your thoughts, ot that you could hear what another person was thinking? | 0 | 1 | 2 | 3 | 4 |
| 14 | Problems with sleep that affected your sleep quality over all? | 0 | 1 | 2 | 3 | 4 |
| 15 | Problems with memory (e.g., learning new information) or with location (e.g., finding your way home)? | 0 | 1 | 2 | 3 | 4 |
| 16 | Unpleasant thoughts, urges, or images that repeatedly enter your mind? | 0 | 1 | 2 | 3 | 4 |
| 17 | Feeling driven to perform certain behaviors or mental acts over and over again? | 0 | 1 | 2 | 3 | 4 |
| 18 | Feeling detached or distant from yourself, your body, your physical surroundings, or your memories? | | | | | |
| 19 | No knowing who you really are or what you want out of life? | 0 | 1 | 2 | 3 | 4 |
| 20 | Not feeling close to other people or enjoying your relationships with them? | 0 | 1 | 2 | 3 | 4 |
| 21 | Drinking at least 4 drinks of any kids of alcohol in a day? | 0 | 1 | 2 | 3 | 4 |
| 22 | Smoking cigarettes, a cigar, or using snuff or chewing tobacco? | 0 | 1 | 2 | 3 | 4 |
| 23 | Using any of the following medicines ON YOUR OWN, that is, without a doctor’s prescription, in greater amounts or longer than prescribed [e.g., painkillers (like Vicodin), stimulants (like Ritalin or Adderall), sedatives or tranquilizers (like sleeping pills or Valium), or drugs like marijuana, cocaine or crack, club drugs (like ecstasy), | 0 | 1 | 2 | 3 | 4 |

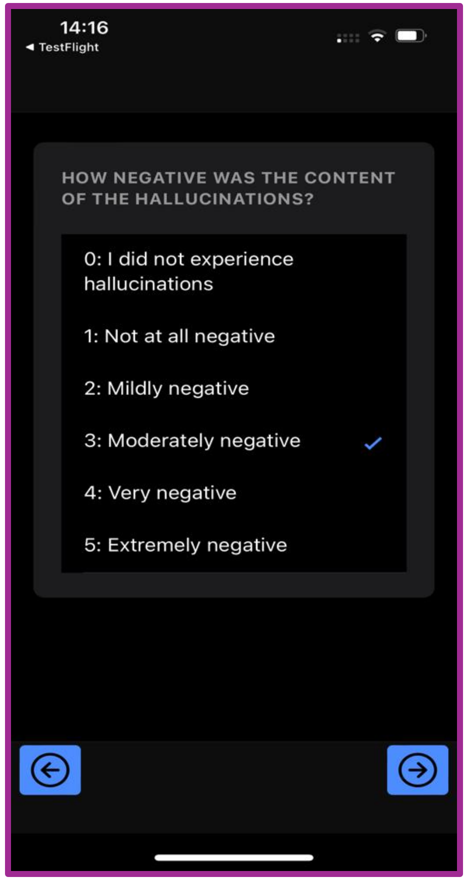
# Smartphone-based behavioral measurement



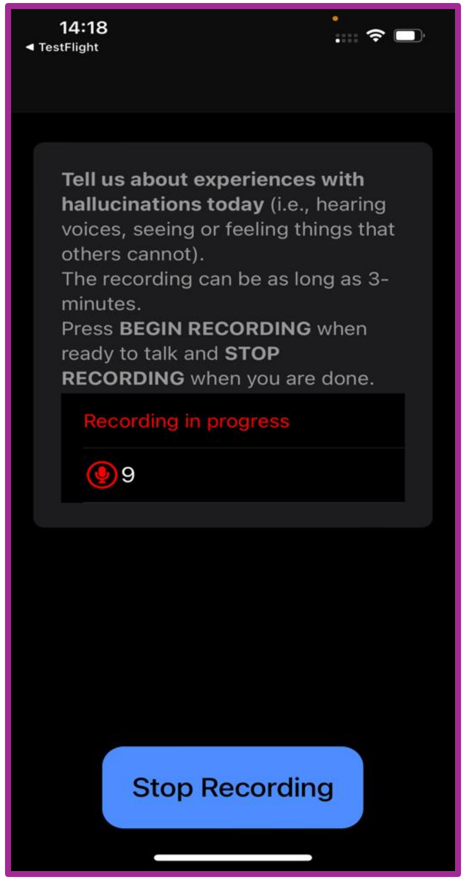

VERBAL RECALL




EMA




AUDIO DIARY



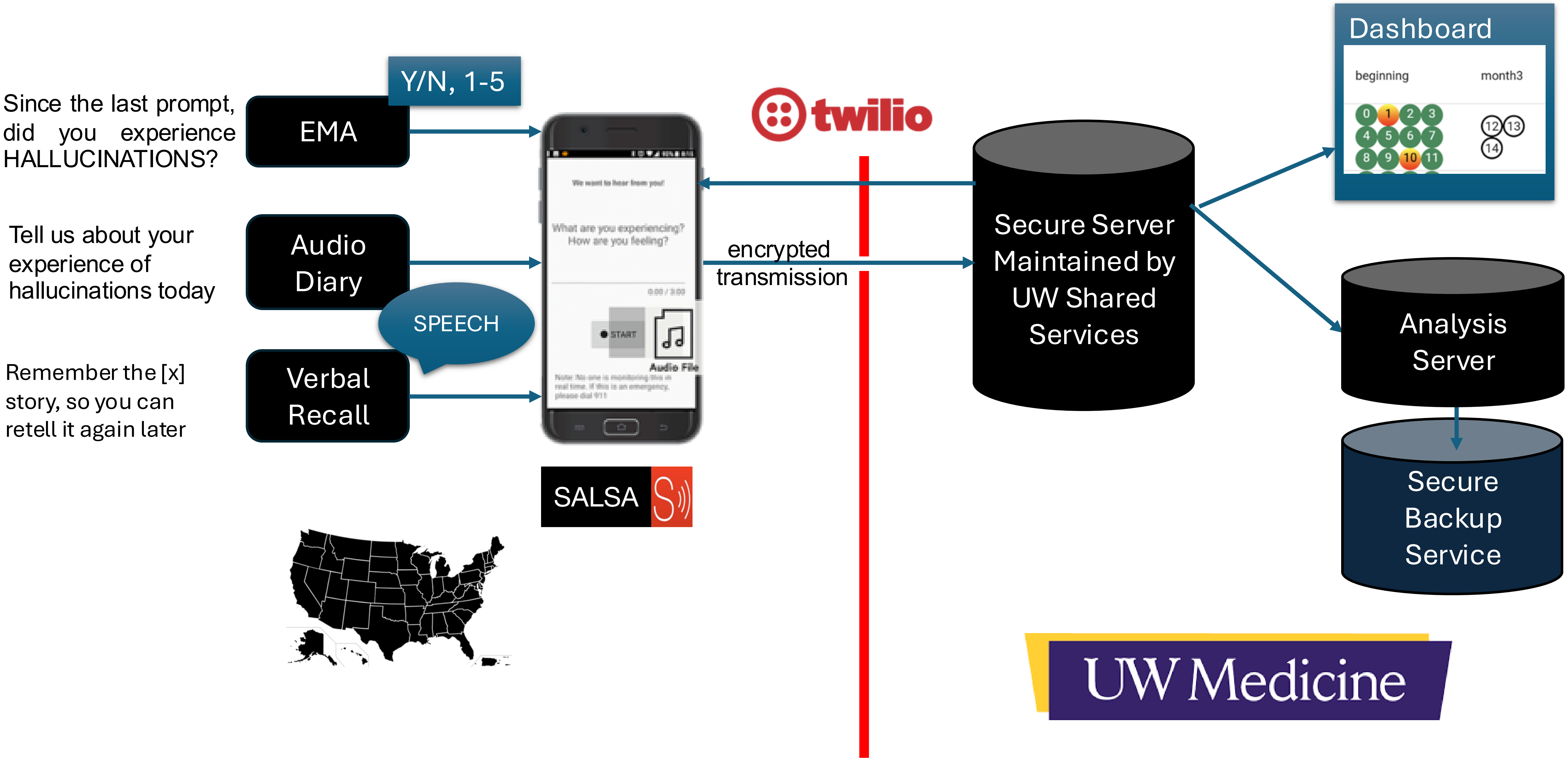

NIH

LSU
LOUISIANA STATE UNIVERSITY

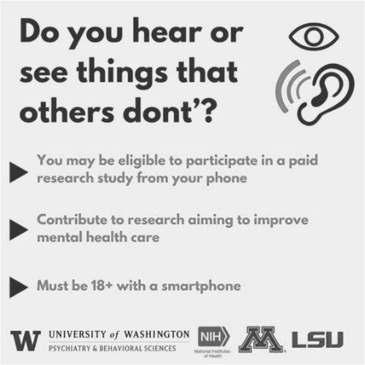
UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON

# Data collection infrastructure

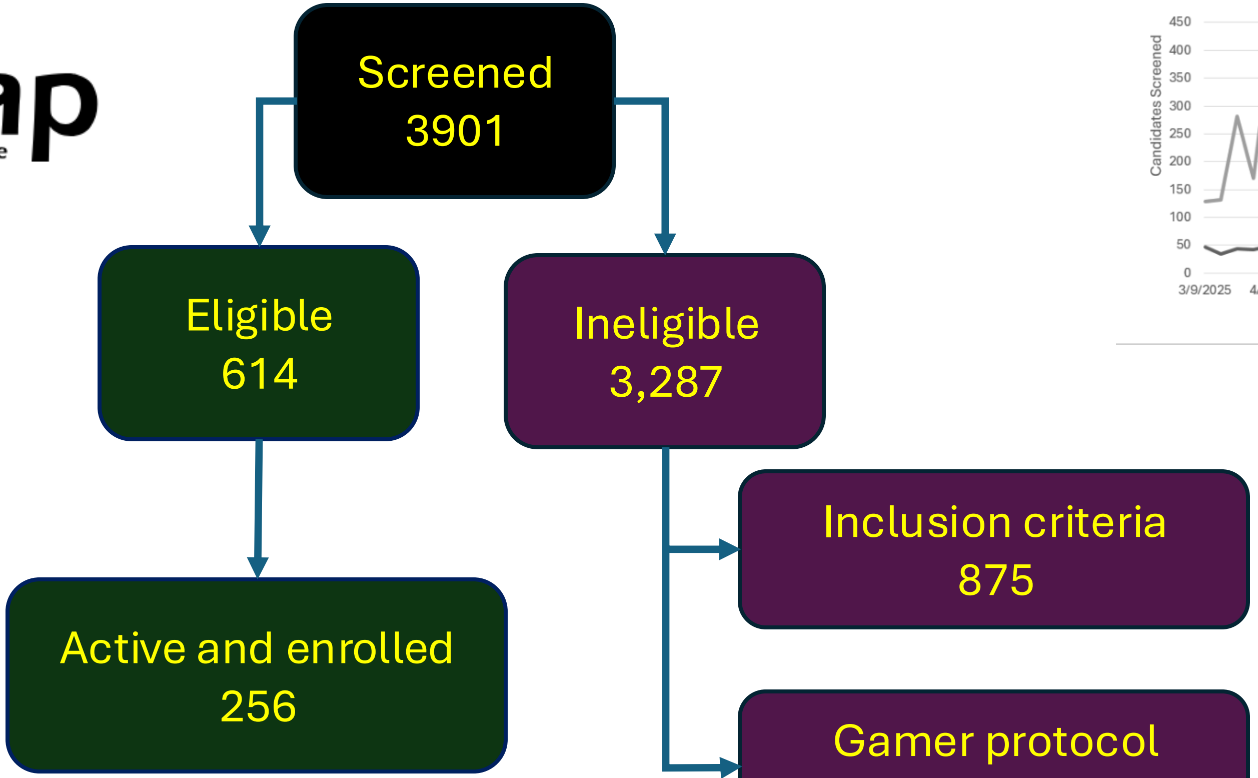


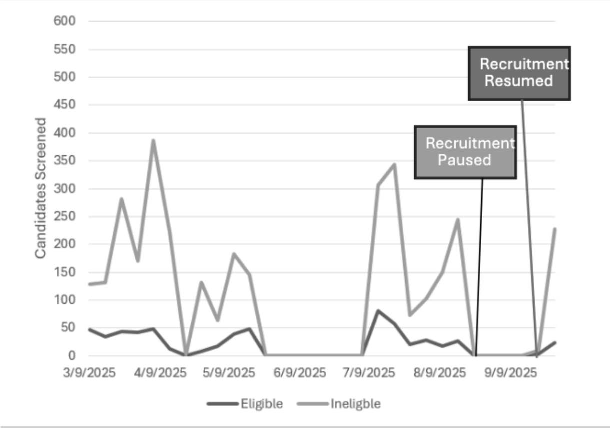
# Data to date




REDCap

Research Electronic Data Capture




NIH

LSU
LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON

# Data to date


REDCap

Screened
3901

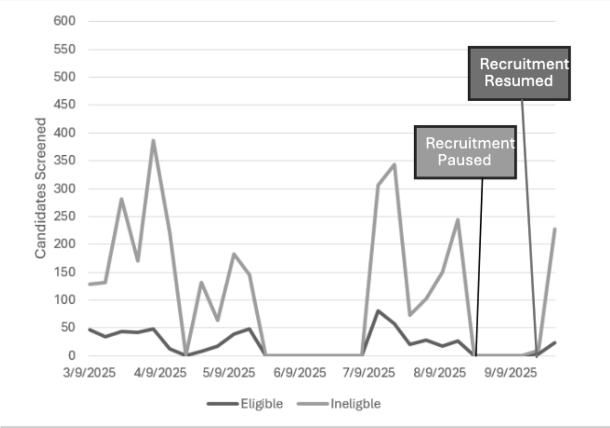
Eligible
614

Active and enrolled
256

Ineligible
3,287

Inclusion criteria
875

Gamer protocol
2,408




# Forthcoming attraction


REDCap
REDCap
Research Electronic Data Capture

Screened
3901

Eligible
614

Ineligible
3,287

Active and enrolled
256

Gamer protocol
2,408

| Gamer Check |
|---|
| Failed duplicate checks |
| Failed IP check |
| Failed time zone check |
| Failed consistency checks |
| Failed attention check |
| Failed source/link check |
| Failed RC intuition check |
| High IP risk |

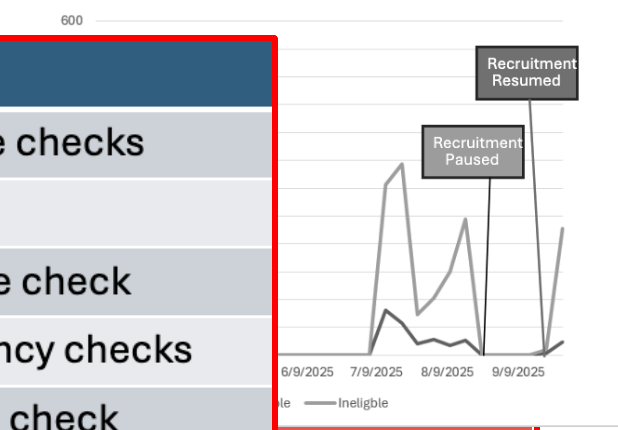
amers

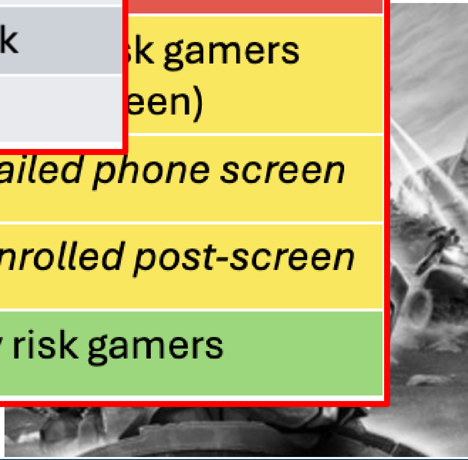
k gamers
een)

*Failed phone screen*

*Enrolled post-screen*

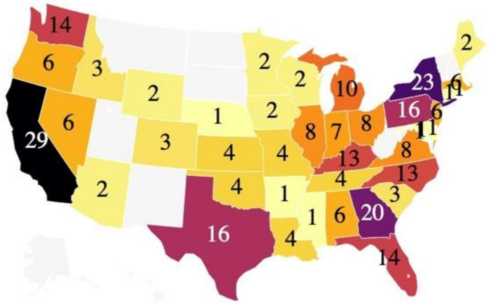
Low risk gamers



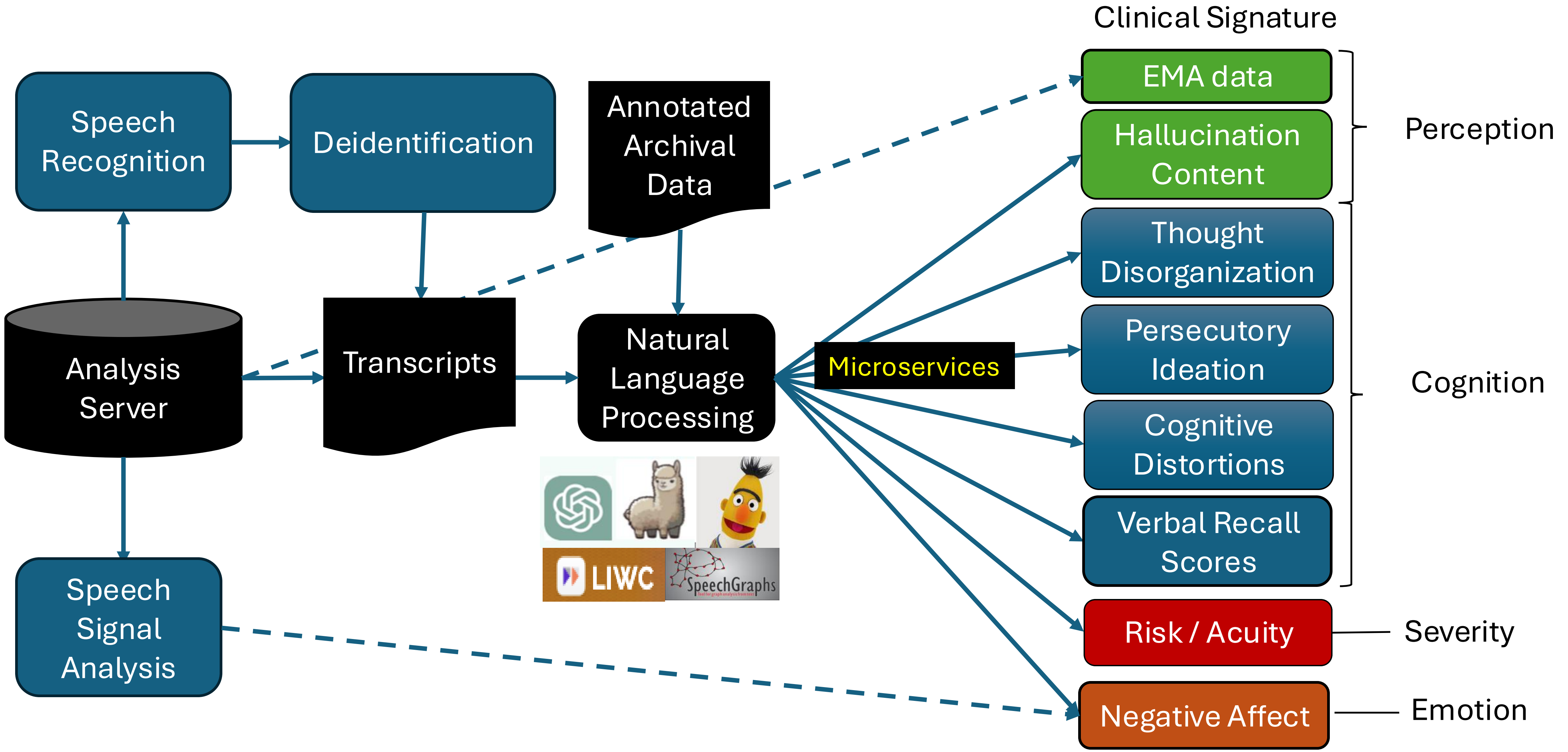



# Cohort to (slightly more recent) date

| Race | N (%) |
|---|---|
| *American Indian/Alaskan Native* | 11 (3.5%) |
| *Asian* | 14 (4.5%) |
| *Black/African American* | 104 (33.2%) |
| *Middle Eastern/North African* | 6 (1.9%) |
| *Native Hawaiian/Pacific Islander* | 5 (1.6%) |
| *White* | 190 (60.7%) |
| *Other* | 8 (2.6%) |
| *Prefer not to say/not reported* | 0 |
| **Ethnicity** | |
| *Hispanic/Latino Origin* | 50 (15.8%) |
| *Not Hispanic/Latino Origin* | 266 (84.2%) |
| *Unknown/Prefer not to say* | 0 |




| What diagnosis (or diagnoses) have you received? (Check all that apply; Among those who reported "yes" to above question) | |
|---|---|
| *Anxiety disorder* | 94 (56.3%) |
| *Obsessive-compulsive disorder* | 13 (7.8%) |
| *Bipolar disorder* | 55 (32.9%) |
| *Depressive disorder* | 56 (33.5%) |
| *Developmental or learning disorder* | 16 (9.6%) |
| *Personality disorder (e.g. borderline, paranoid, schizotypal, antisocial)* | 19 (11.4%) |
| *Schizophrenia spectrum or psychotic disorder* | 65 (38.9%) |
| *Trauma and stressor related disorder* | 50 (29.9%) |

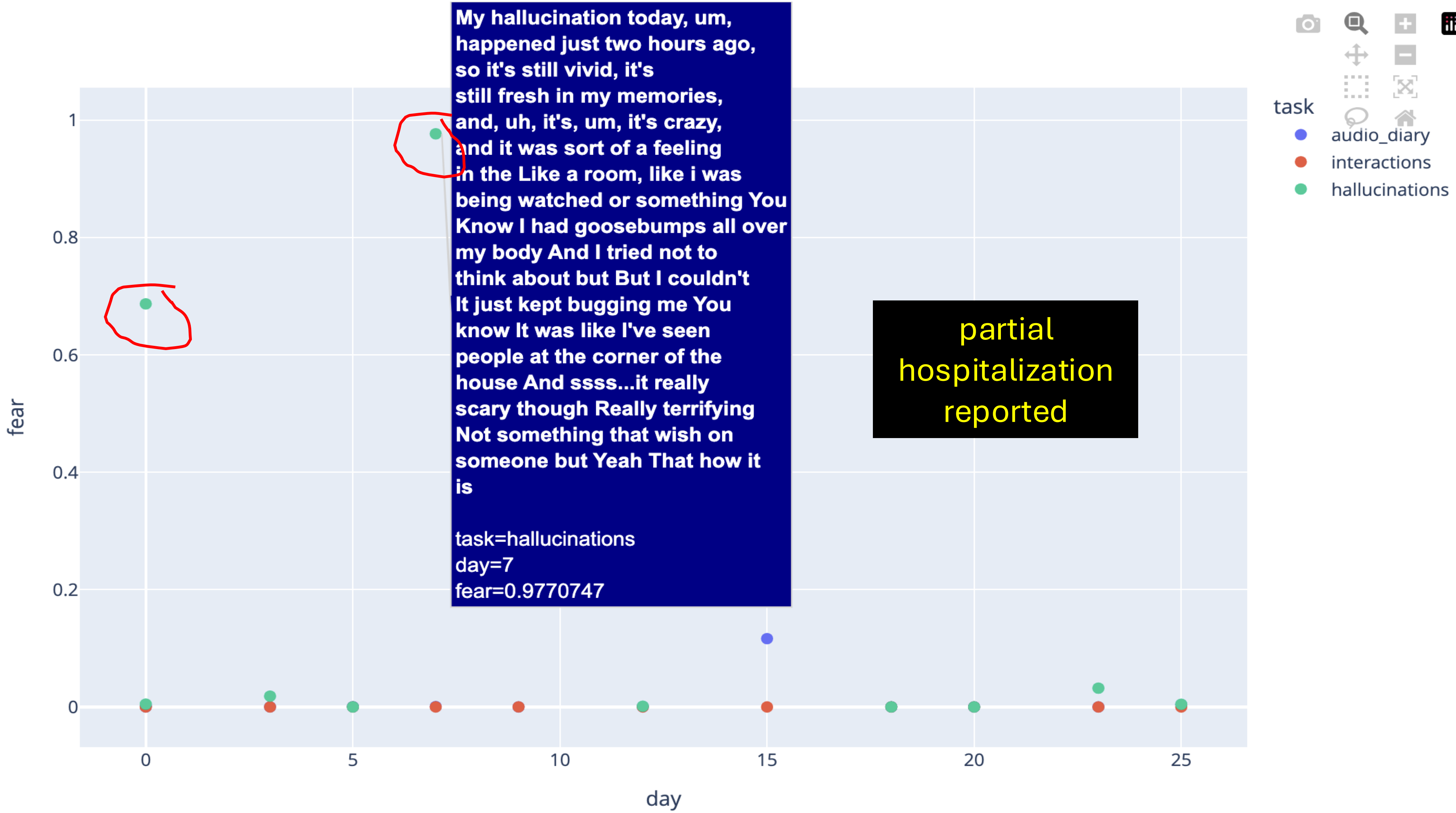



NIH

LSU
LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W
UNIVERSITY of WASHINGTON




## Clinical Signature

- EMA data
- Hallucination Content
- Thought Disorganization
- Persecutory Ideation
- Cognitive Distortions
- Verbal Recall Scores
- Risk / Acuity
- **Negative Affect**

NIH

LSU
Louisiana State University

University of Minnesota

W
University of Washington


Clinical Signature

- EMA data
- Hallucination Content
- Thought Disorganization
- Persecutory Ideation
- Cognitive Distortions
- Verbal Recall Scores
- Risk / Acuity
- Negative Affect

NIH
LSU
LOUISIANA STATE UNIVERSITY
UNIVERSITY OF MINNESOTA
W
UNIVERSITY of WASHINGTON

The memo materials have been infiltrated.

They're jumping around.

I used to be good at the broad jump, because I'm tall.

I fall.

*Excerpt From: The Center Cannot Hold: My Journey through Madness*
*By Elyn R. Saks*

## Clinical Signature

- Perception
  - EMA data
  - Hallucination Content
- Cognition
  - **Thought Disorganization**
  - Persecutory Ideation
  - Cognitive Distortions
  - Verbal Recall Scores
- Severity
  - Risk / Acuity
- Emotion
  - Negative Affect


Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. Schizophrenia research. 2007 Jul 1;93(1-3):304-16.

Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, Ribeiro S, Javitt DC, Copelli M, Corcoran CM. Automated analysis of free speech predicts psychosis onset in high-risk youths. npj Schizophrenia. 2015 Aug 26;1(1):1-7..

Journal of Biomedical Informatics 170 (2025) 104899

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

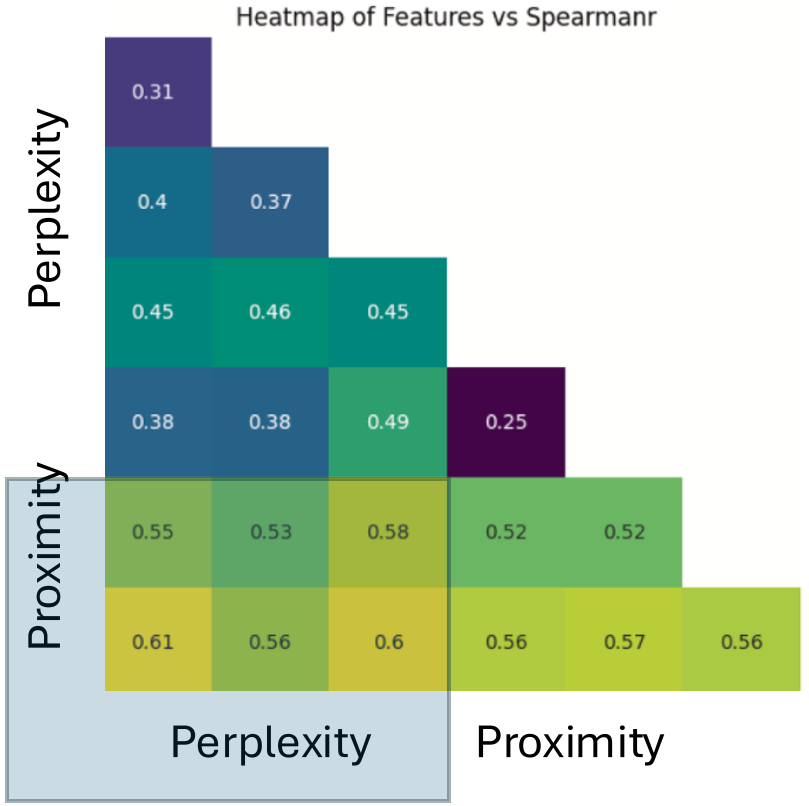
journal homepage: www.elsevier.com/locate/yjbin

ELSEVIER

Original Research

# Perplexity and proximity: Large language model perplexity complements semantic distance metrics for the detection of incoherent speech

Weizhe Xu [a,*], Serguei Pakhomov [b], Patrick Heagerty [c], Eric Horvitz [d], Ellen R. Bradley [e], Josh Woolley [e], Andrew Campbell [f], Alex Cohen [g], Dror Ben-Zeev [h], Trevor Cohen [a,h]



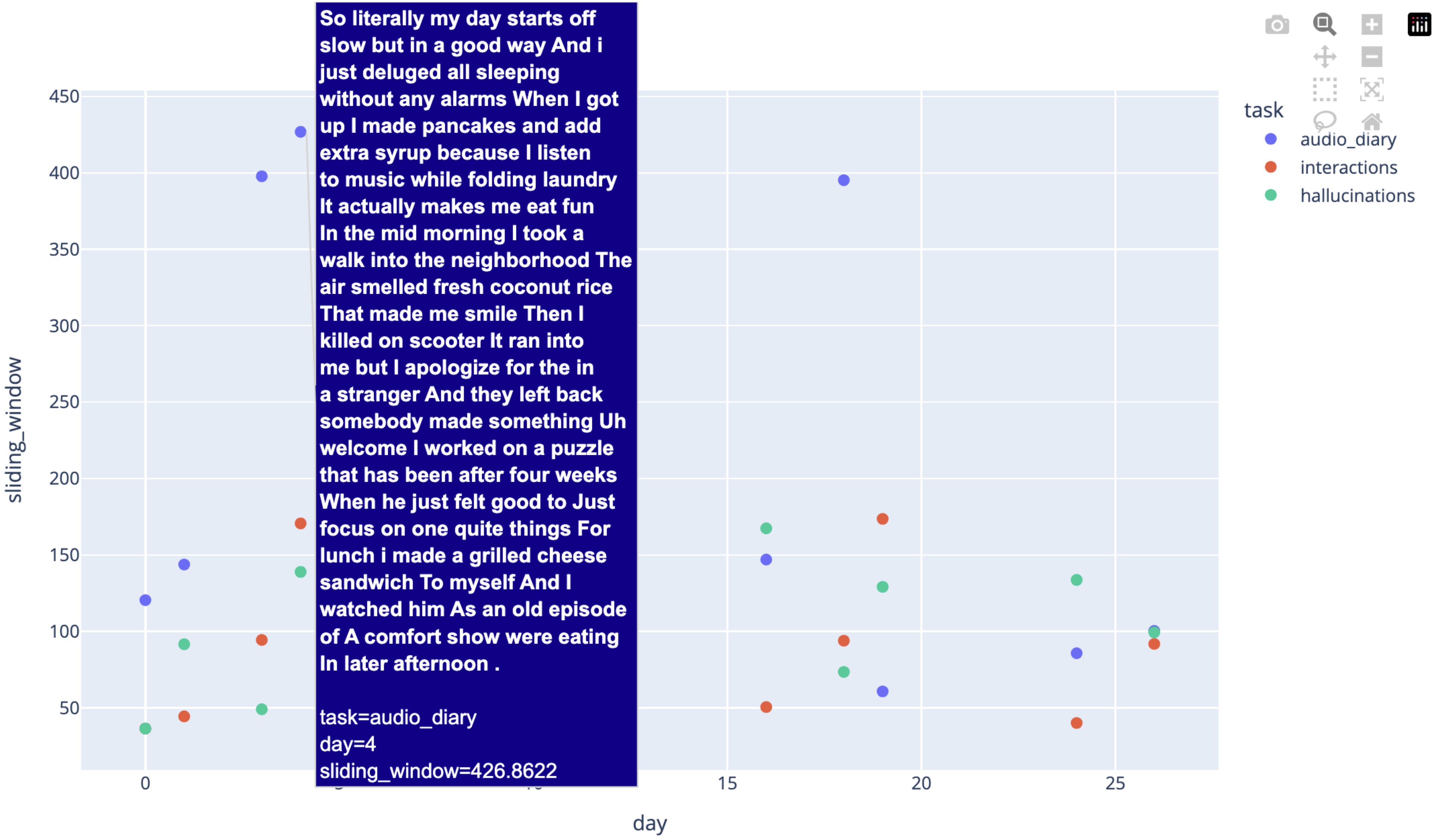

CLPsych
Workshop on Computational Linguistics and Clinical Psychology

# Bigger But Not Better: Small Neural Language Models Outperform Large Language Models in Detection of Thought Disorder

Changye Li [1*] Weizhe Xu [1*] Serguei Pakhomov [2]
Ellen Bradley [3] Dror Ben-Zeev [1] Trevor Cohen [1]

| Model | Sliding windows | | | | |
|---|---|---|---|---|---|
| | 8 | 16 | 32 | 64 | 128 |
| 70m | 0.265 | **0.482***** | 0.338** | 0.356** | 0.344** |
| 160m | 0.414*** | 0.433*** | 0.316* | 0.354*** | 0.325** |
| 410m | 0.385* | 0.433*** | 0.316* | 0.354** | 0.325** |
| 1b | 0.415*** | 0.352** | 0.380** | 0.410*** | 0.313* |
| 1.4b | 0.458*** | 0.370*** | 0.382** | 0.418*** | 0.326** |
| 2.8b | 0.425*** | 0.385** | 0.369** | **0.428***** | 0.348*** |
| 6.9b | **0.478***** | 0.352** | 0.394** | 0.414*** | **0.368**** |
| 12b | 0.441*** | 0.313* | **0.404**** | 0.412*** | 0.357* |
| LLaMA | – | – | – | 0.249 | – |

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$




## Clinical Signature

EMA data

Hallucination Content

**Thought Disorganization**

Persecutory Ideation

Cognitive Distortions

Verbal Recall Scores

Risk / Acuity

Negative Affect

NIH

LSU

Louisiana State University

University of Minnesota

W

University of Washington

# Why so perplexed?

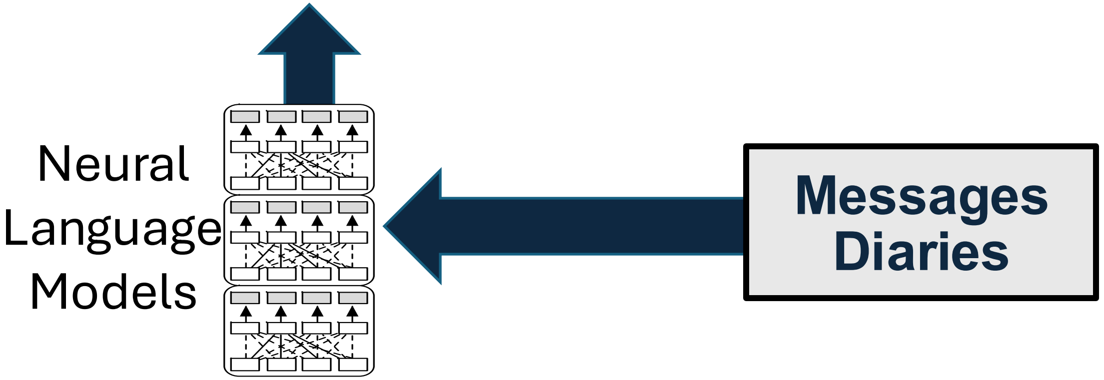
**Our objective is to leverage novel data sets and cutting-edge computational methods to generate clinical signatures that can be used to inform scalable and timely risk detection and clinical decision making.**

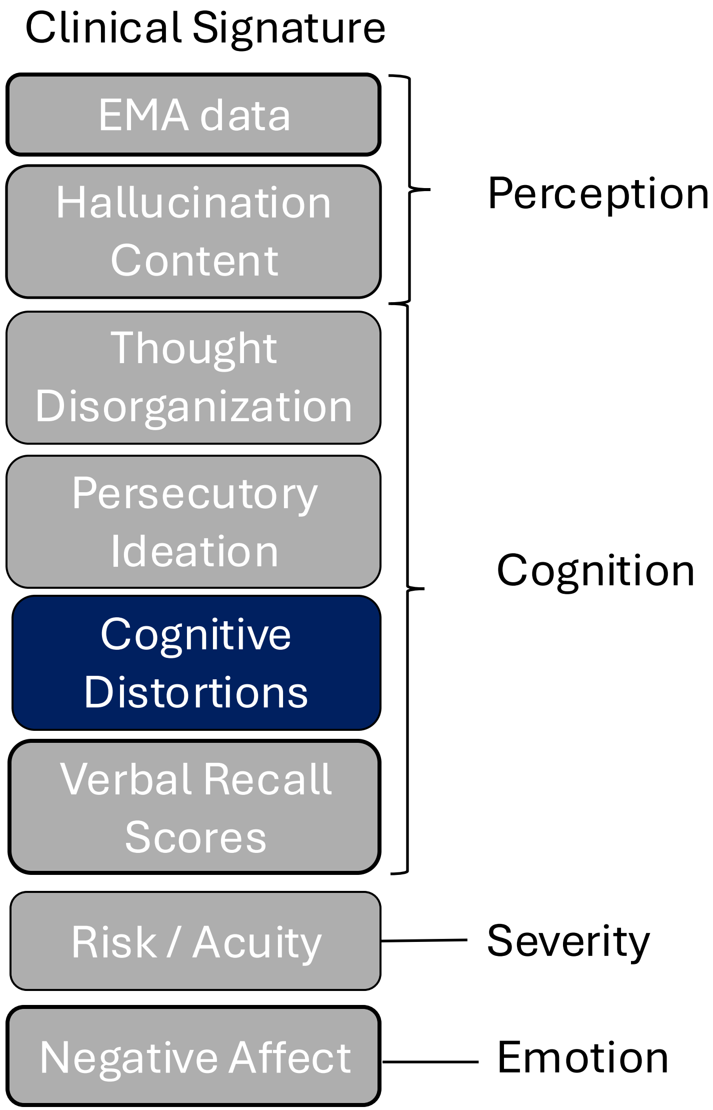
Specifically, we aim to:

1) **Derive data-driven clinical signatures** from mobile behavioral tasks to predict individual differences in severe negative outcomes among people experiencing hallucinations;
2) **Identify and mitigate inaccuracies** in modelling across groups;
3) Examine whether **adding smartphone-captured behavioral data** to information that is typically available in the clinical record **improves model clinical utility**; and
4) Produce **machine learning-ready data structures** that adhere to FAIR (Findable, Accessible, Interoperable, Reusable) principles.

NIH LSU LOUISIANA STATE UNIVERSITY UNIVERSITY OF MINNESOTA UNIVERSITY of WASHINGTON

| Distortion | SHAP Explainer |
|---|---|
| mental filter | Good morning how are you? I couldn't fall asleep last night until 1am. And this morning when I got up I got very dizzy almost like I was going to pass out. I passed out once when I was 11. I remember how I felt right before and that is how I felt this morning. It was very disconcerting, so I don't think I'm doing to do much today. |
| jumping to conclusions | Maybe they really don't like me. |
| catastrophizing | Feeling unsure of myself right now. Desperate. Idk if I can handle my money or tackle my goals of saving for a rainy day. |





**Tauscher JS** et al. Automated Flagging of Cognitive Biases in the Spoken Language of People Hallucination Experiences. Journal of Technology in Behavioral Science. 2025 Sep 11:1-0./

**Tauscher JS** et al. Automated detection of cognitive distortions in text exchanges between clinicians and people with serious mental illness. Psychiatric services. 2023. https://pubmed.ncbi.nlm.nih.gov/36164769/


NIH LSU LOUISIANA STATE UNIVERSITY UNIVERSITY OF MINNESOTA W UNIVERSITY of WASHINGTON

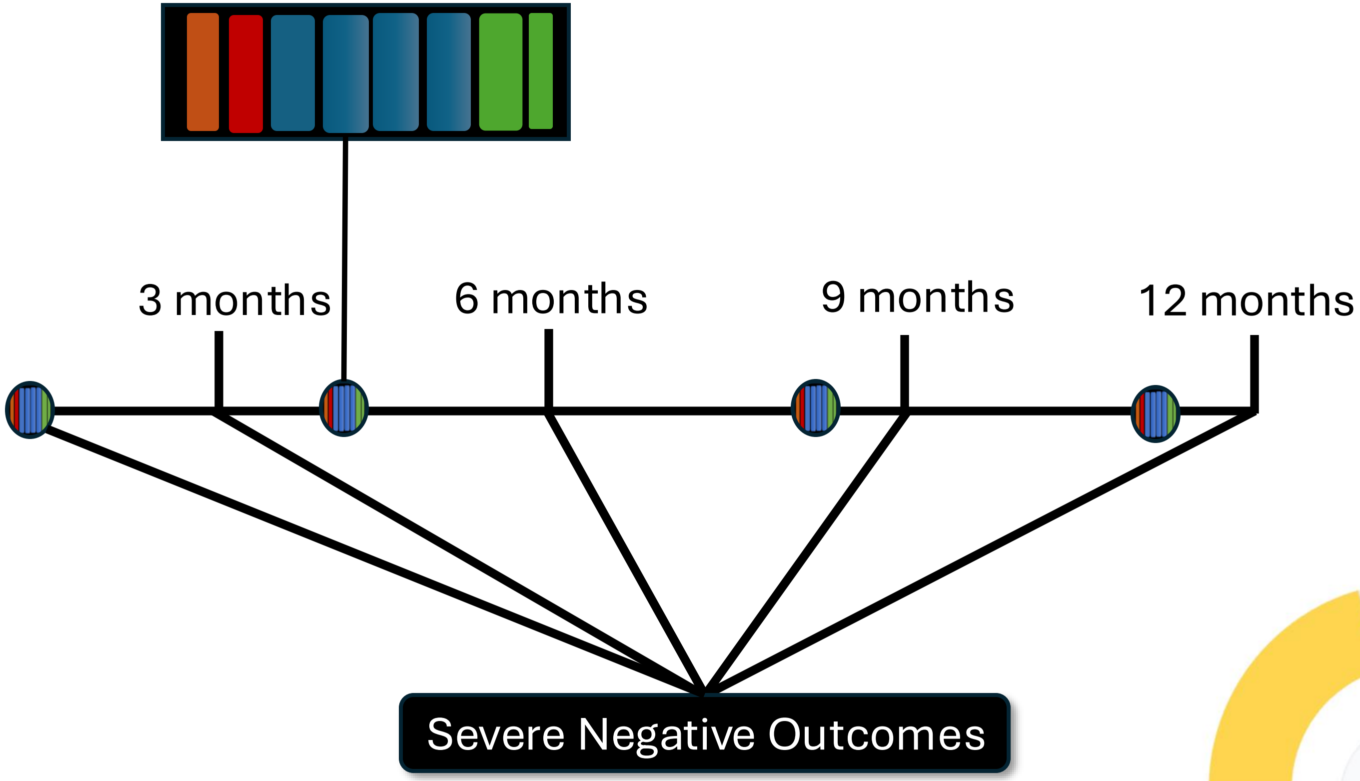

Clinical Signature

- EMA data
- Hallucination Content
- Thought Disorganization
- Persecutory Ideation
- Cognitive Distortions
- Verbal Recall Scores
- Risk / Acuity
- Negative Affect

# Computational psychiatry in the wild?

**Clinical Signature**

- EMA data
- Hallucination Content
- Thought Disorganization
- Persecutory Ideation
- Cognitive Distortions
- Verbal Recall Scores
- Risk / Acuity
- Negative Affect




Psychiatric symptoms
Functional impairment


NIH
LSU
Louisiana State University
University of Minnesota
W
University of Washington

# Acknowledgements

- Now:
  - NIMH, NLM, IMPACT-MH, Garvey Institute for Brain Health Solutions
- Then:
  - NIMH, NLM, NIA, NIH Common Fund
  - Garvey Institute for Brain Health Solutions


NIH

LSU

LOUISIANA STATE UNIVERSITY

UNIVERSITY OF MINNESOTA

W

UNIVERSITY of WASHINGTON