camh
Centre for Addiction and Mental Health

Krembil Centre for Neuroinformatics

Using big data, artificial intelligence and brain modelling to fundamentally change our understanding of mental illness.

SUMMER SCHOOL 2021
Day 7

Digital Health and Population-based Data Resources

CAMH Land Acknowledgement

CAMH is situated on lands that have been occupied by First Nations for millennia; lands rich in civilizations with knowledge of medicine, architecture, technology, and extensive trade routes throughout the Americas. In 1860, the site of CAMH appeared in the Colonial Records Office of British Crown as the council grounds of the Mississaugas of the New Credit, as they were known at the time.

Today, Toronto is covered by the Toronto Purchase, treaty No. 13 of 1805 with the Mississaugas of the Credit.

Toronto is now home to a vast diversity of First Nations, Inuit, and Métis who enrich this city.

CAMH is committed to reconciliation. We will honour the land through programs and places that reflect and respect its heritage. We will embrace the healing traditions of the Ancestors, and weave them into our caring practices. We will create new relationships and partnerships with First Nations, Inuit, and Métis and share the land and protect it for future generations.

camh | Krembil Centre for Neuroinformatics

camh Shkaabe Makwa

Summer School Schedule

camh | Krembil Centre for Neuroinformatics

Mon 05/07

Understanding clinical research questions and reproducible science.

S. Hill & E. Dickie
B. Jones & V. Tang

Tu 06/07

D. Buchman, L. Sikstrom & M Maslej

Applied ethics in machine learning and mental health

Wed 07/07

Genetics and transcriptomics

S.Tripathy, & D. Felsky

Wed 08/07

E. Hay & F. Mazza

Brain Microcircuit Simulations of Depression

Fri 09/07

Whole-Brain Modelling and Neuroimaging Connectomics

J. Griffiths & E Dickie

Mon 12/07

A Diaconescu + Lab

Bayesian Models of Learning and Integration of Neuroimaging Data

Tues 13/07

Day 7

Digital Health and Population-based data resources

A. Pratap, J Yu & D. Felsky

Wed 14/07

A. Pratap & D. Felsky
Group Panel

Integrative research methods
Panel Discussion

Remember - many ways to engage

camh | Krembil Centre for Neuroinformatics

crowdcast

(during sessions) Use the chat or the ask question!

You can always return to the session and re-watch the vidos after the session ends

slack

come chat with us in KCNI Summer School Slack :)

virtually meet with us in gather.town

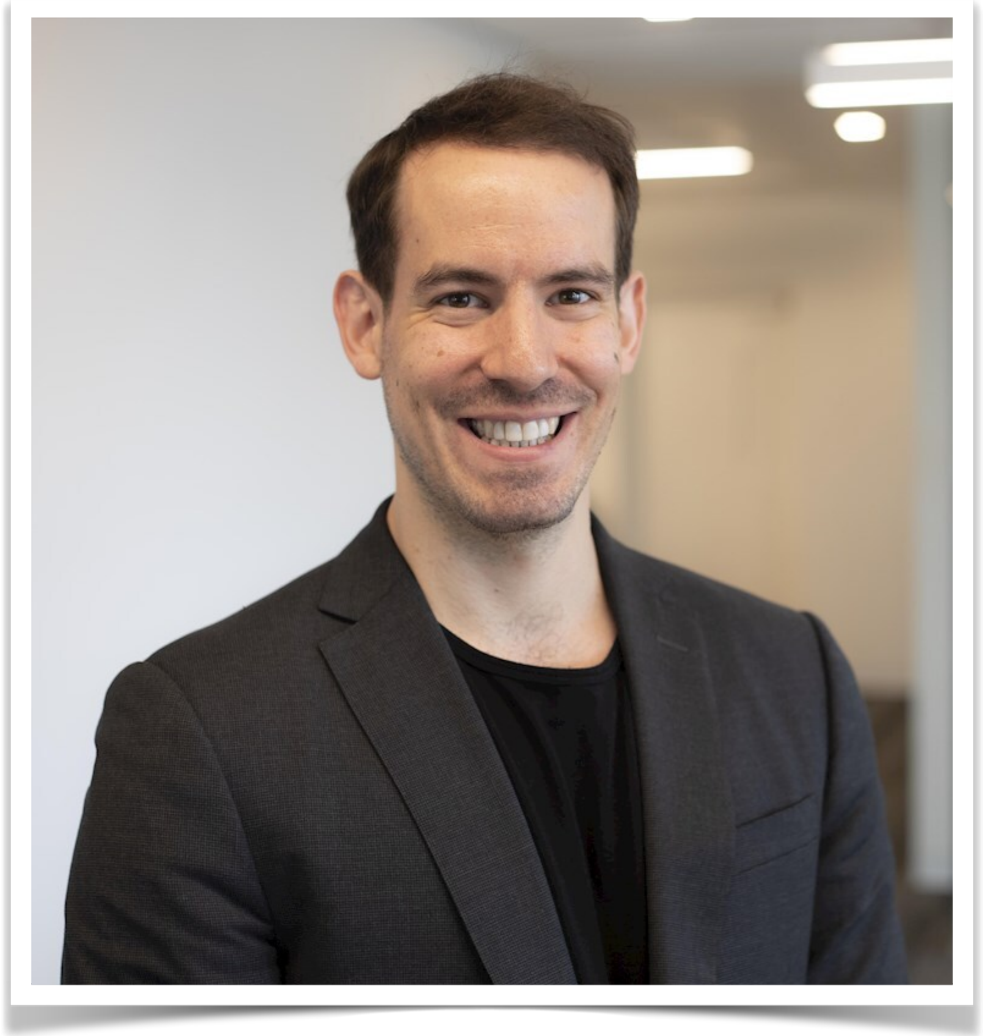
Gather

KCNISchool@camh.ca

Instructors for today

camh | Krembil Centre for Neuroinformatics

Joanna Yu, PhD
Team Lead, Brain Health Databank
KCNI



Dan Felsky, PhD
Lab Head | Independent Scientist, KCNI
Assistant Professor of Psychiatry,
Associate Member, Institute of Medical Science,
University of Toronto.

Abhi Pratap, PhD
Lab Head - Digital Health & AI
Independent Scientist, KCNI
Faculty Affiliate, Vector Institute

Today's Agenda

camh
Krembil Centre for Neuroinformatics

Day 7

Digital Health
&
Population-based data resources

9:00 am - 10:30 am	Digital Health for Mental Health - Opportunities & Challenges Dr. Abhi Pratap
10:45 am- 12:15 pm	Population-based resources and the BrainHealth Databank Drs. Daniel Felsky, Joanna Yu & Abhi Pratap
1:00 pm - 2:30 pm	Workshop/Demo: Reproducible analysis using Synapse as part of an integrated workflow Dr. Abhi Pratap
2:45 pm - 4:15 pm	Workshop: Introduction to interactive methods Dr. Daniel Felsky

9:00 am - 10:30 am

Digital Health for Mental Health - Opportunities & Challenges

Dr. Abhi Pratap

Mental Health

Technology

Why

How

When

Why

Digital Health for Mental Health

Opportunities

How

Using digital health to assess CNS symptoms
“in the real world”

Feasibility & Predictability

When

If we build tech, communities will embrace it

Challenges & Solutions

Why

Digital Health for Mental Health

Opportunities

How

Using digital health to assess CNS symptoms
“in the real world”

Feasibility & Predictability

When

If we build tech, communities will embrace it

Challenges & Solutions

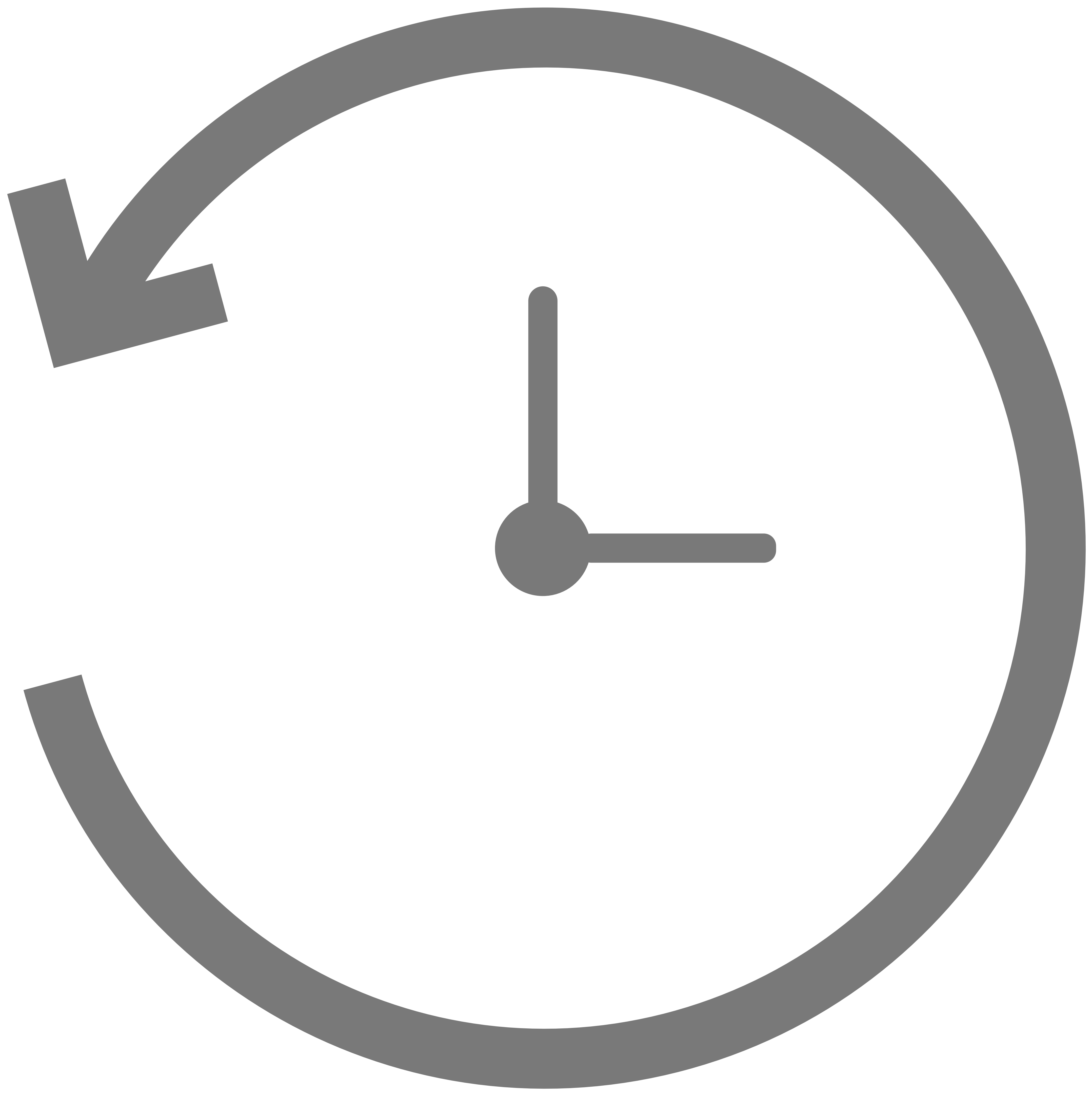
Fresh table setting with leaves on dark plates, with fruit on wooden table by Anna Ivanova from Noun Project

"...... *trying to understand another human being's emotional life is fraught with potential error* .....

.....*one's own prejudices and needs* ....

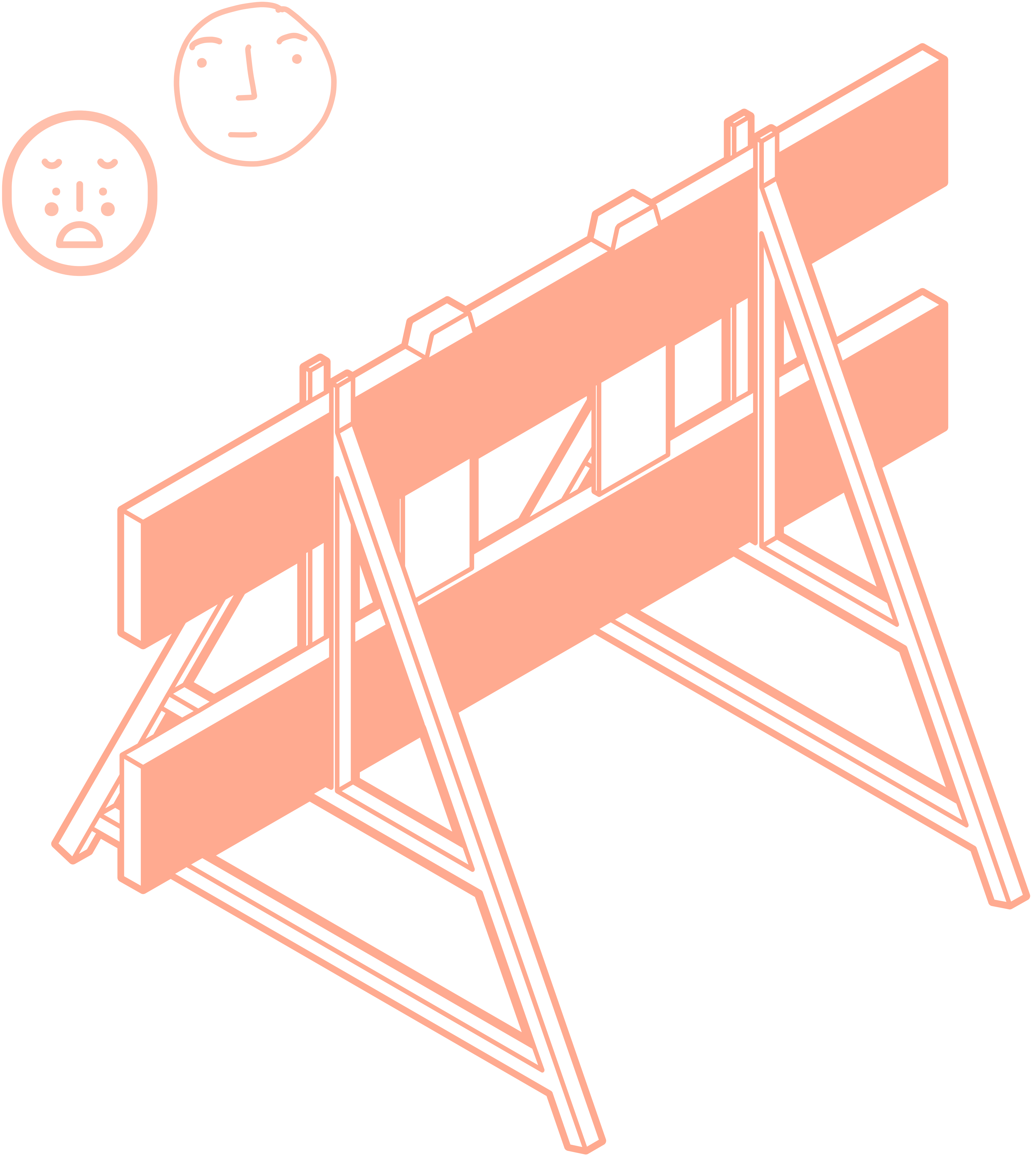
.......*we have no objective yardstick for this confidence*."

– Emil Kraepelin, The Manifestations of Insanity, 1920



Most people with mental illness are not able to get minimally adequate and timely care

https://www.nature.com/polopoly_fs/1.19694!/menu/main/topColumns/topLeftColumn/pdf/532020a.pdf?origin=ppub



Self-realization / acceptance

Stigma

Privacy / Trust Concerns

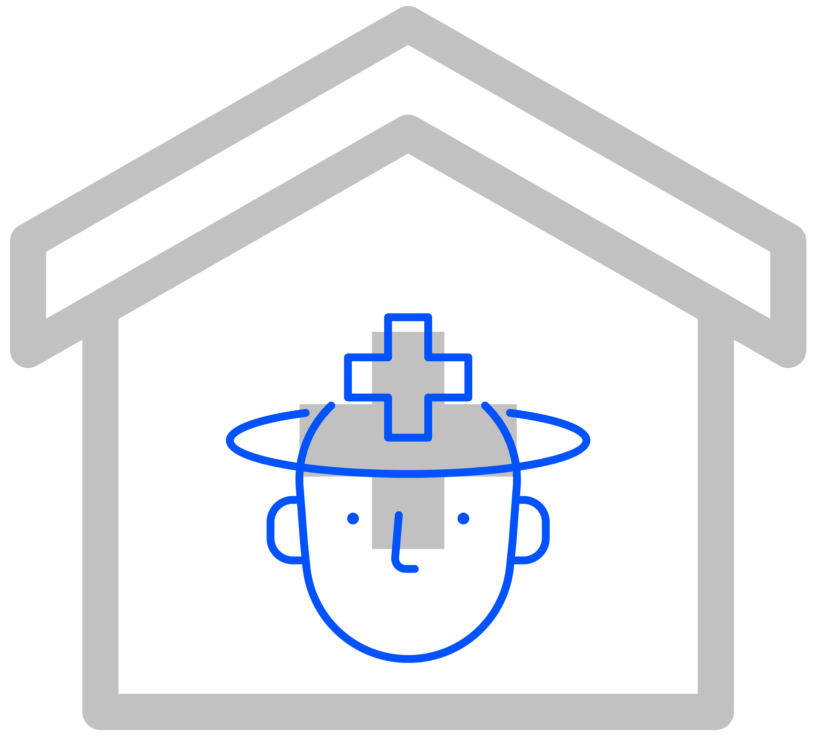
Cost

Wait Time

Variabile experience

Language/Cultural

……. and many more……..



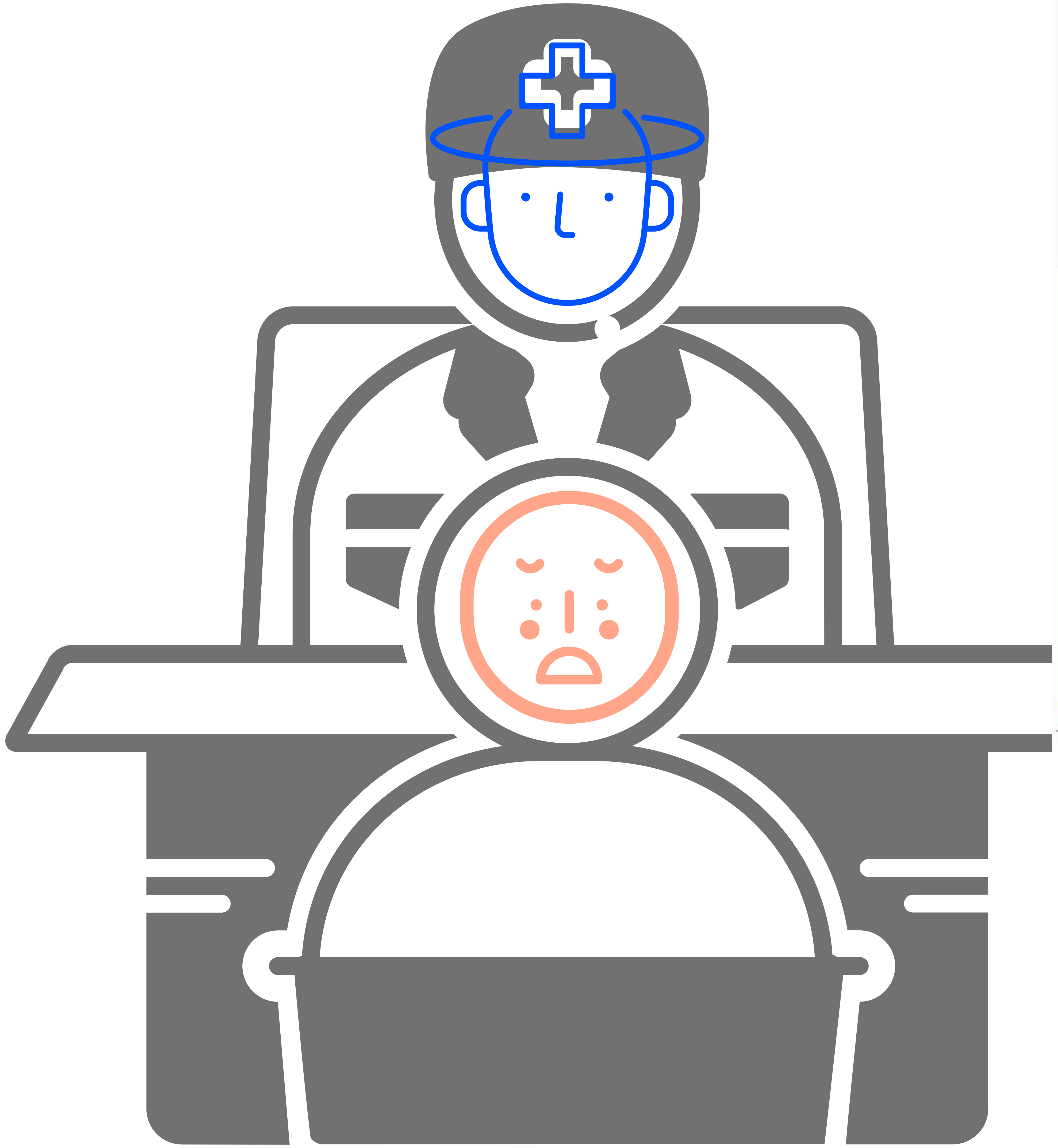


TABLE 1 DSM-5 criteria for major depressive disorder and persistent depressive disorder

Major depressive disorder (in children and adolescents, mood can be irritable)

5 or more of 9 symptoms (including at least 1 of depressed mood and loss of interest or pleasure) in the same 2-week period; each of these symptoms represents a change from previous functioning

- Depressed mood (subjective or observed)
- Loss of interest or pleasure
- Change in weight or appetite
- Insomnia or hypersomnia
- Psychomotor retardation or agitation (observed)
- Loss of energy or fatigue
- Worthlessness or guilt
- Impaired concentration or indecisiveness
- Thoughts of death or suicidal ideation or suicide attempt

Persistent depressive disorder (in children and adolescents, mood can be irritable and duration must be 1 year or longer)

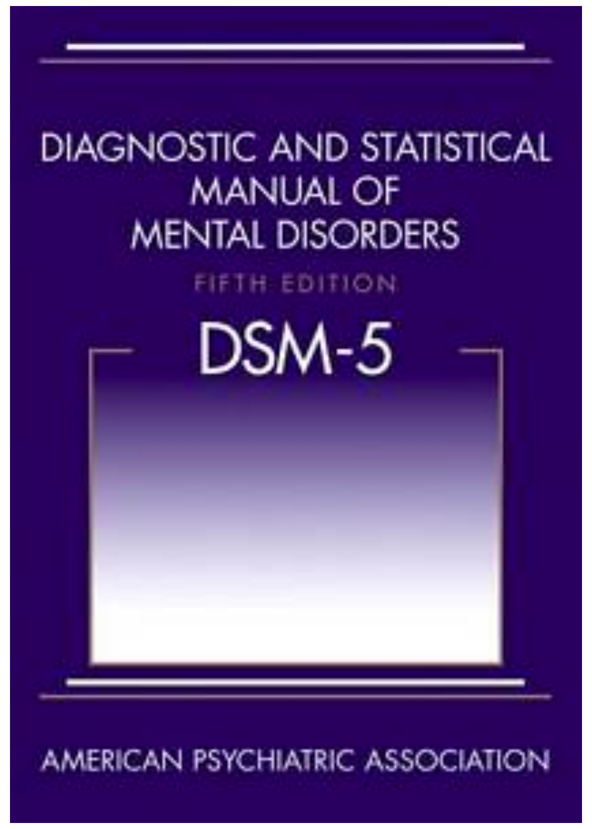
Depressed mood for most of the day, for more days than not, for 2 years or longer

Presence of 2 or more of the following during the same period

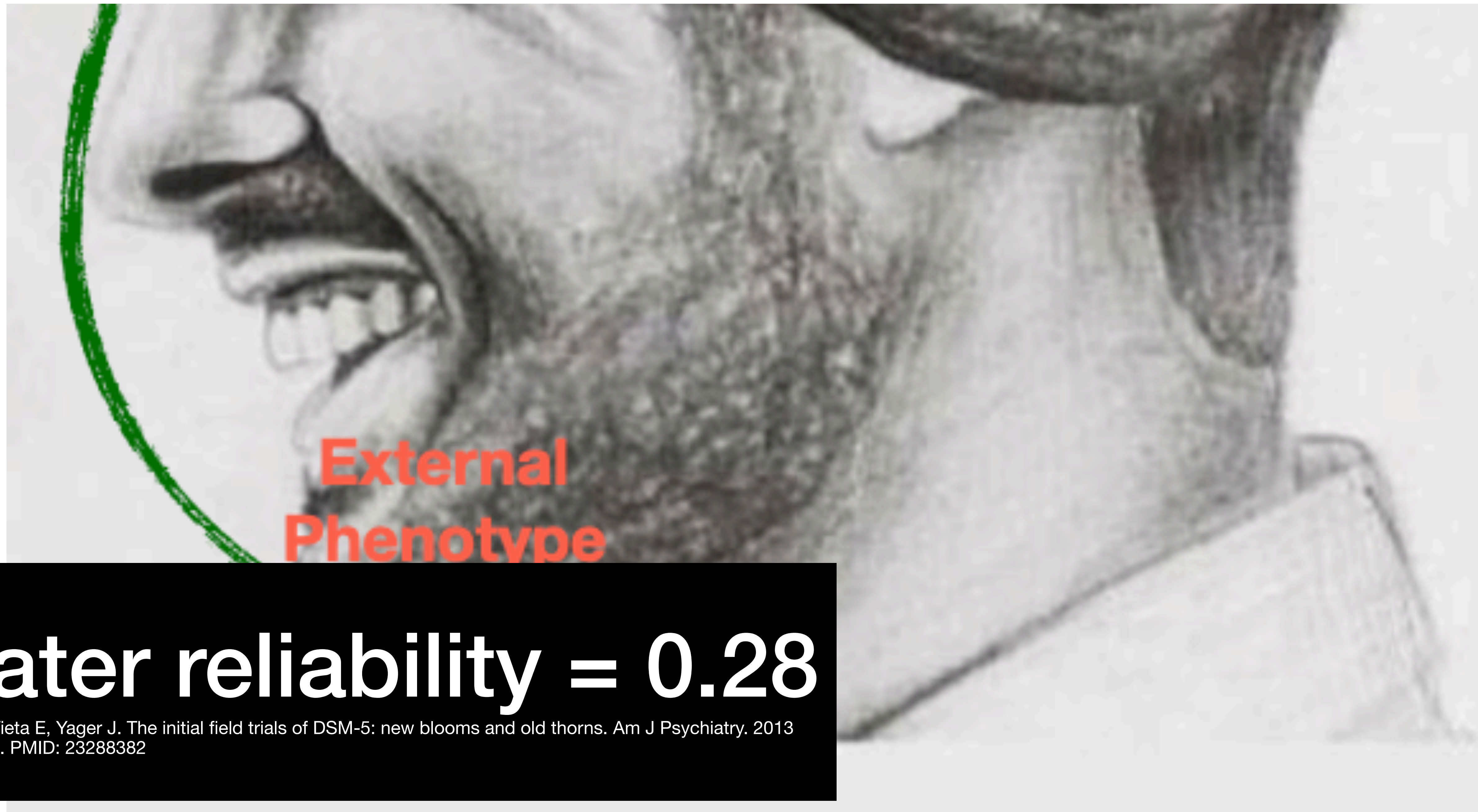
- Poor appetite or overeating
- Insomnia or hypersomnia
- Low energy or fatigue
- Low self-esteem
- Impaired concentration or indecisiveness
- Hopelessness

Never without symptoms for more than 2 months

Table 1: DSM-5 criteria for major depressive disorder and persistent d...



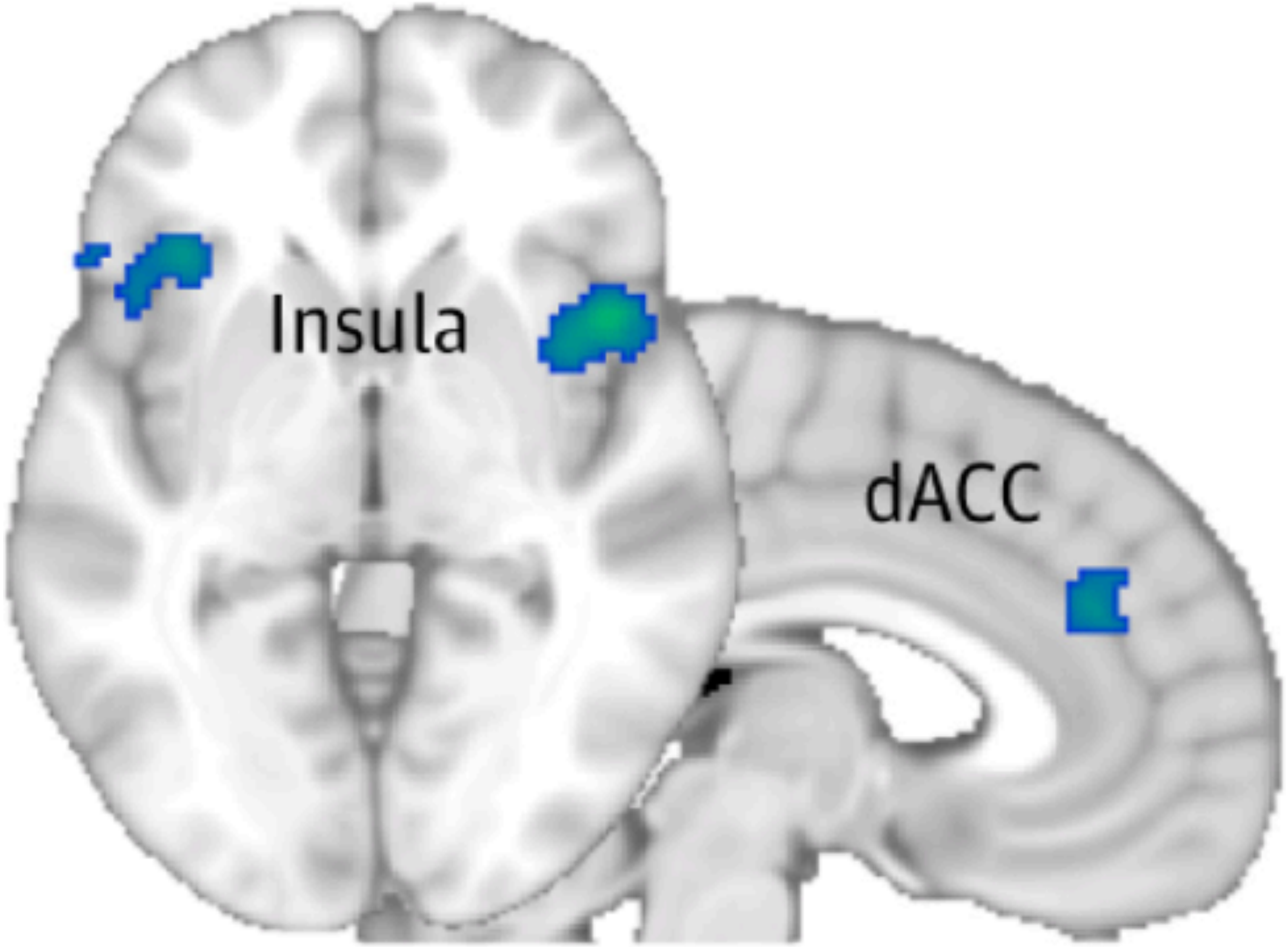
Clinical assessment / diagnosis of mental health disorders can be imprecise

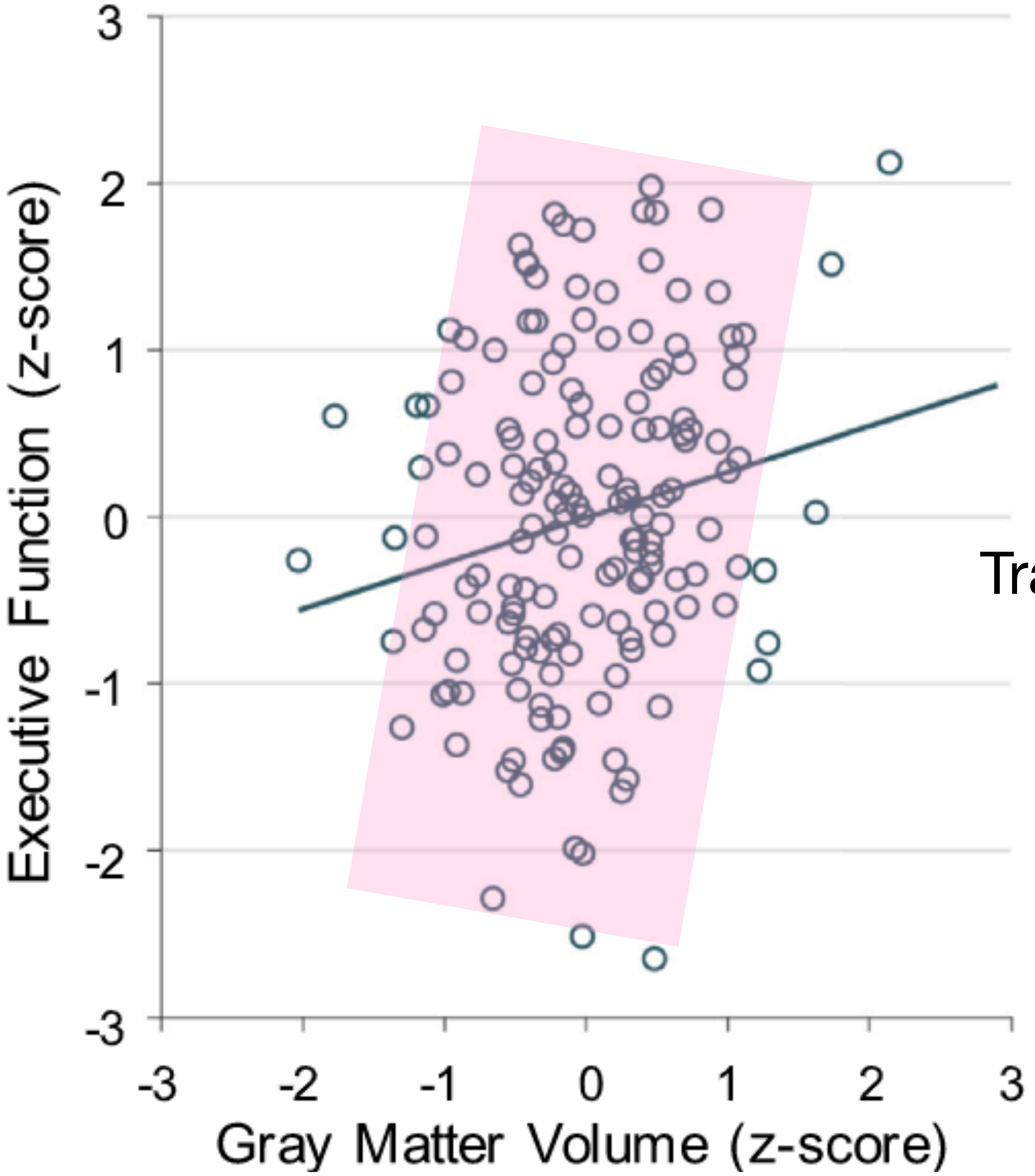


Clinician to clinician inter-rater reliability = 0.28

Freedman R, Lewis DA, Michels R, Pine DS, Schultz SK, Tamminga CA, Gabbard GO, Gau SS, Javitt DC, Oquendo MA, Shrout PE, Vieta E, Yager J. The initial field trials of DSM-5: new blooms and old thorns. Am J Psychiatry. 2013 Jan;170(1):1-5. doi: 10.1176/appi.ajp.2012.12091189. PMID: 23288382

Neuroimaging & Liquid CNS Biomarkers





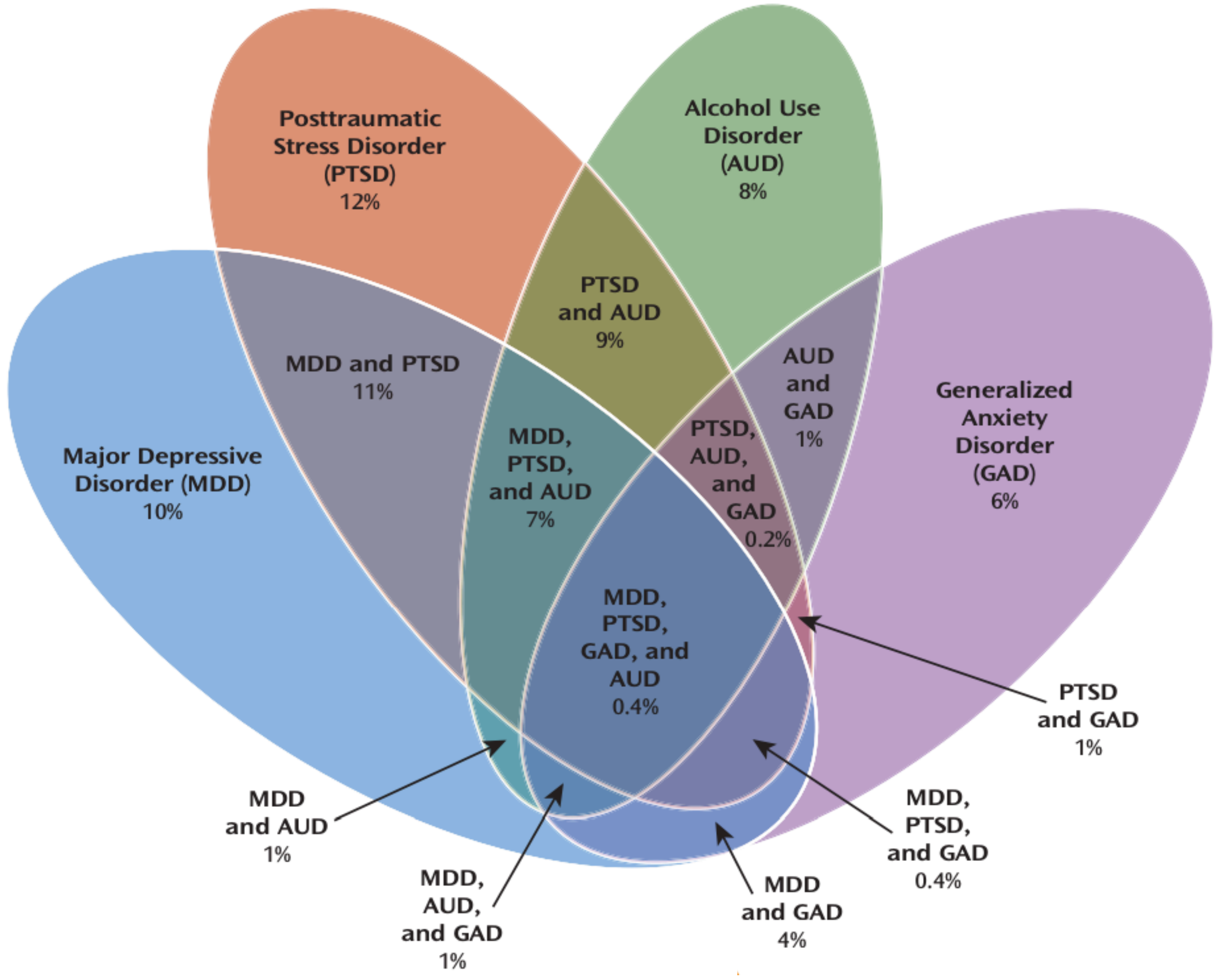
Transdiagnostic Variability

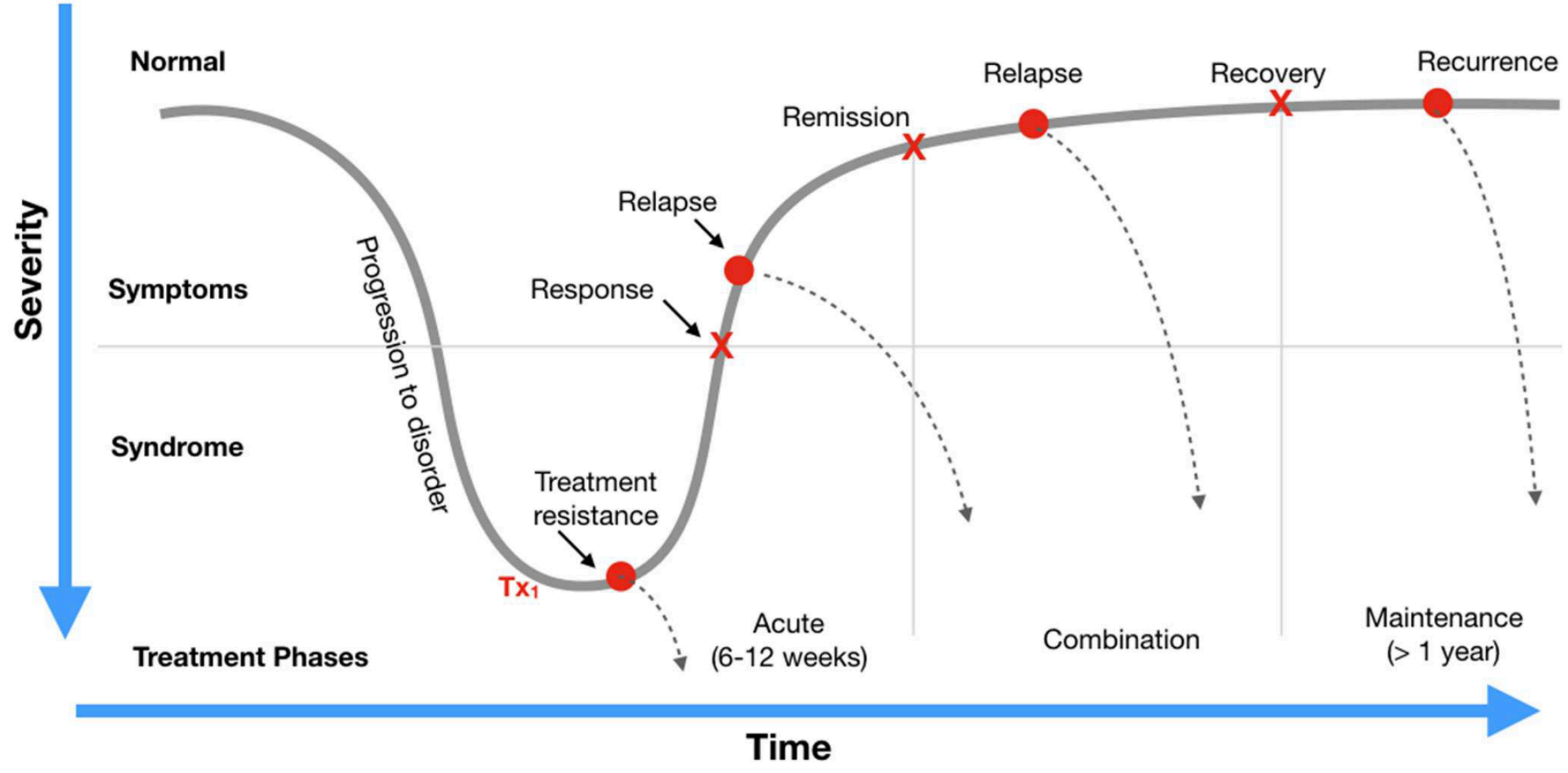
Lynch CJ, Gunning FM, Liston C. Causes and Consequences of Diagnostic Heterogeneity in Depression: Paths to Discovering Novel Biological Depression Subtypes. Biol Psychiatry. 2020 Jul 1;88(1):83-94. doi: 10.1016/j.biopsych.2020.01.012. Epub 2020 Jan 28. PMID: 32171465.

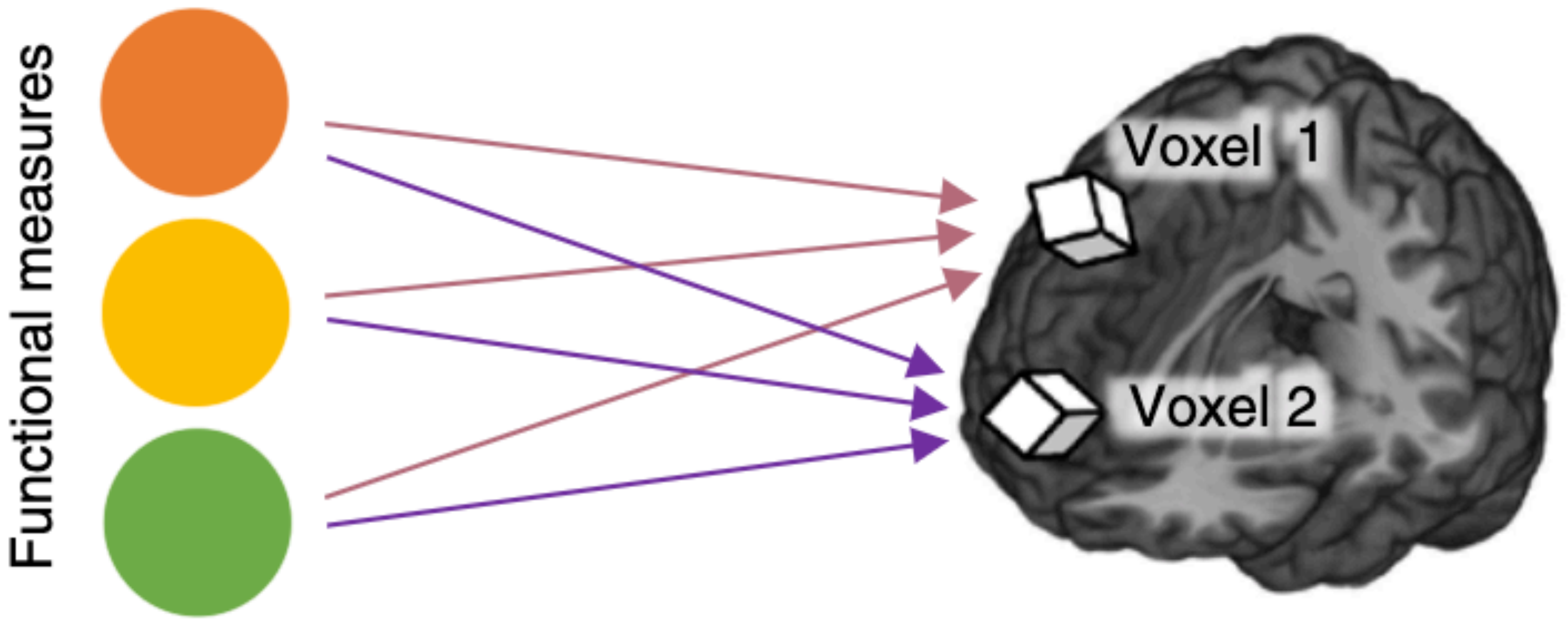
Hampel H, Vergallo A, Caraci F, Cuello AC, Lemercier P, Vellas B, Giudici KV, Baldacci F, Hänisch B, Haberkamp M, Broich K, Nisticò R, Emanuele E, Llavero F, Zugaza JL, Lucía A, Giacobini E, Lista S; Alzheimer Precision Medicine Initiative. Future avenues for Alzheimer's disease detection and therapy: liquid biopsy, intracellular signaling modulation, systems pharmacology drug discovery. Neuropharmacology. 2021 Mar 1;185:108081. doi: 10.1016/j.neuropharm.2020.108081. Epub 2020 May 11. PMID: 32407924.

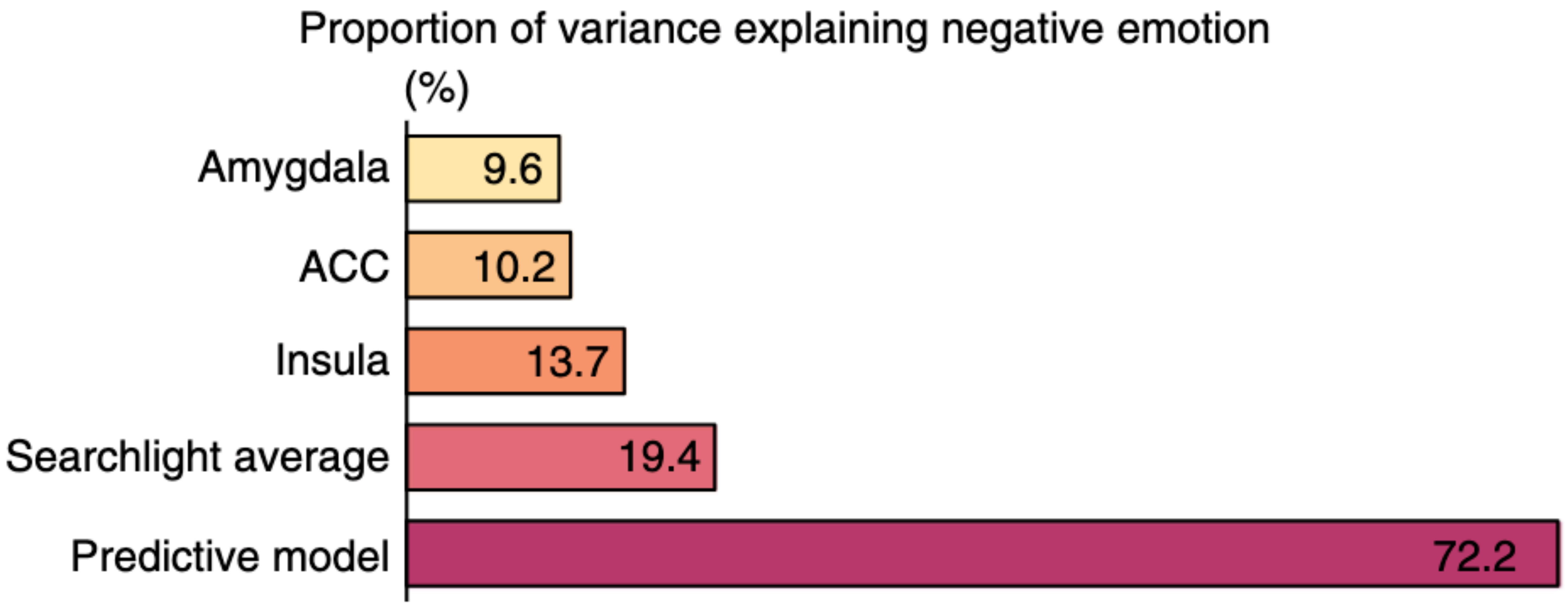
Many sources of heterogeneity

- Overlapping Symptoms
- Lived-experience
- Behavior
- Clinical outcomes
- Socio-demographics
- Genetic
- Temporal variability
- Socio-environmental
- Cognitive functioning
- ....









P(Brain | Stimulus or Symptom)

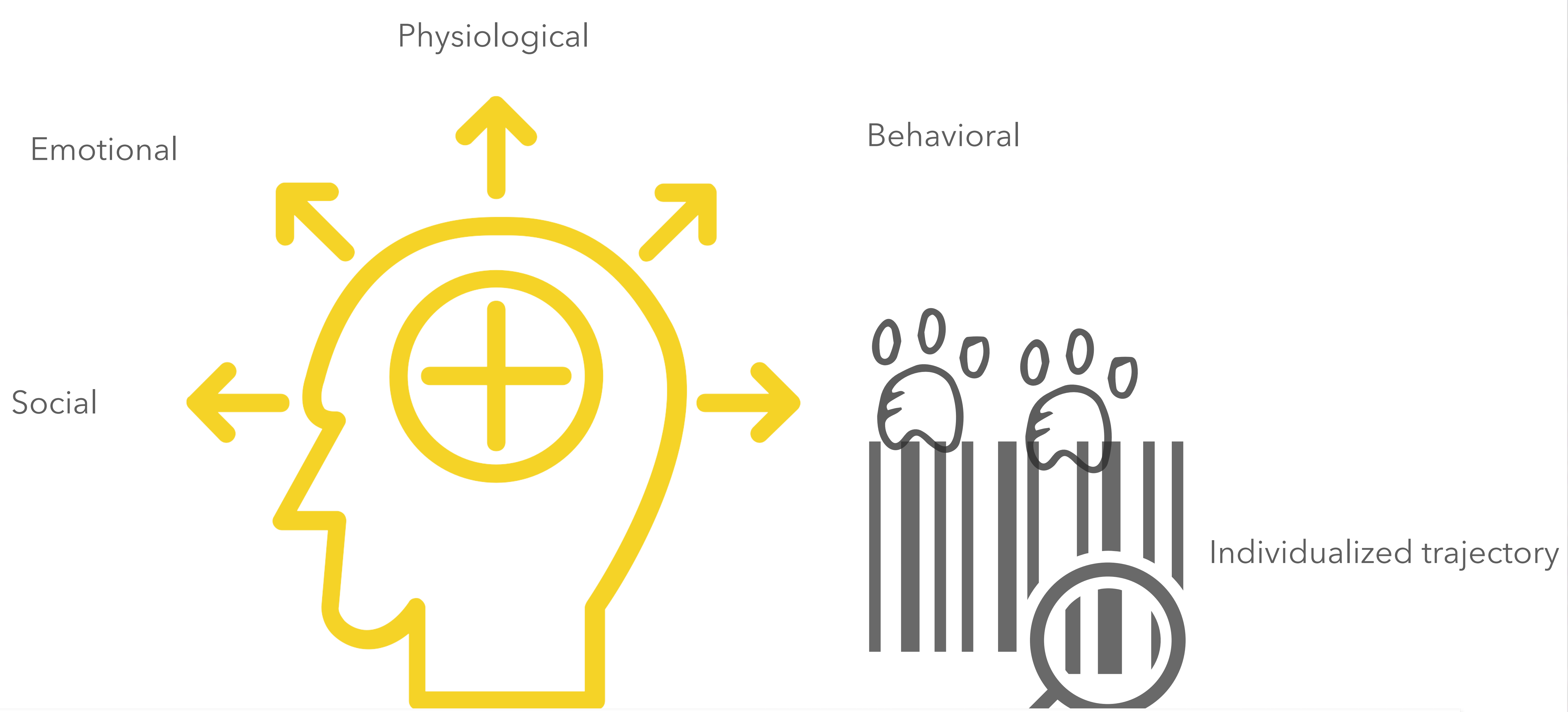
Connecting Lab to
Real-world behavior

P(Symptoms, Outcomes | F$_{1}$ ..... Sex, Genes, Behavior, Social$_{temporal}$, ... Fn)

P(Symptom, Outcome | Brain+)

Woo, CW., Chang, L., Lindquist, M. *et al.* Building better biomarkers: brain models in translational neuroimaging. *Nat Neurosci* **20,** 365–377 (2017). https://doi.org/10.1038/nn.4478

Digital Thermometer for Mental Health

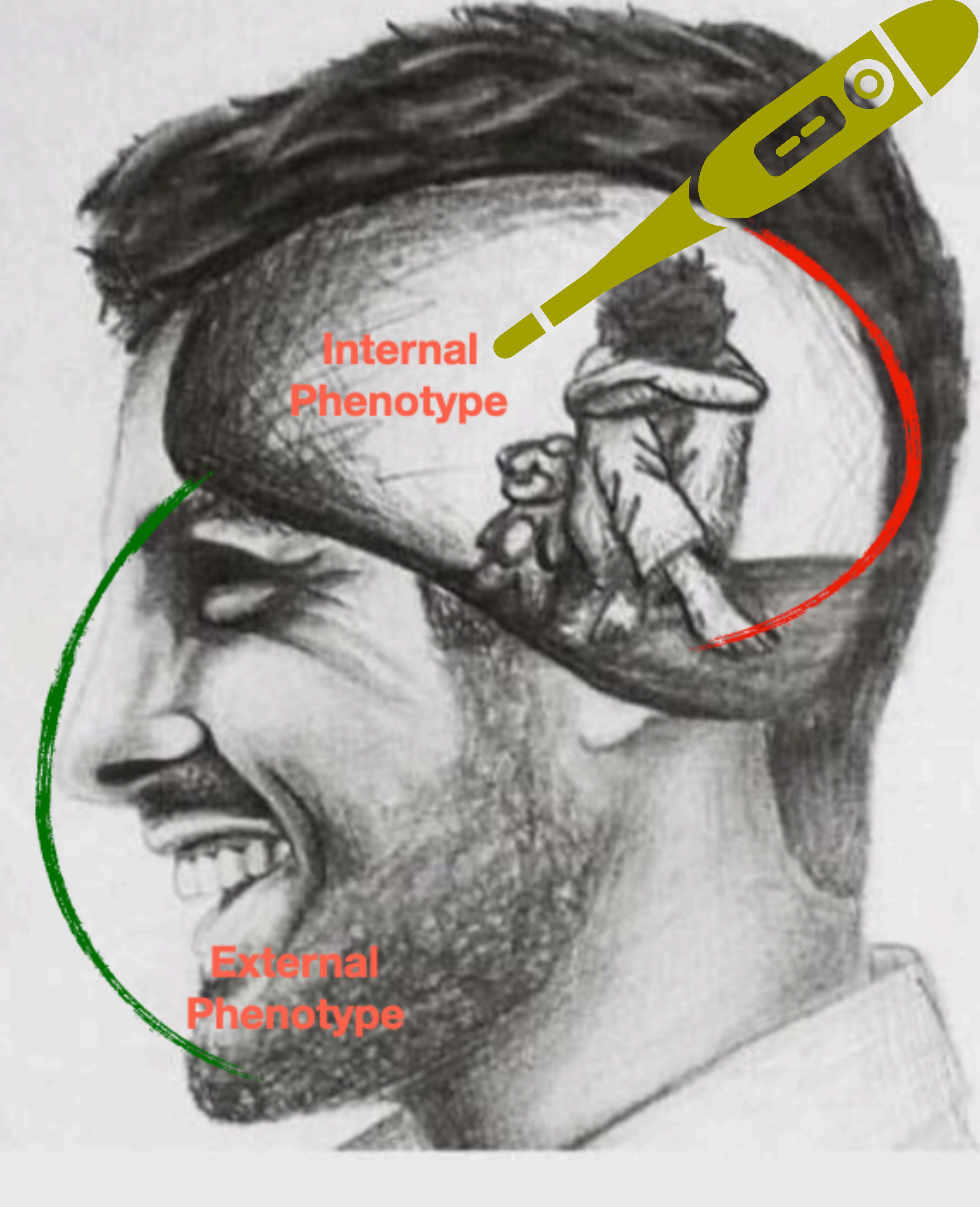


"...... *trying to understand another human being's emotional life is fraught with potential error* .....

.....*one's own prejudices and needs* ....

.......*we have no objective yardstick for this confidence*."

– **Emil Kraepelin, The Manifestations of Insanity, 1920**



> 99%
real-world lived experience of mental health
is NOT captured
clinic
assessments
Missed clinic visit
Assess in the clinic
(episodic)

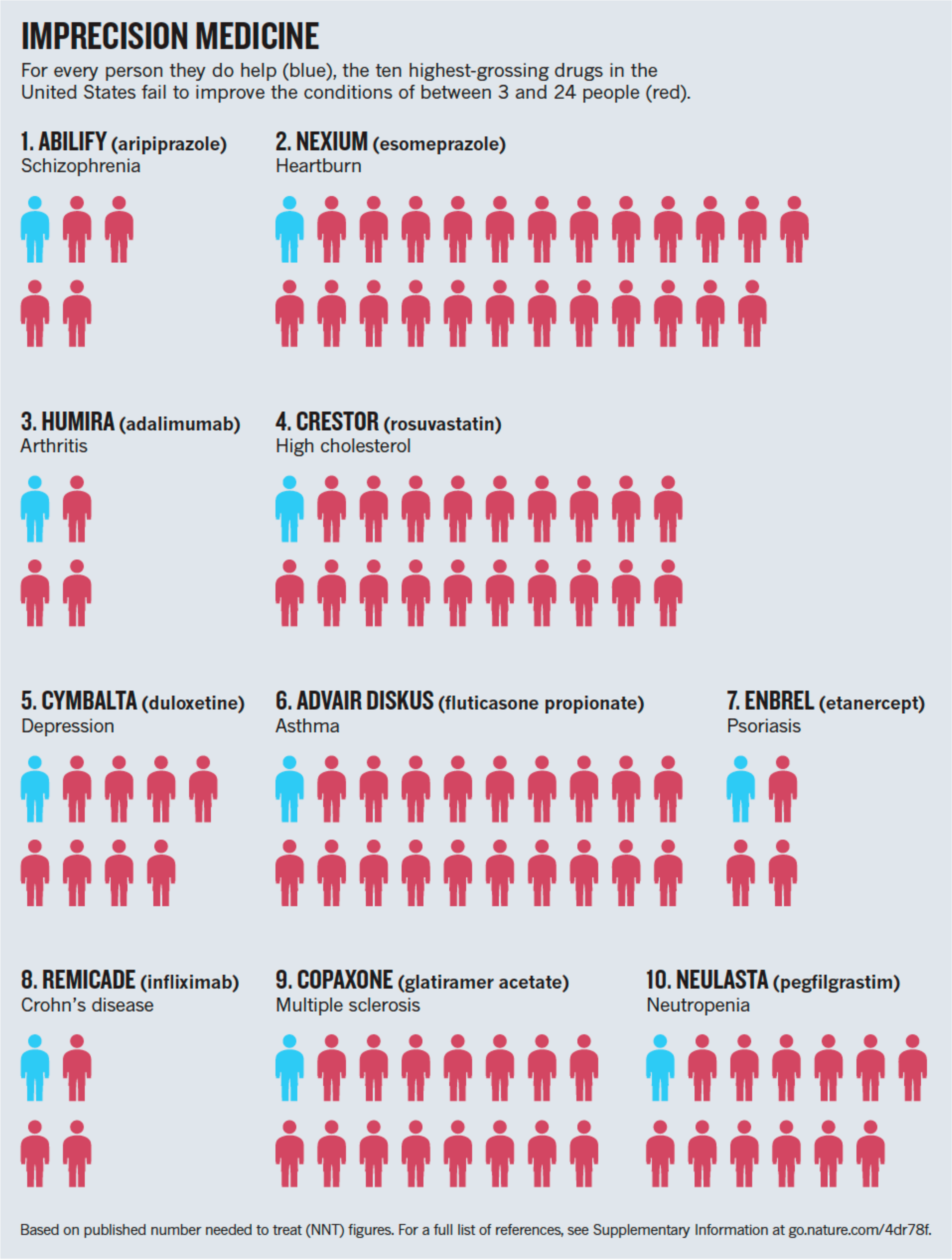
Sense individualized lived experience for the target population

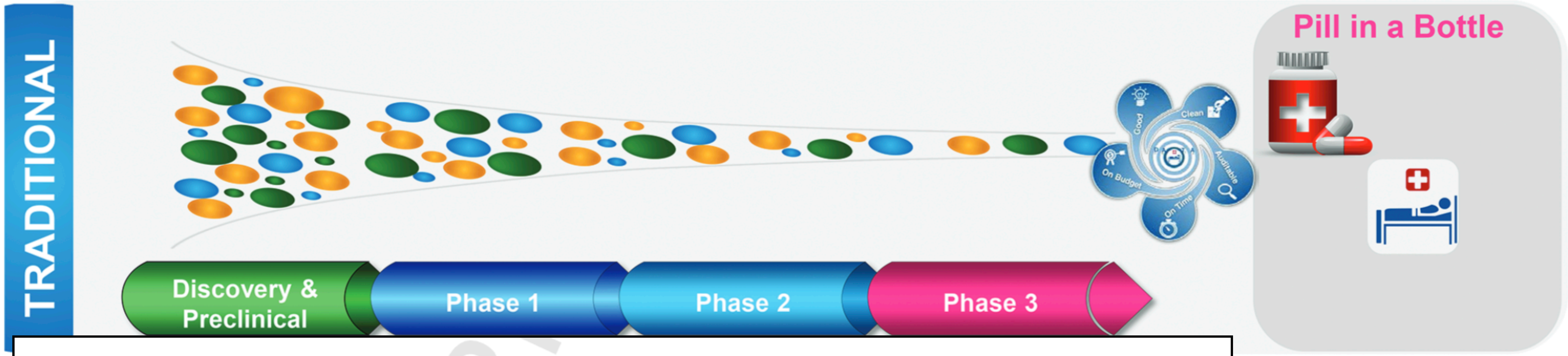
Sense individualized lived experience for the target population

Digitally Augmented

- Fully remote
- Hybrid - remote + episodic in-clinic

Augmenting digital tech* in clinical research is also a necessity





ulation in time

Wessels, A.M. et al. Cognitive Go/No-Go decision-making criteria in Alzheimer's disease drug development, Drug Discov Today (2021), https://doi.org/ 10.1016/j.drudis.2021.01.012
Clinical and Translational Science, Volume: 11, Issue: 5, Pages: 450-460, First published: 16 May 2018, DOI: (10.1111/cts.12559)

Why

Digital Health for Mental Health

Opportunities

How

Using digital health to assess CNS symptoms
“in the real world”

Feasibility & Predictability

When

If we build tech, communities will embrace it

Challenges & Solutions

Example 1:
Creating a digital health study

Digitally Augmented (Decentralized*) studies/trials

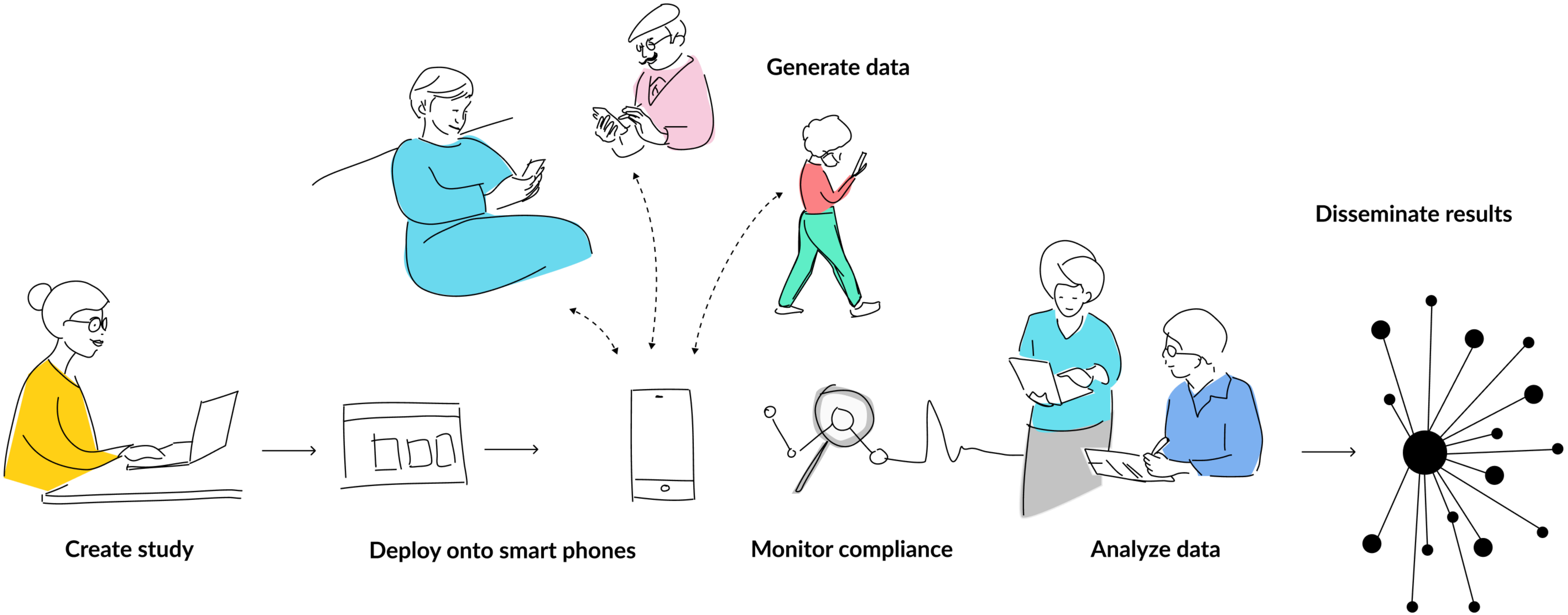


Image - Lynn Bui @ Sage

Digitally Augmented (Decentralized*) studies/trials

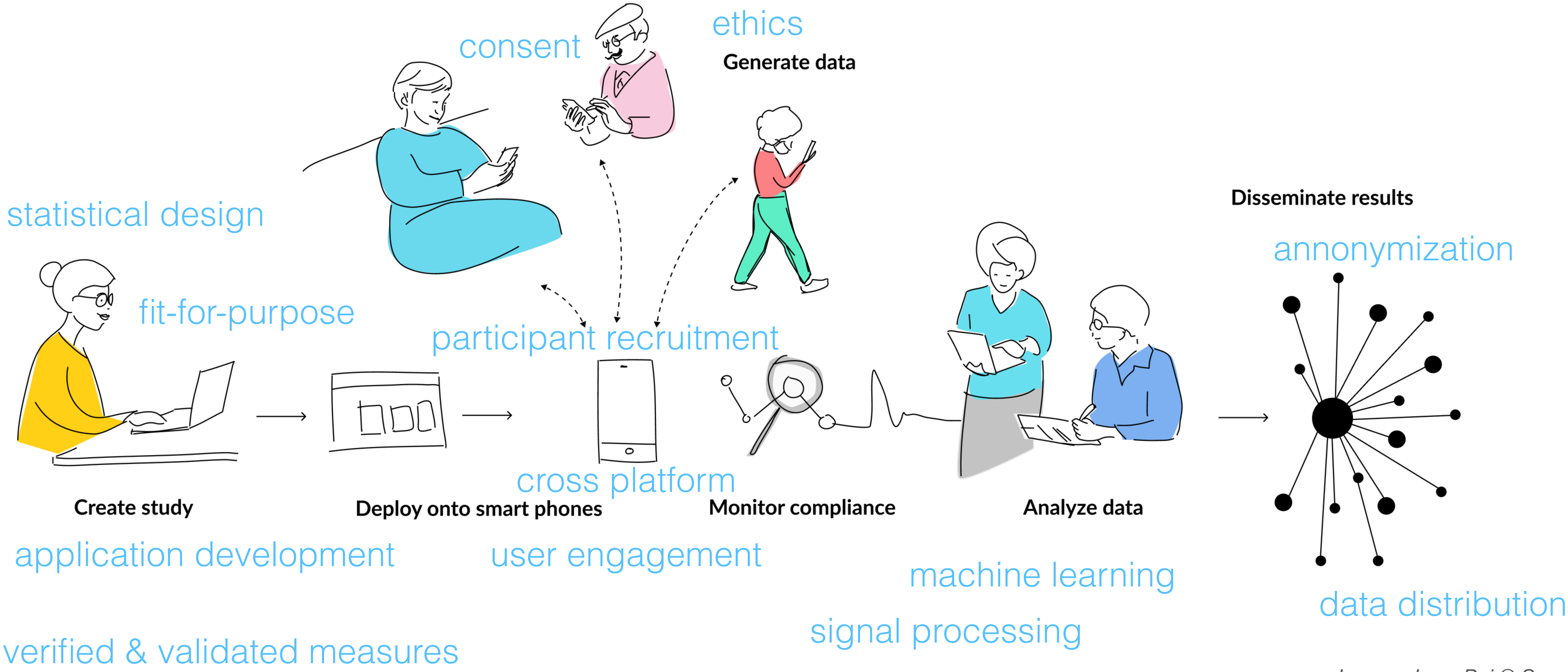
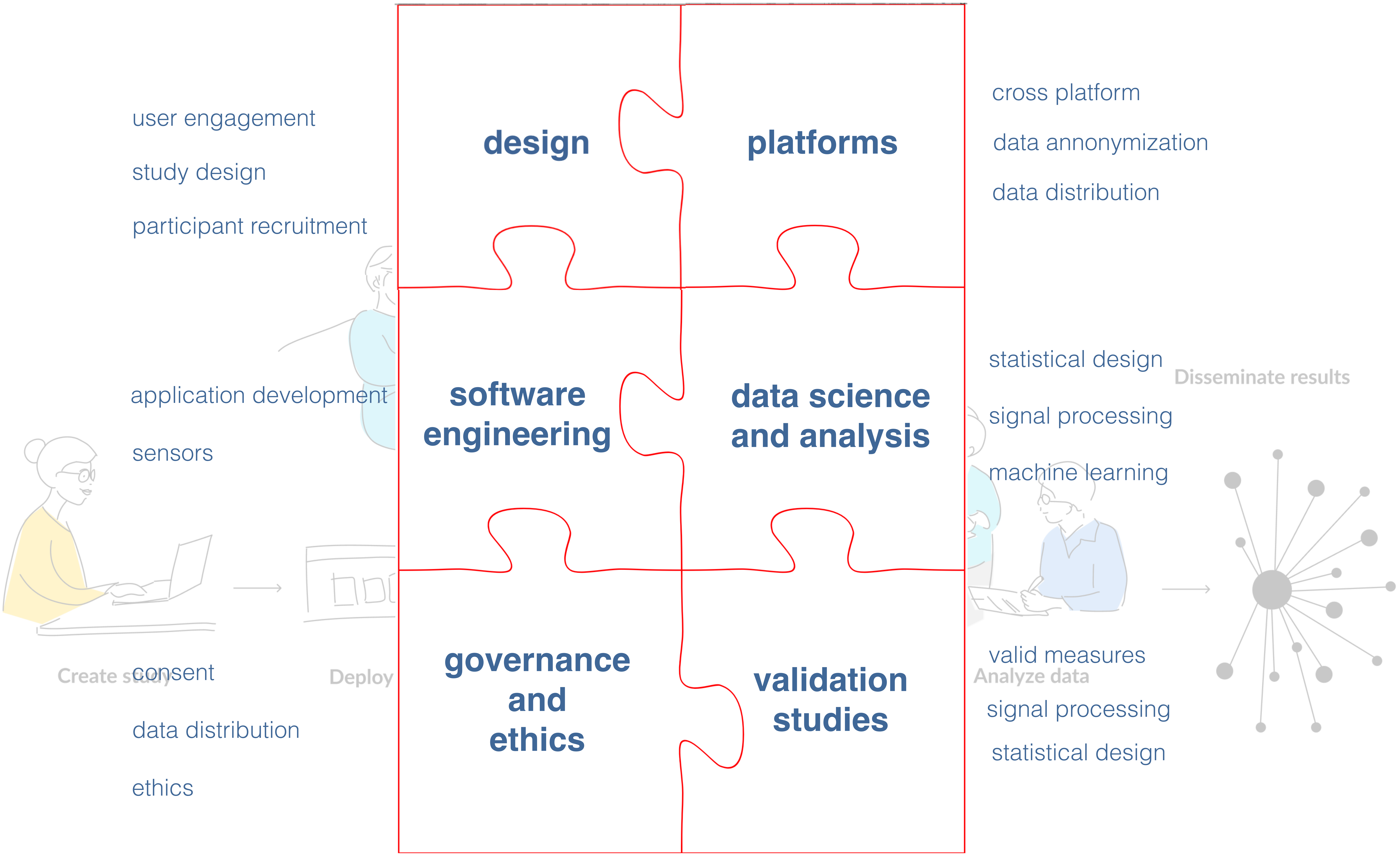
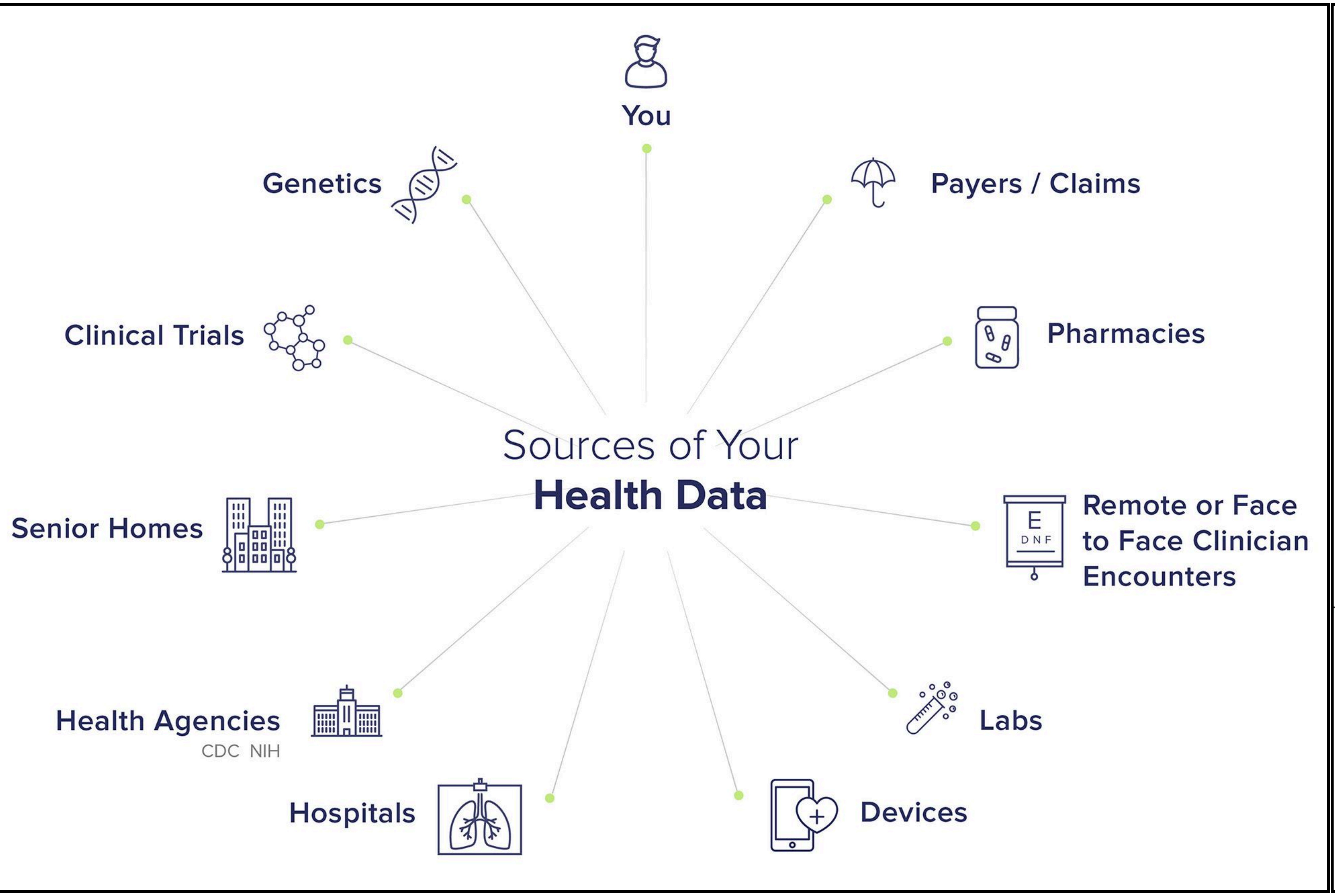


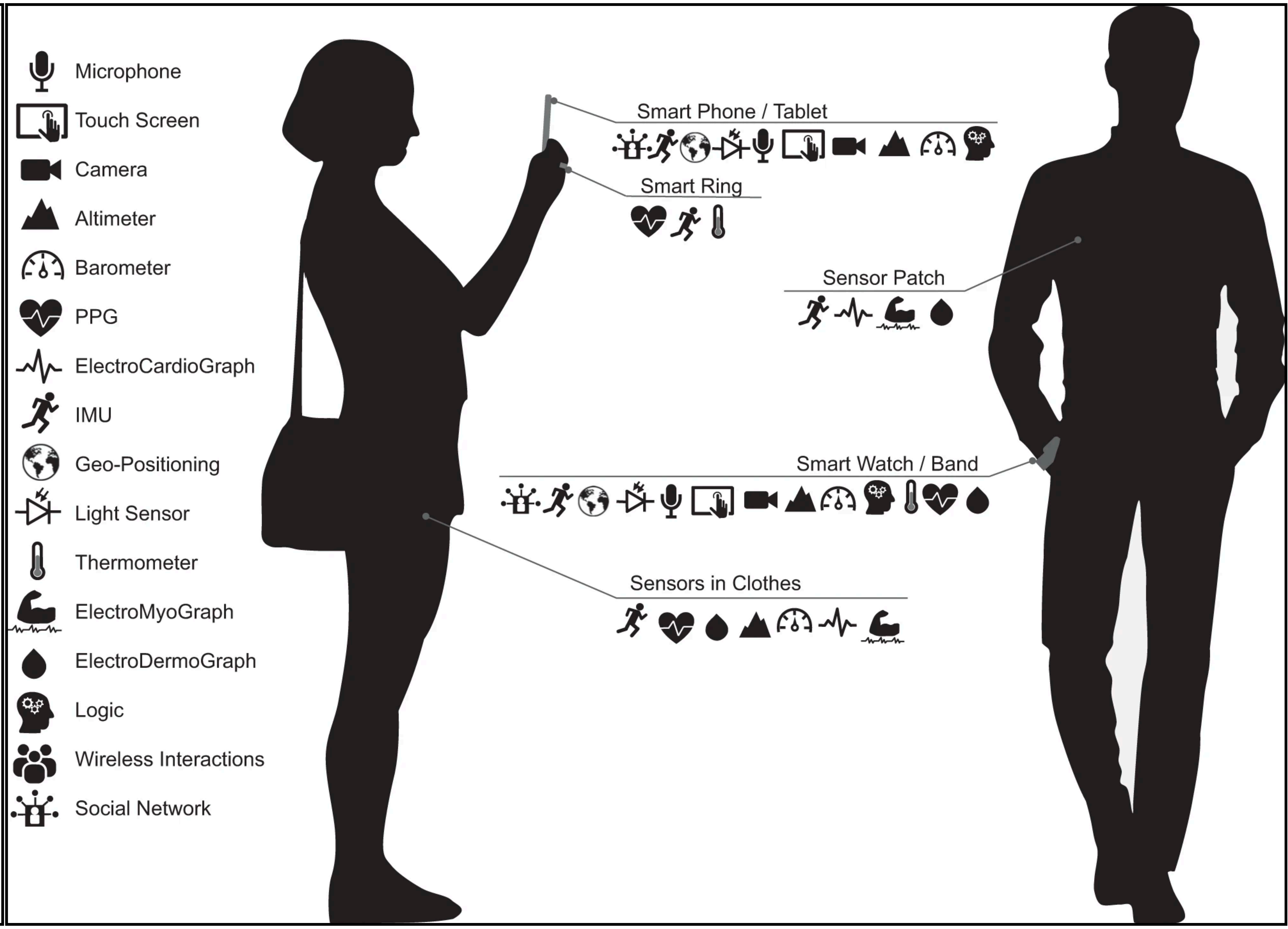
Image - Lynn Bui @ Sage


user engagement
study design
participant recruitment
design
platforms
cross platform
data annonymization
data distribution
application development
sensors
software engineering
data science and analysis
statistical design
signal processing
machine learning
Disseminate results
Create study
Deploy
Analyze data
consent
data distribution
ethics
governance and ethics
validation studies
valid measures
signal processing
statistical design

Collect **real-world mental health data from 3 countries**

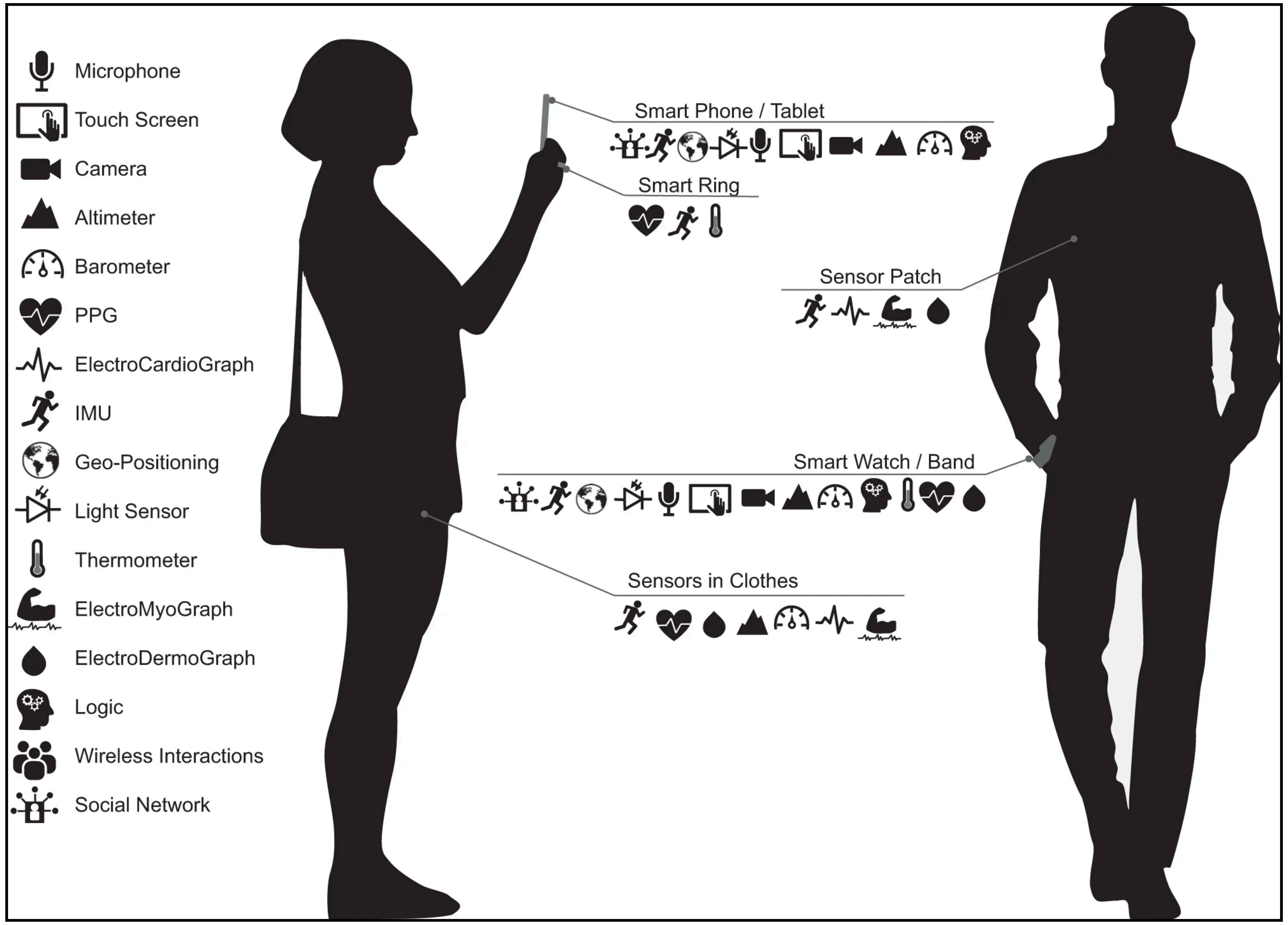
What data do we need to assess outcomes of interest?



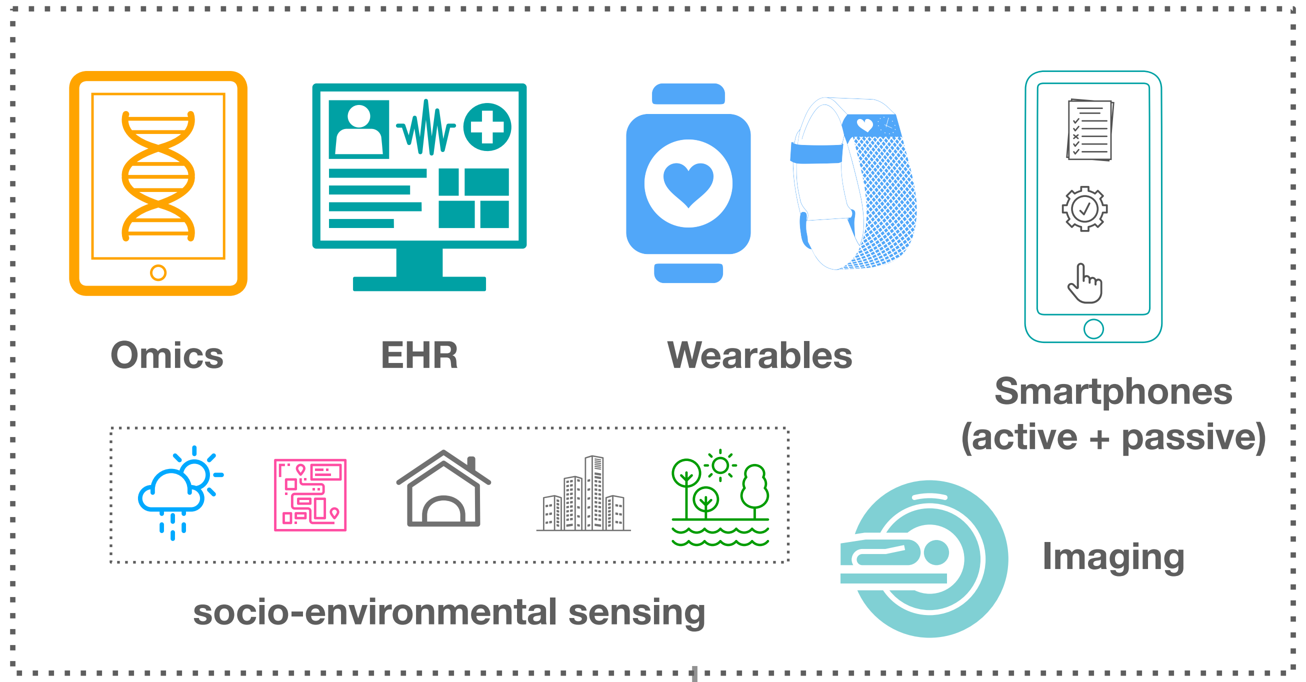


Its not about deploying a lot of measures rather right measures

Depression & Anxiety



How COA's will be collected



Tech's
fit-for-purpose

- Relevance to COA's
- Verification and validity of selected data streams
- Account for temporal changes in tech
- Reaching the target population
- How much data and for how long
- Participant Burden
  - Active | Passive | Hybrid
  - At-home | In-person | Hybrid
- Data governance & privacy

https://www.ichom.org/portfolio/depression-anxiety/

Who
will we enroll

How
will we reach them

What/How
will we collect data

Recruitment Sites
- India
- UK
- South Africa

Localized site-specific recruitment

Social media

Reddit

Forums

Reach

Enrollment Website
+
Site-specific alternatives

Enrollment

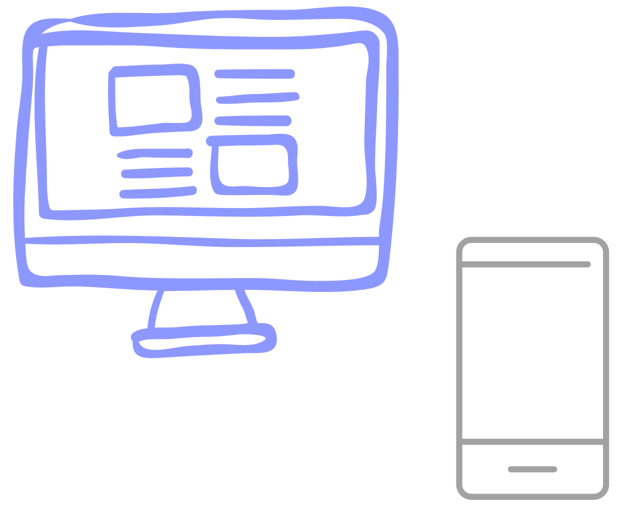
Triage

BYOD
Bring-your-own-device

BYOD - proximal
Survey only*

Participant Onboarding

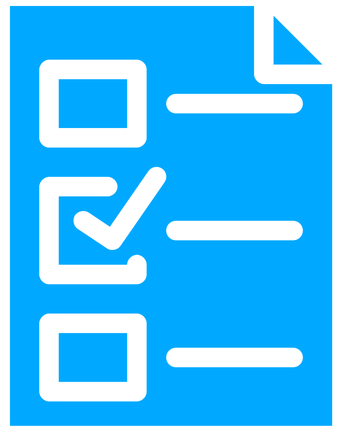
Enrollment
(Post inclusion/exclusion criteria)



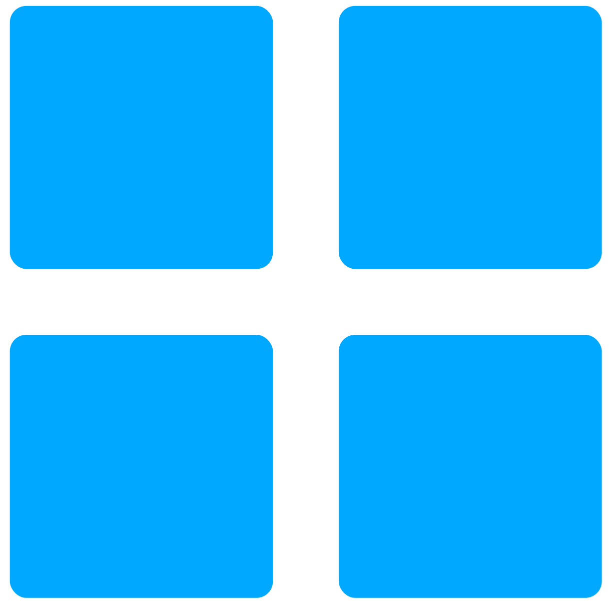
eConsent

Baseline Assessments

Demographics
&
Socioeconomic status

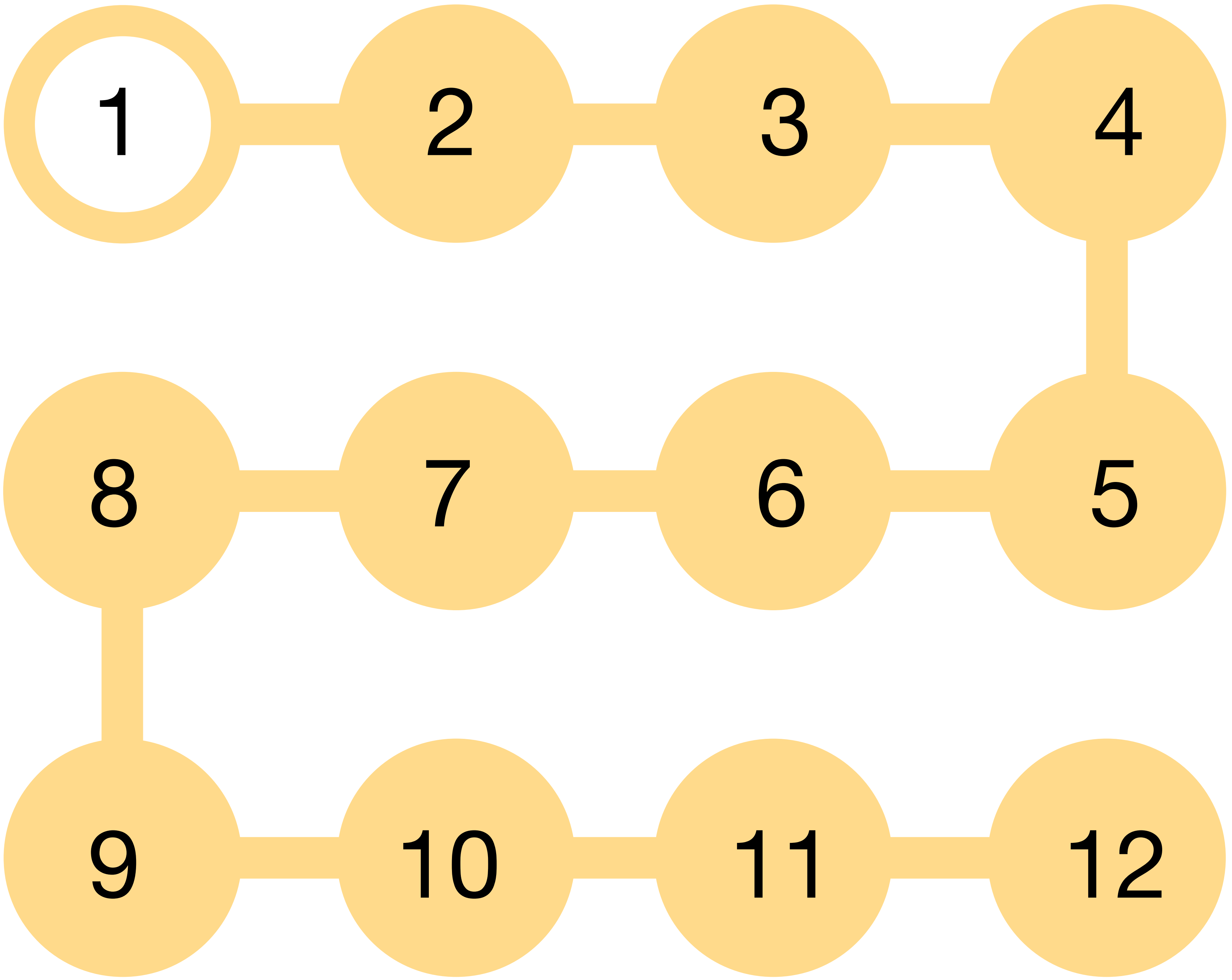


Mental Health History
(Focus on
Depression & Anxiety)



+ Mobile phone usage
+ any other surveys
(a.k.a widgets)

How long will this study last ?



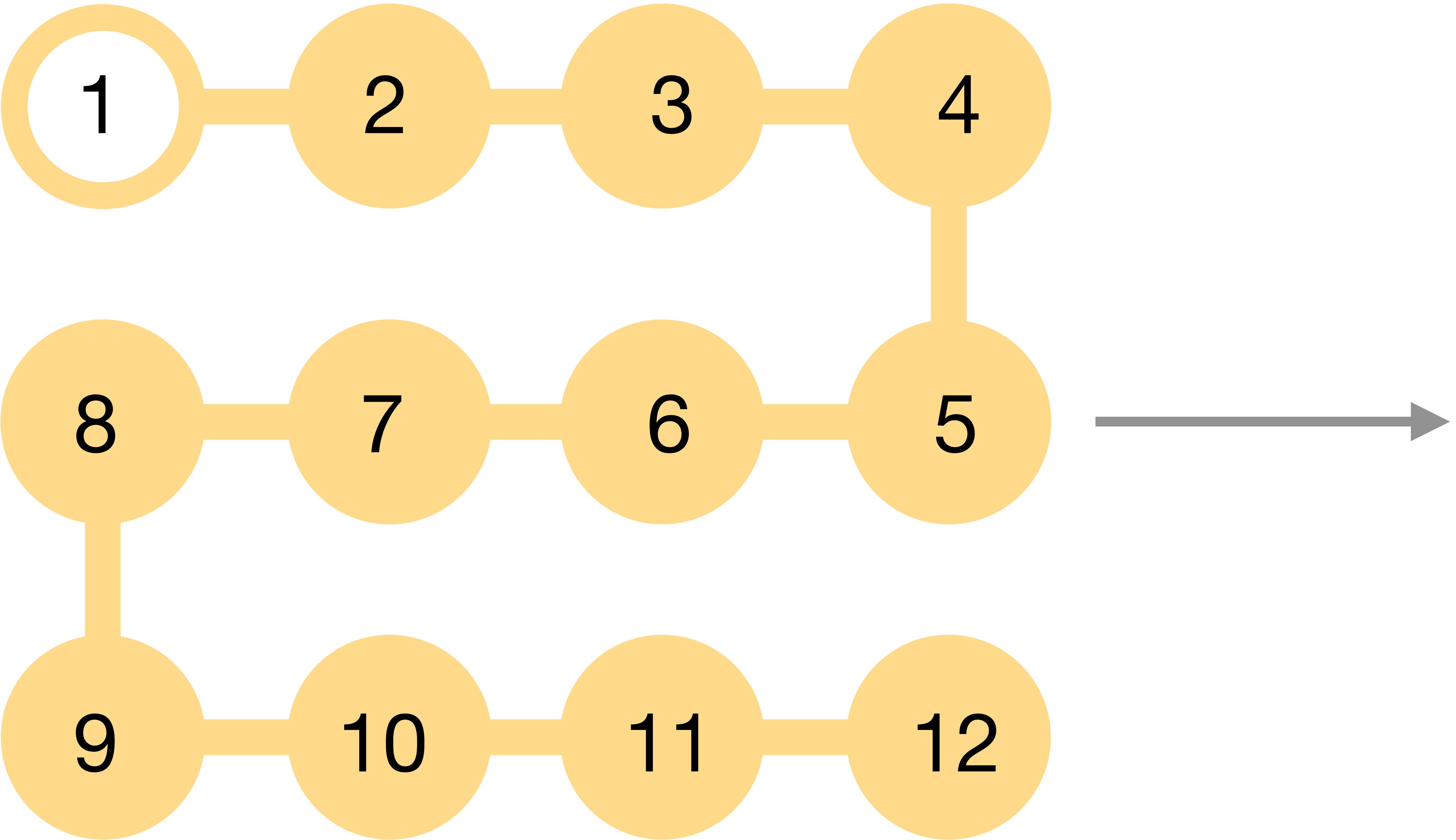
Study participants are expected to be engage in the study for 12 weeks

(Participants will be able skip any question/s or leave the study at any time)

How will the study collect data?

Participants will install the study app on their smartphone

Longitudinal data collection protocol ?

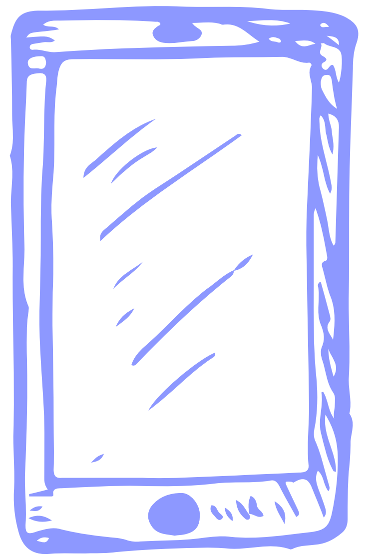


Remote assessments (survey)

+

Sensor-based measures

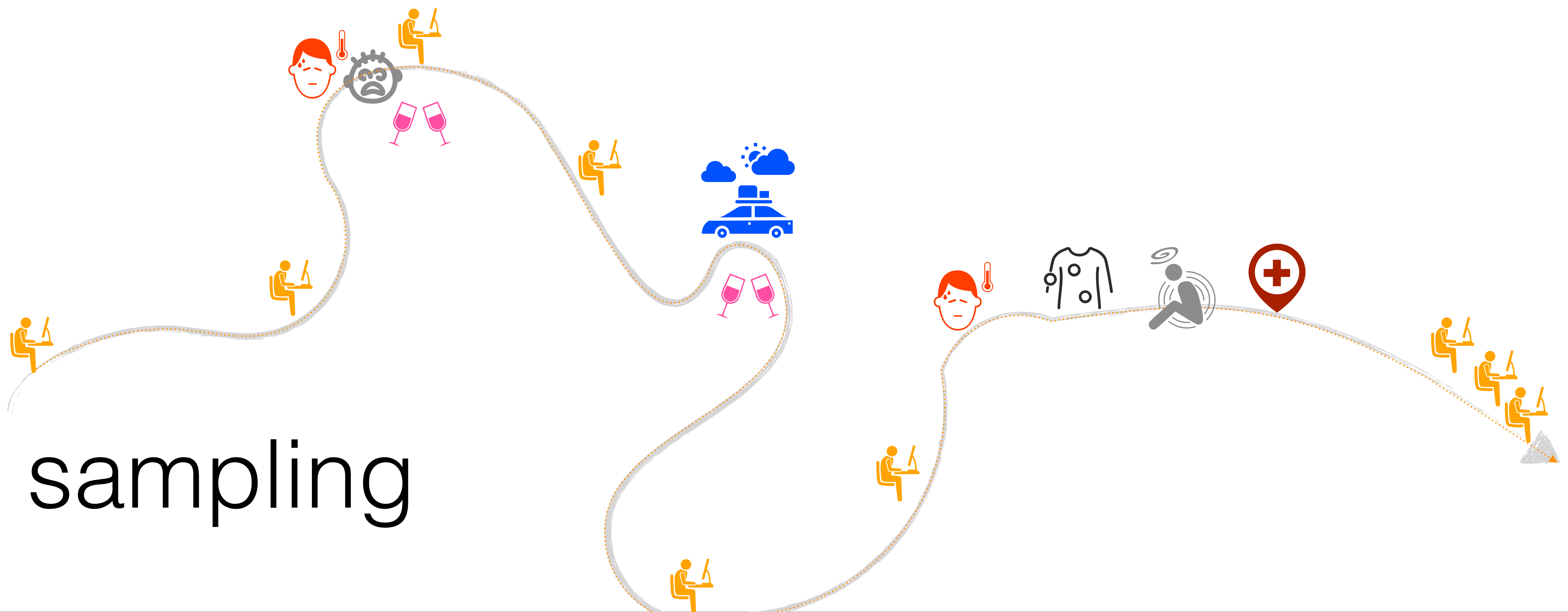
Okay, but can we be more specific ?



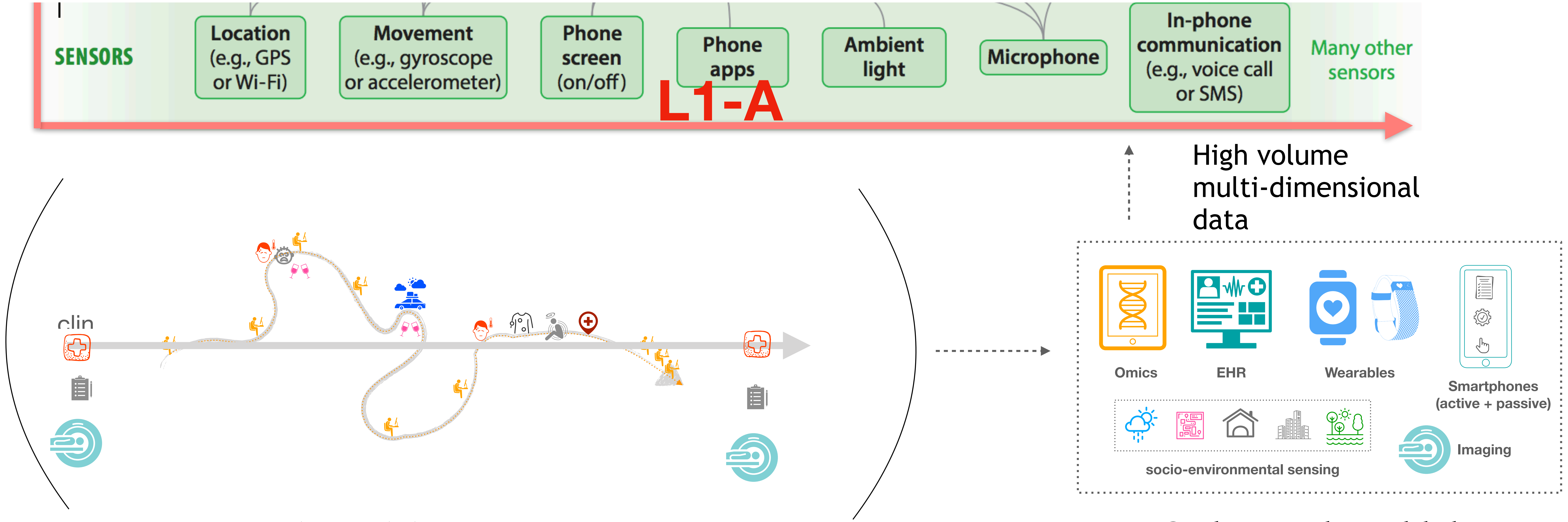
Daily & Weekly Surveys

e.g. - Weekly assessment of mood, physical activity etc

(Participants will be able to skip any question/s or opt-out of sharing phone usage data)



long term continuous sampling

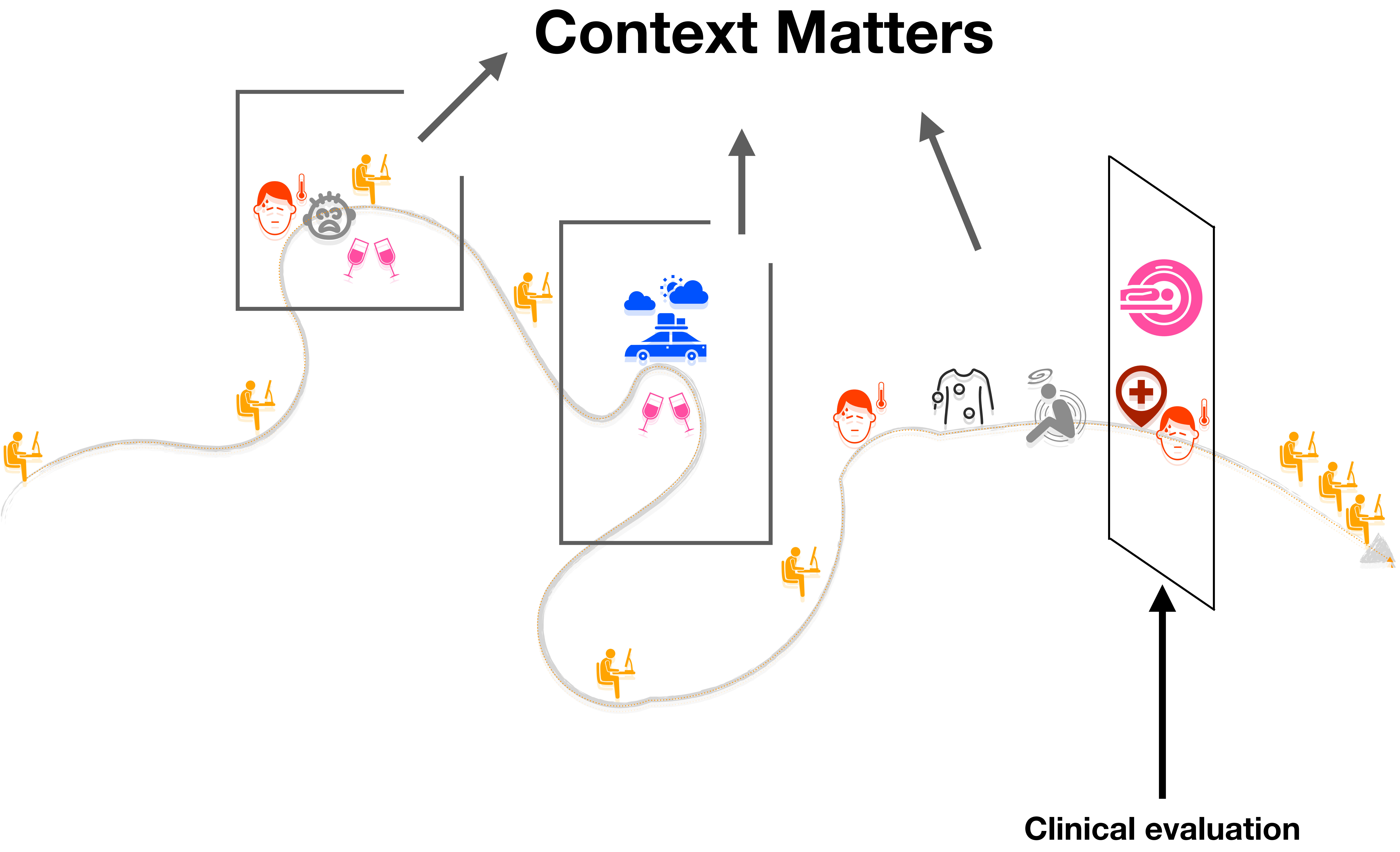


multimodal sensing

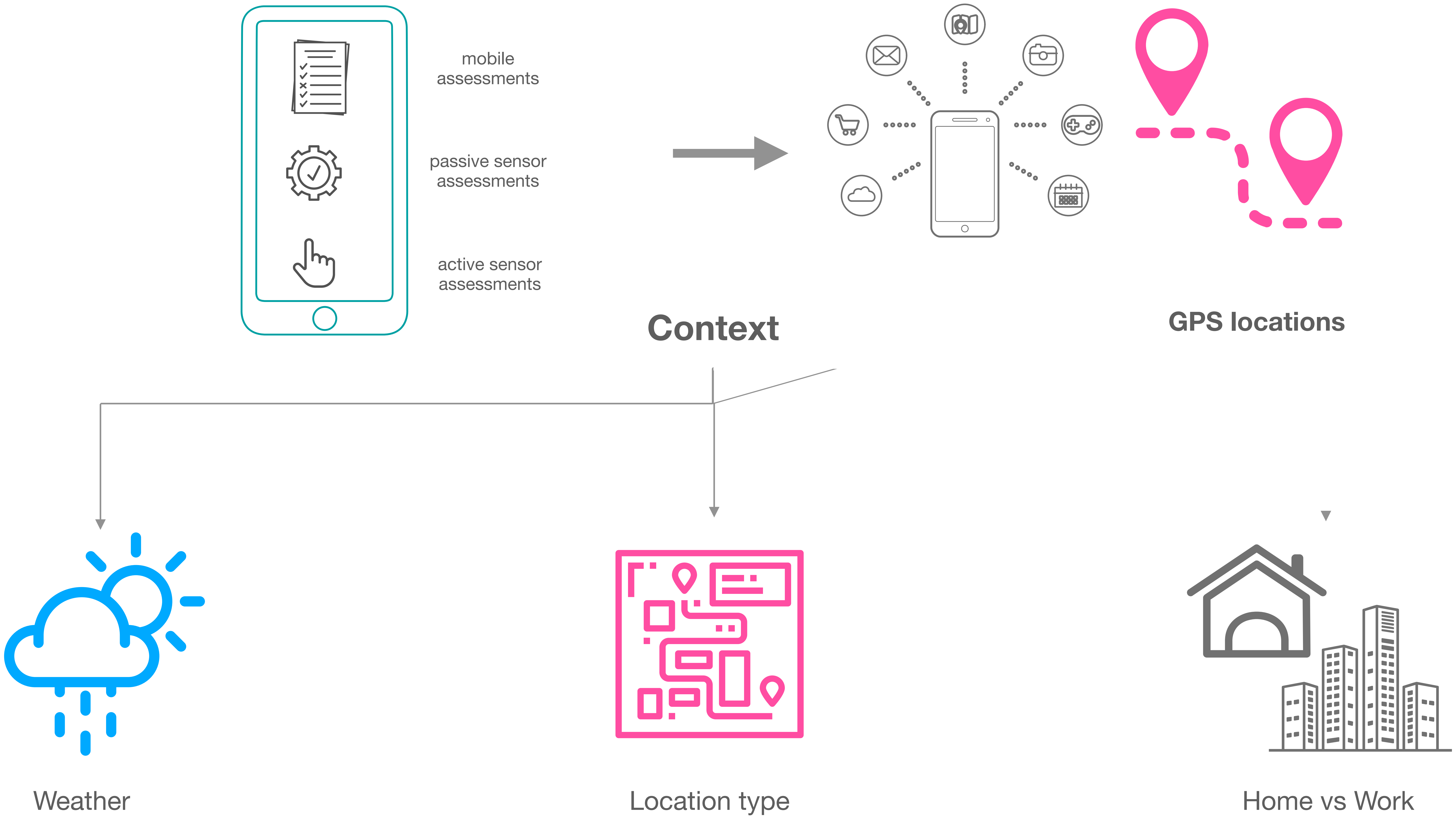
Gather real-world data

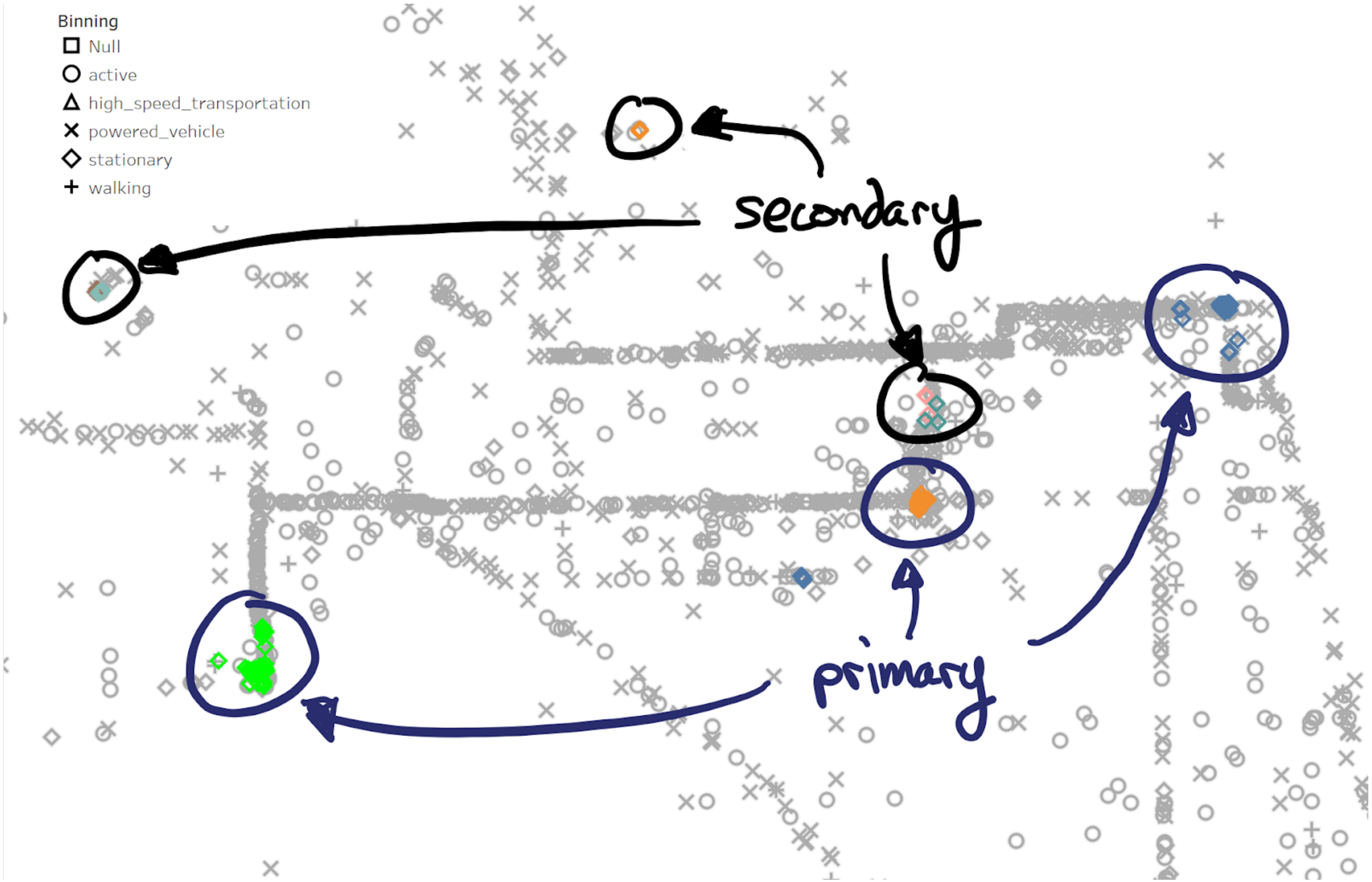
Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning | David C. Mohr, Mi Zhang, Stephen M. Schueller | Annual Review of Clinical Psychology 2017 13:1, 23-47

Outcomes may also be impacted by external factors that are uniquely linked to our daily lives

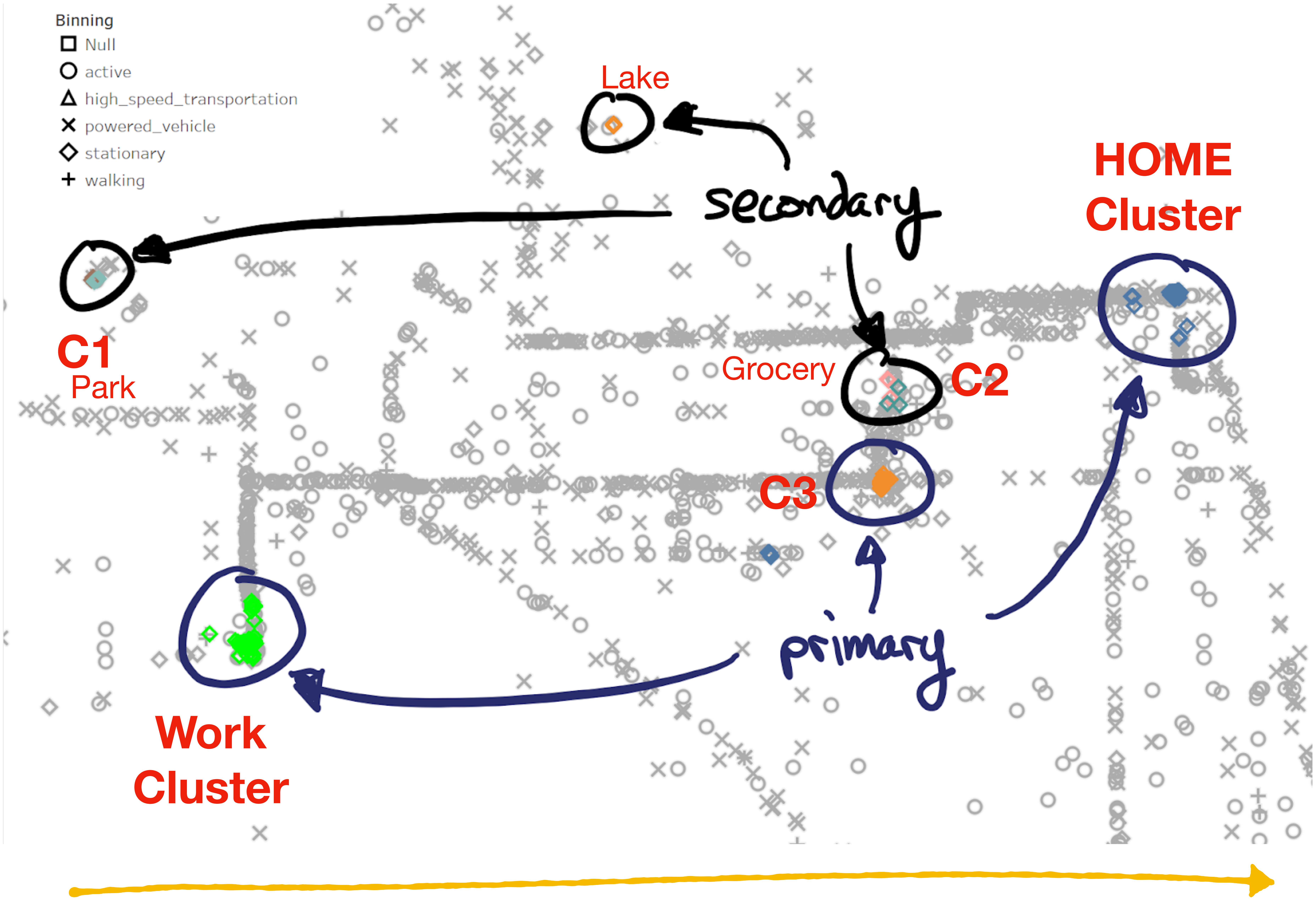


Deriving individualized geospatial context

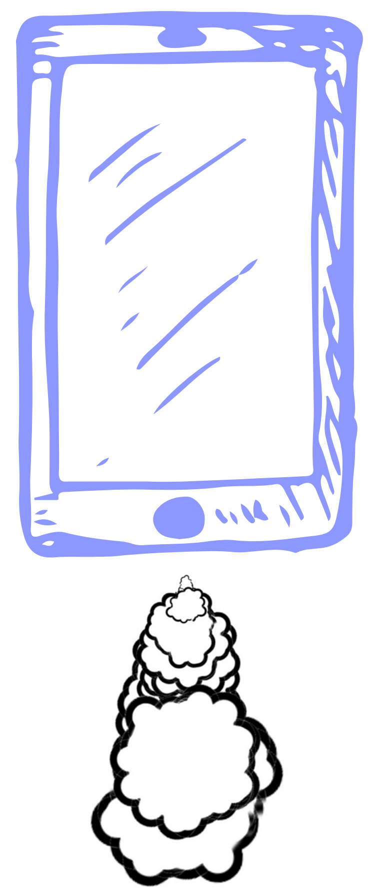




Using spatial density-based clustering (DBSCAN)



Daily location-based weather patterns



GPS

social media interaction

step counts / day

#messages/ day
#calls / day

ambient light levels

Screen Time

Wifi Networks

Weather

Air quality

Data usage

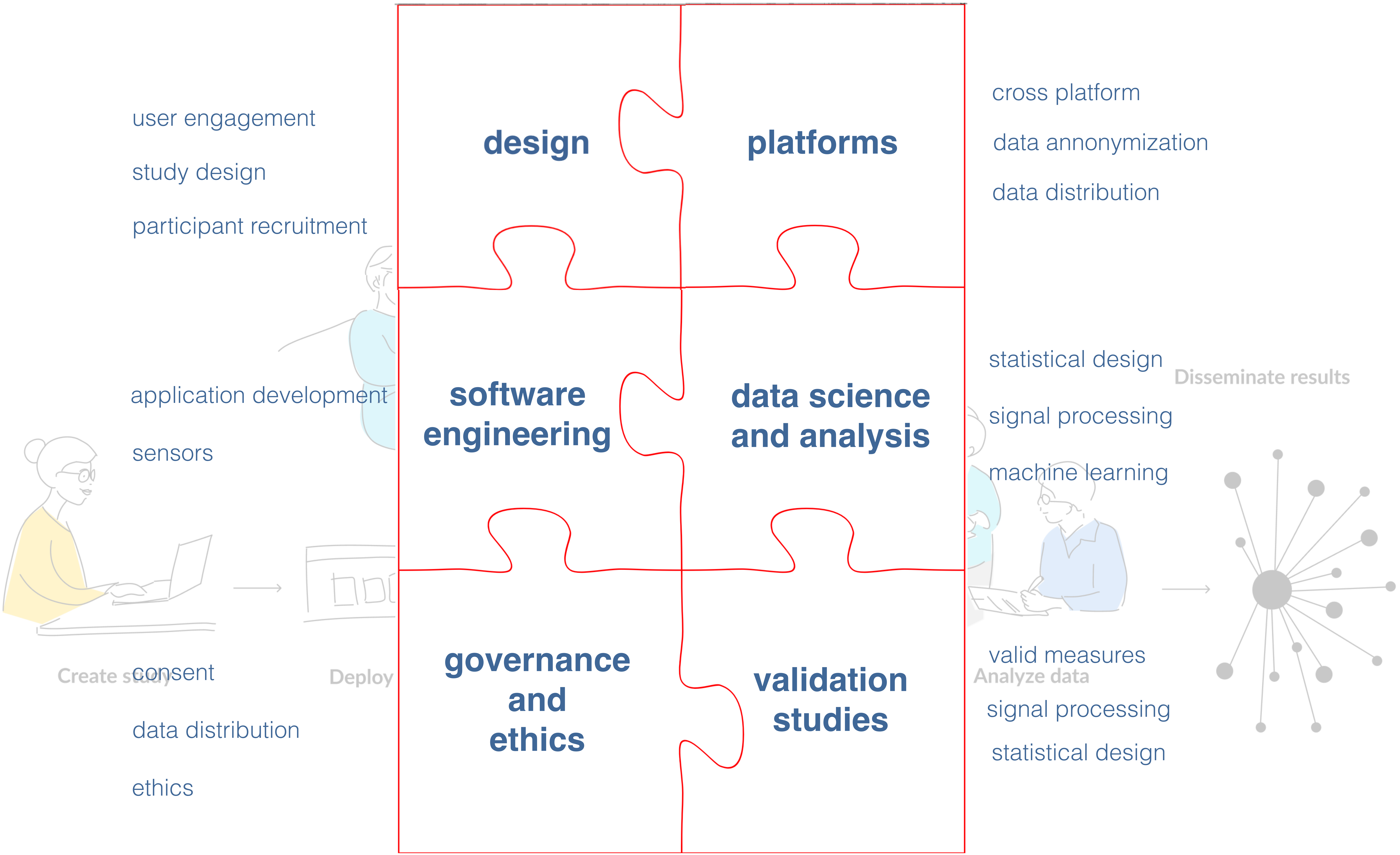
Screen lock | unlock

Background noise levels

Battery Usage

What data we will NOT collect ?

- Raw GPS data | your location (latitude, longitude)
- Voice calls, messages
- Any data stream/s that we hear your concerns about
- Any phone usage data without participant's explicit consent



user engagement
study design
participant recruitment
design
platforms
cross platform
data annonymization
data distribution
application development
sensors
software engineering
data science and analysis
statistical design
Disseminate results
signal processing
machine learning
Create study
consent
Deploy
governance and ethics
validation studies
valid measures
Analyze data
signal processing
statistical design
data distribution
ethics

Longitudinal Data Collection

Subjective

Episodic MH assessments

Continual Passive Data Collection (opt-in)

Digital Diary (text | voice | photos)

Surveys assessing exposure to known risks/protective factors

Objective

Environment

Ambient Light
Ambient Noise
Location Semantic
Daily Weather
Air Quality

Physical Activity

Step Count
Activity Recognition

Phone Usage

Screen time
WiFi
Battery Drain
Charging
Device on/off
Data usage
Lock/Unlock

Social

Instagram
Facebook
Calls
Messages

Feasibility & Predictability

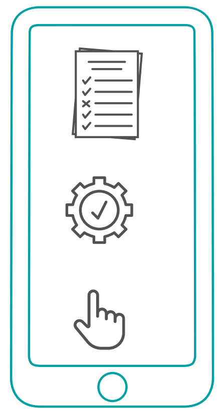
Example 2

Smartphone-based behavioral sensing in the real world

Brighten Studies

two fully remote RCT for assessing and mediating in Depression

A cohort of depressed people was recruited fully remotely



Passive

GPS
Phone Usage

Surveys

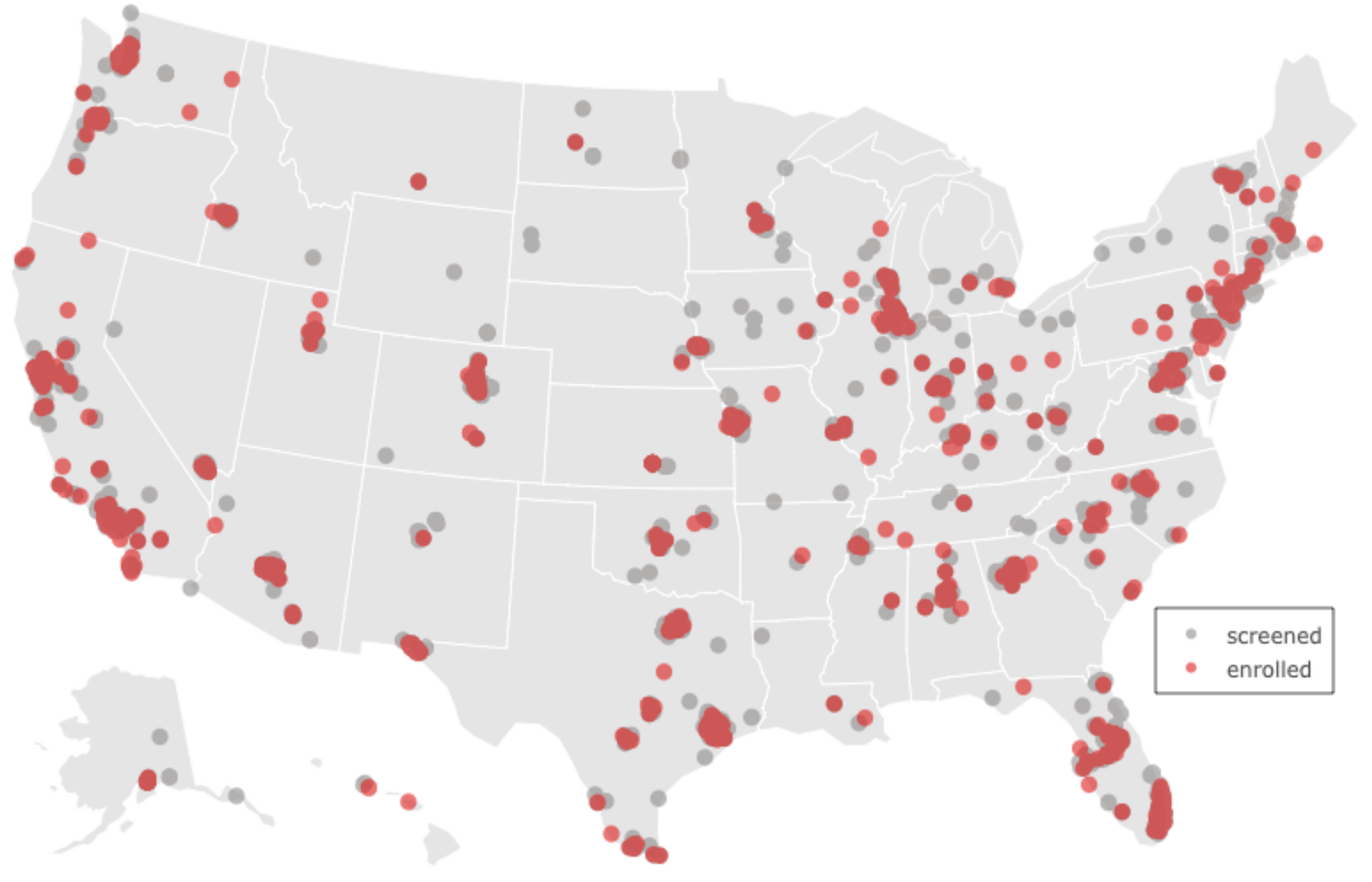
Daily Mood
Weekly PHQ-9

Interventions

Akili's EVO - Neuromodulation
iPST - Problem solving therapy
Health Tips - Placebo

12 weeks remote monitoring

Participants were paid up to $90 for completing self-assessments remotely



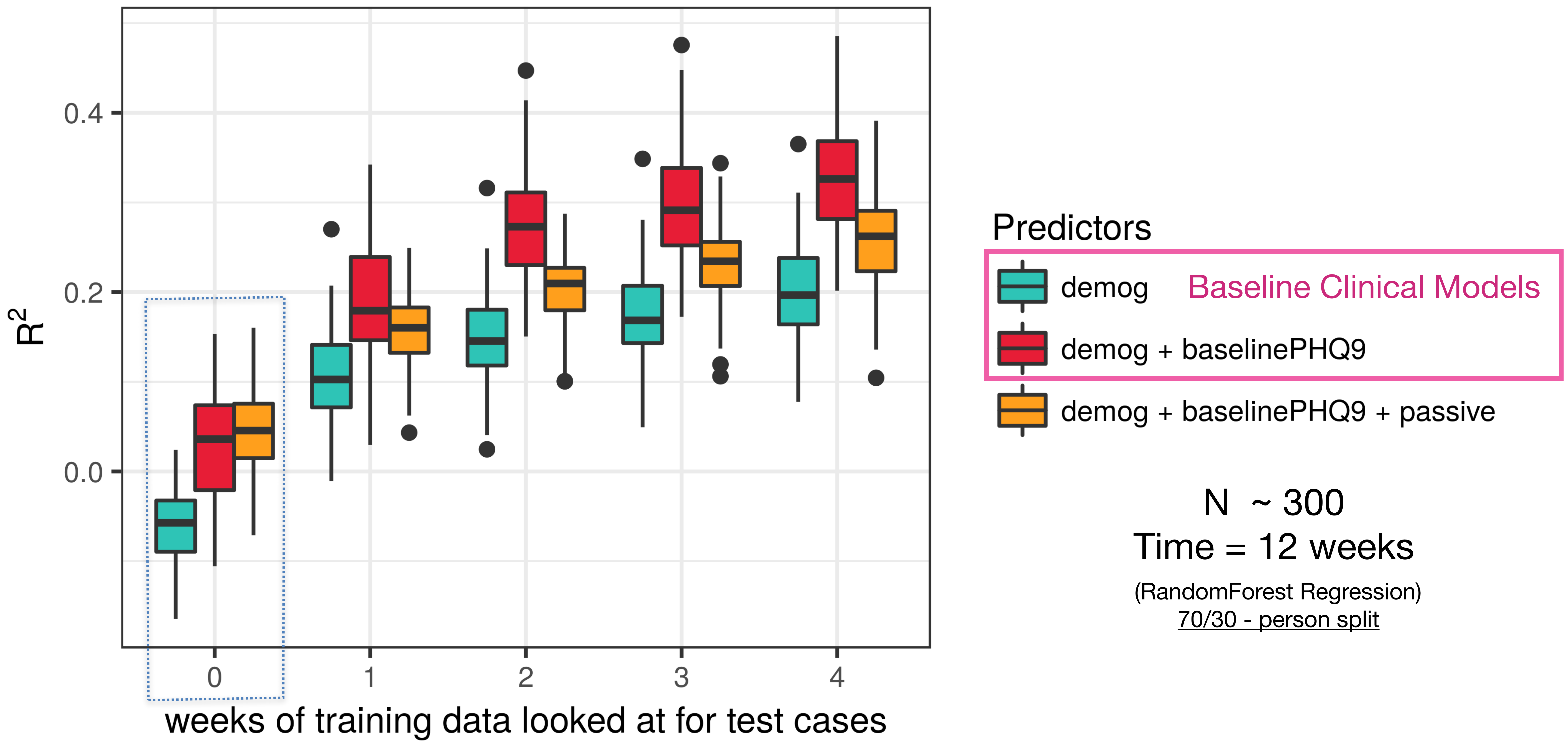
Screened - ~7000

Enrolled - ~2000

Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using Mobile Apps to Assess and Treat Depression in Hispanic and Latino Populations: Fully Remote Randomized Clinical Trial. **J Med Internet Res. 2018**;20: e10130

Pratap A, Anguera JA, Renn BN, Neto EC, Volponi J, Mooney SD, et al. The feasibility of using smartphones to assess and remediate depression in Hispanic/Latino individuals nationally. ***UbiComp '17. 2017***. *doi:10.1145/3123024.3127877*

Predicting daily mood at cohort level remains challenging



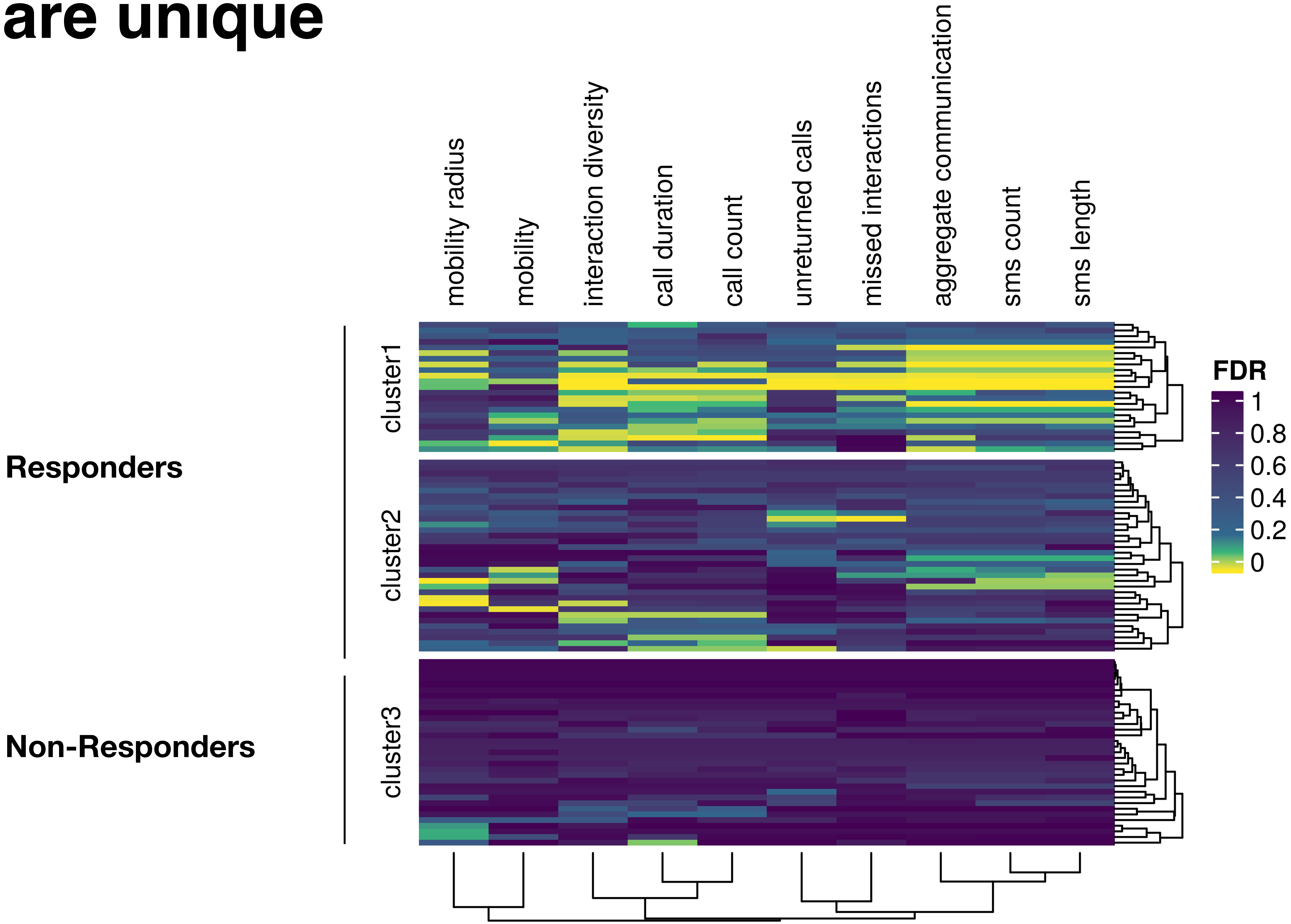
Further research is needed to predict behavioral health using passive data above and beyond demographics and baseline clinical assessment at **cohort level**

Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, Arean PA. The accuracy of passive phone sensors in predicting daily mood. **Depress Anxiety. 2019**; doi:10.1002/da.22822

1

People are unique

N ~ 200
Time = 12 weeks

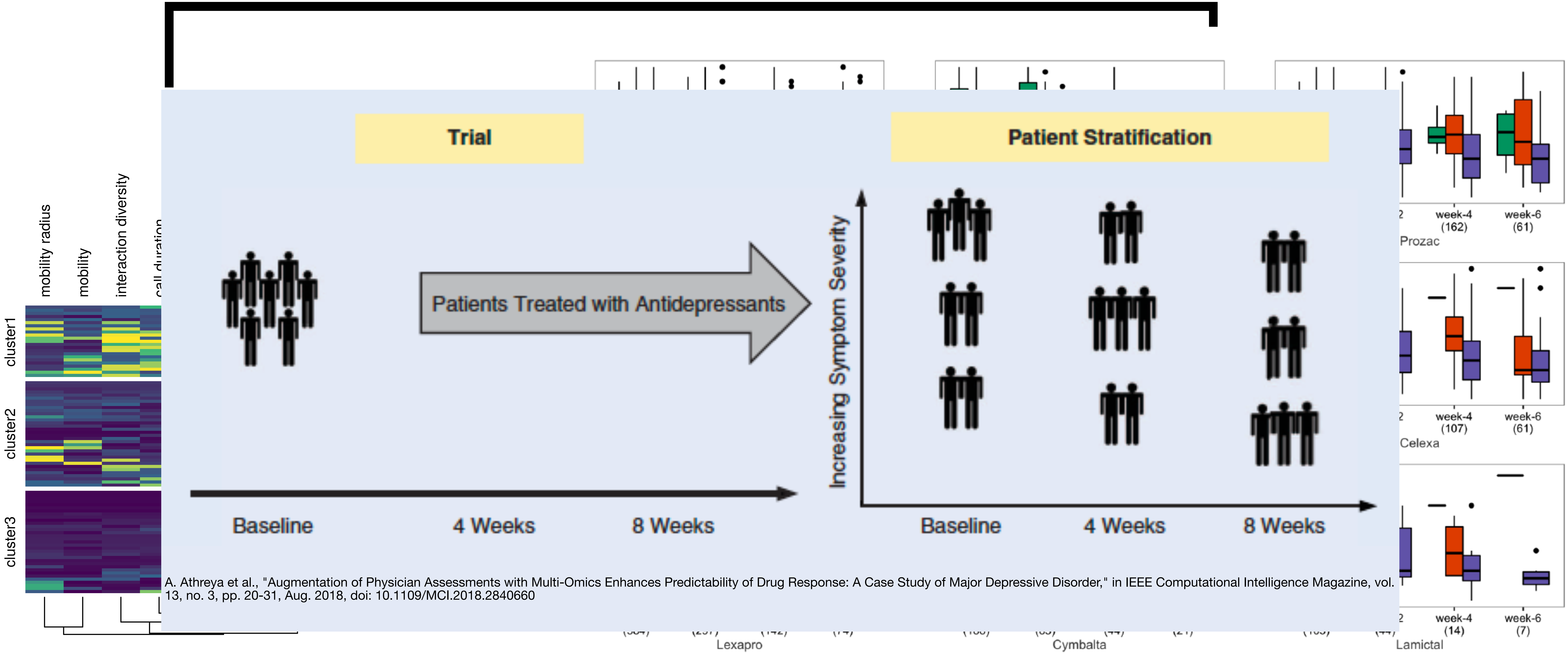


1

Understanding the underlying heterogeneity in Depression

Real-world behavior

Response to antidepressants



A. Athreya et al., "Augmentation of Physician Assessments with Multi-Omics Enhances Predictability of Drug Response: A Case Study of Major Depressive Disorder," in IEEE Computational Intelligence Magazine, vol. 13, no. 3, pp. 20-31, Aug. 2018, doi: 10.1109/MCI.2018.2840660

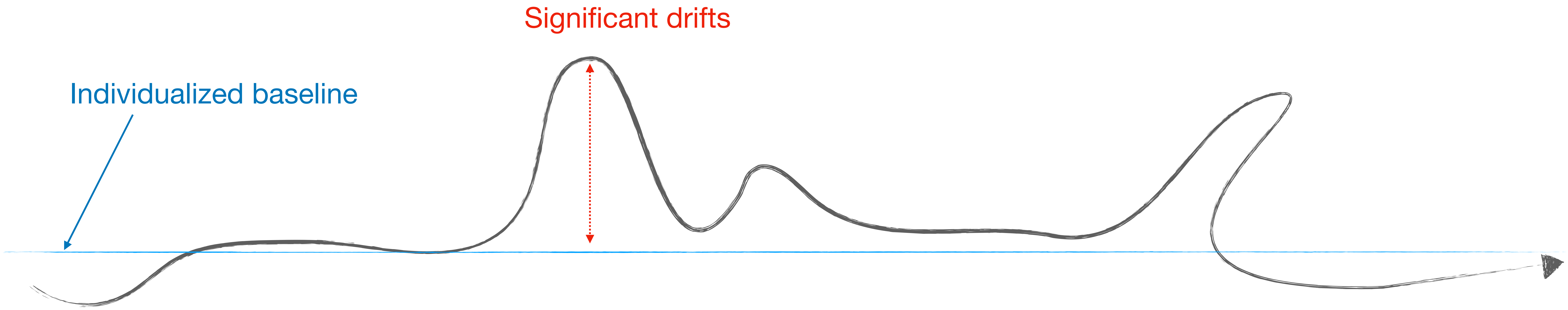
Study 1

Study 2 (unpublished)

(work-in-progress)

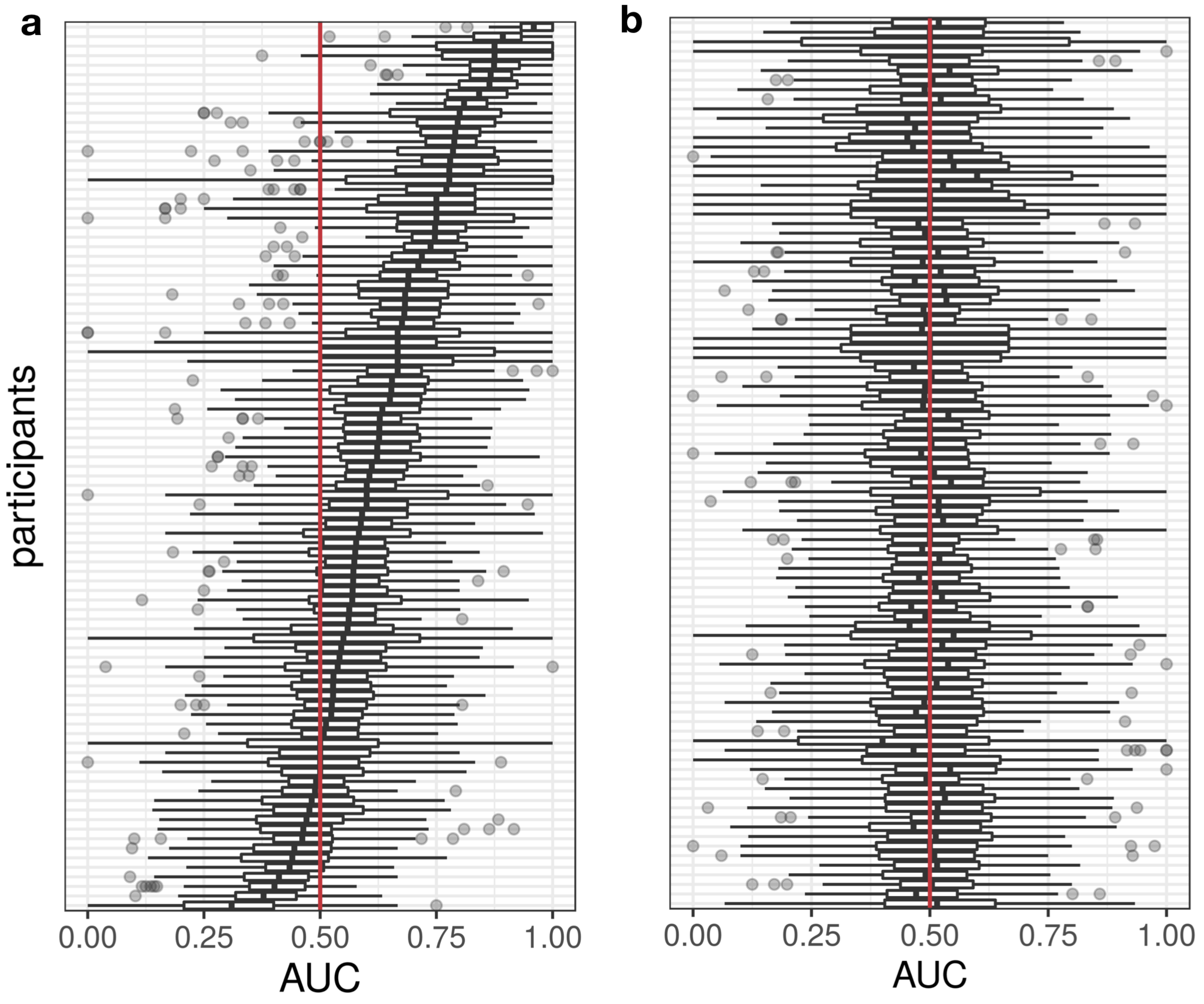
N-of-1

association between drifts from one's own digital "me"(baseline) and behavior anomalies



Time Series - ARIMA-based
Random Forest - Regression | Classification

Predicting mood for an individual based on their unique smartphone usage characteristics looks more promising



N ~ 120

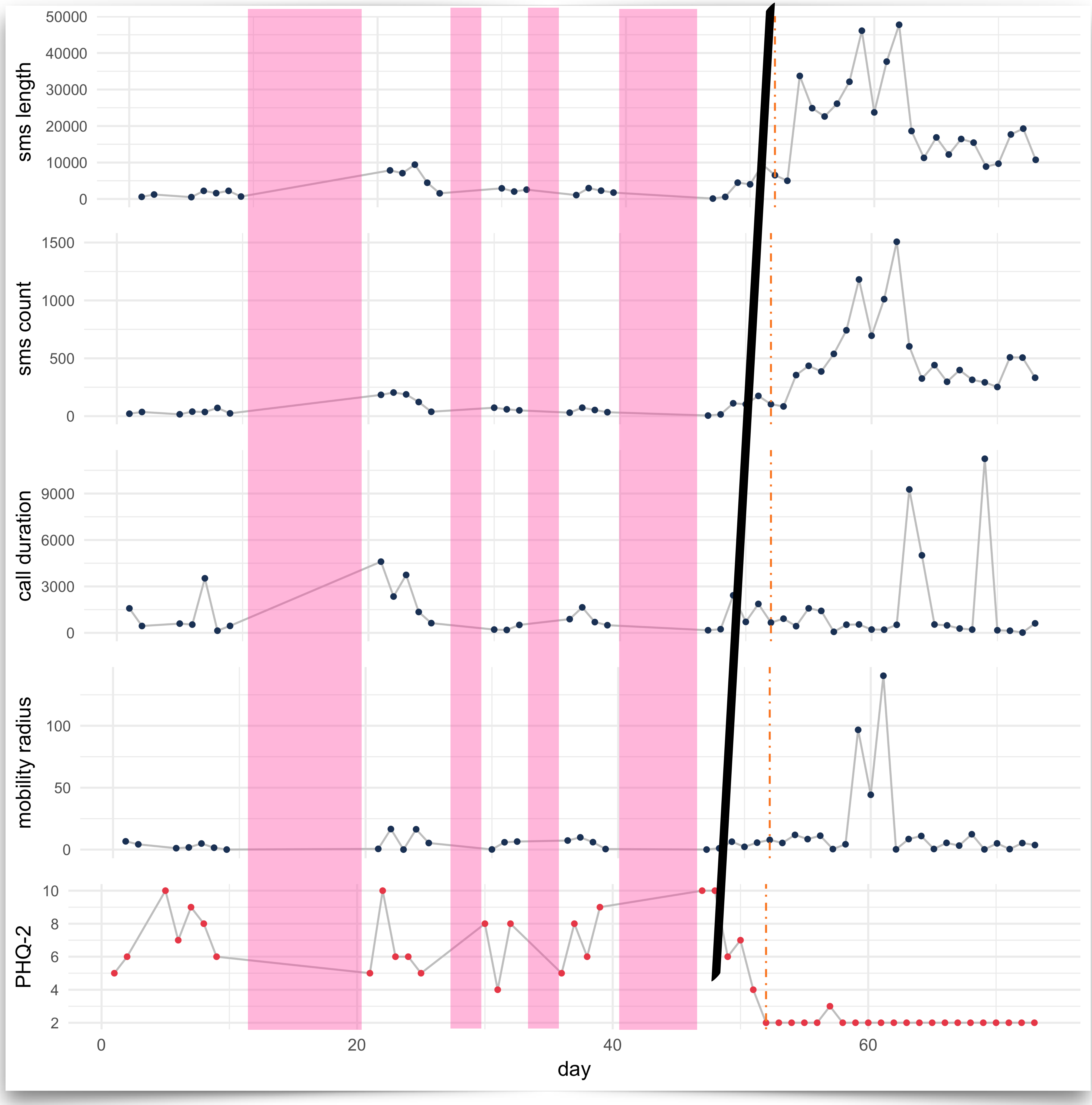
Random Forest - Classification **Time = 12 weeks**

Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, Arean PA. The accuracy of passive phone sensors in predicting daily mood. **Depress Anxiety. 2019**; doi:10.1002/da.22822

Jane's 12 week data profile shows distinct patterns

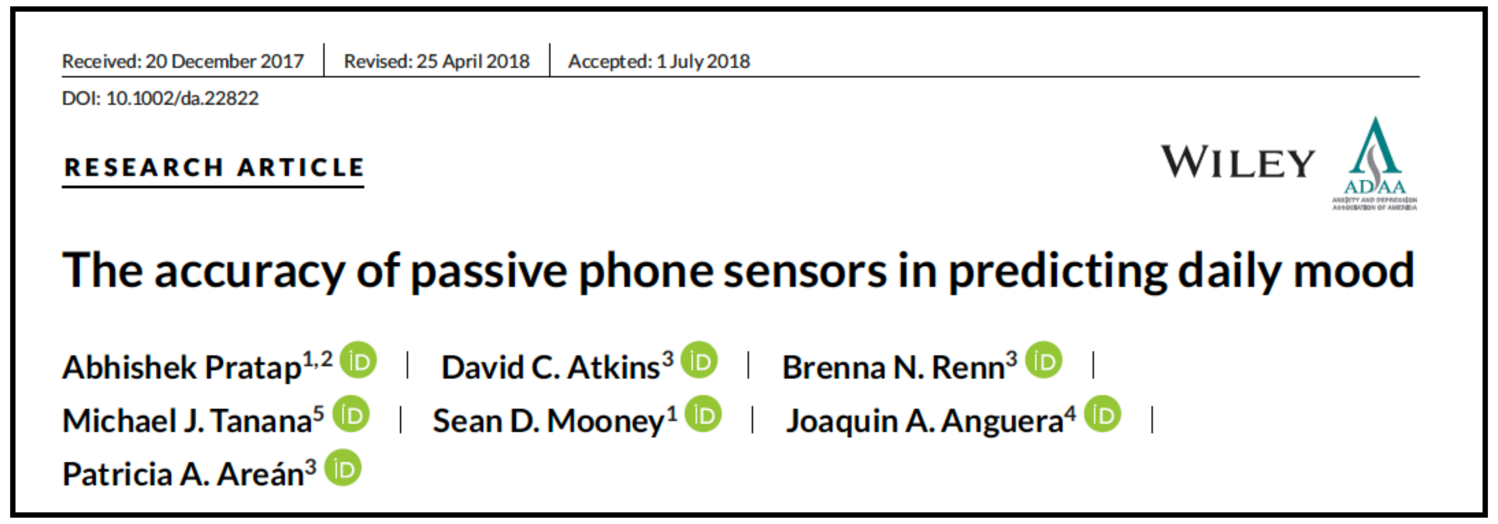
Social & Physical Mobility

Daily mood



Missing Data

- Missing at random
- Technical noise
- Biological signal



Received: 20 December 2017 | Revised: 25 April 2018 | Accepted: 1 July 2018
DOI: 10.1002/da.22822

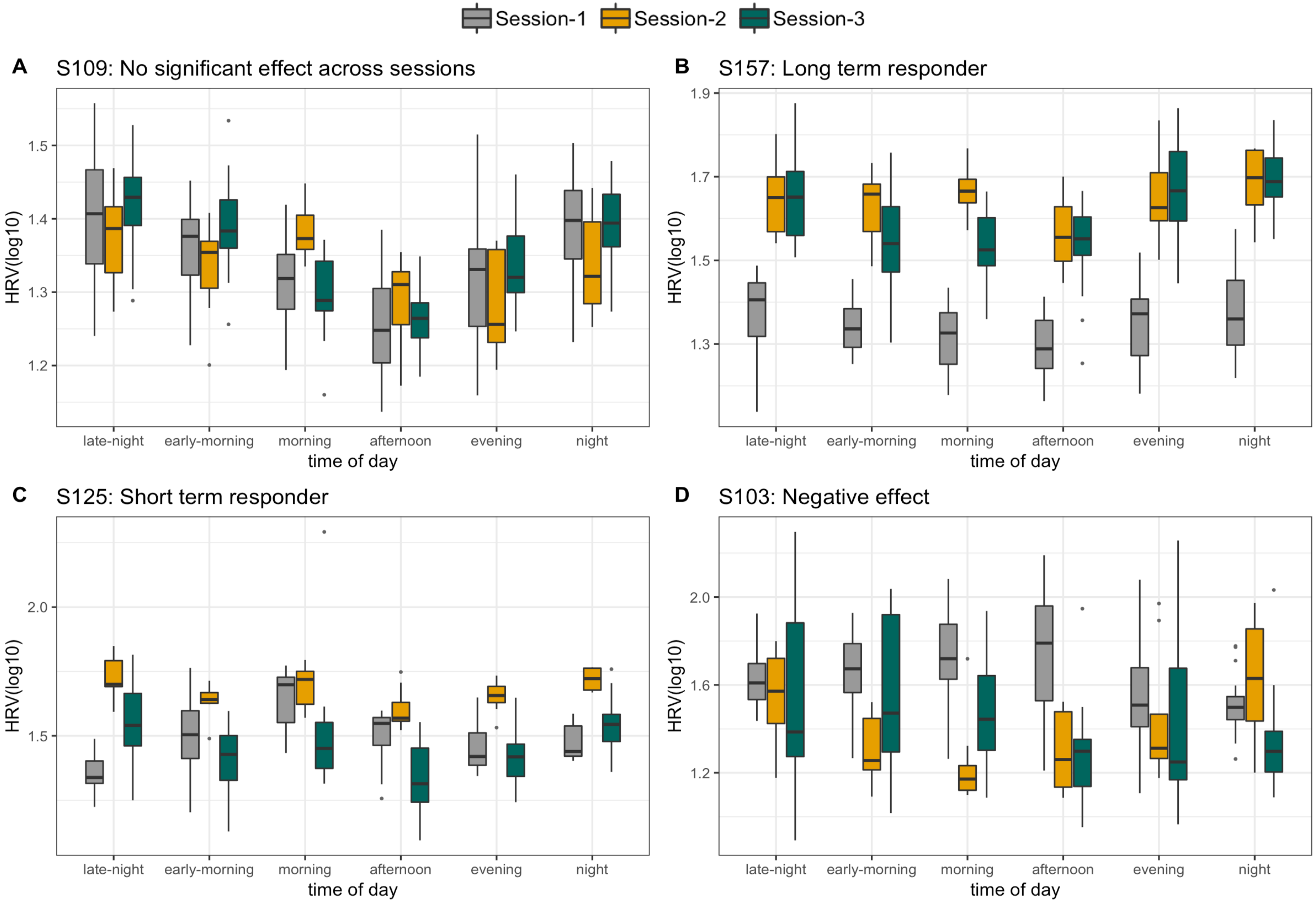
RESEARCH ARTICLE

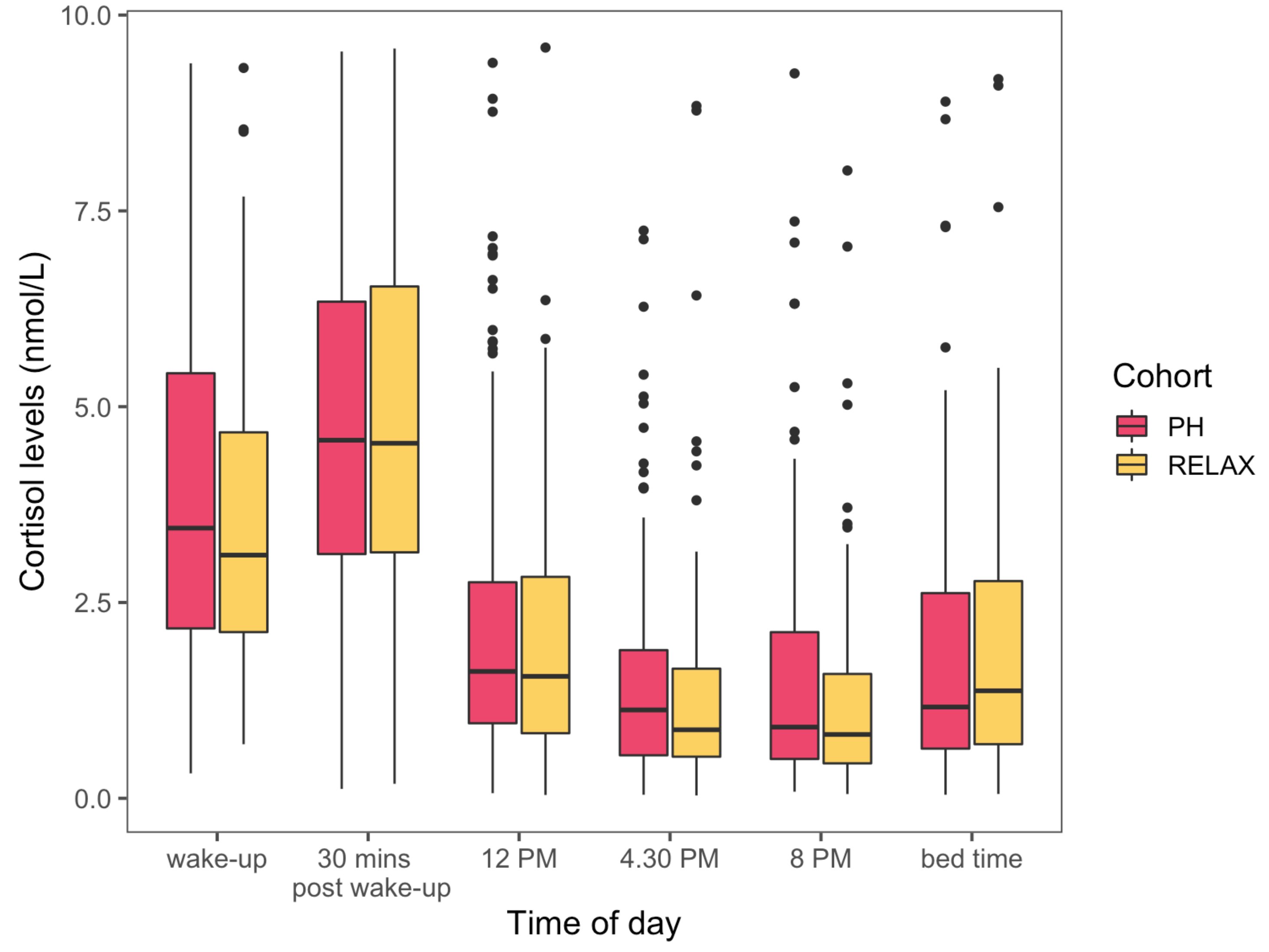
WILEY

The accuracy of passive phone sensors in predicting daily mood

Abhishek Pratap[1,2] | David C. Atkins[3] | Brenna N. Renn[3] | Michael J. Tanana[5] | Sean D. Mooney[1] | Joaquin A. Anguera[4] | Patricia A. Areán[3]

Individualized differences in physiological response to relaxation intervention






frontiers in Cardiovascular Medicine

ORIGINAL RESEARCH
published: 31 July 2020
doi: 10.3389/fcvm.2020.00120

Changes in Continuous, Long-Term Heart Rate Variability and Individualized Physiological Responses to Wellness and Vacation Interventions Using a Wearable Sensor

OPEN ACCESS

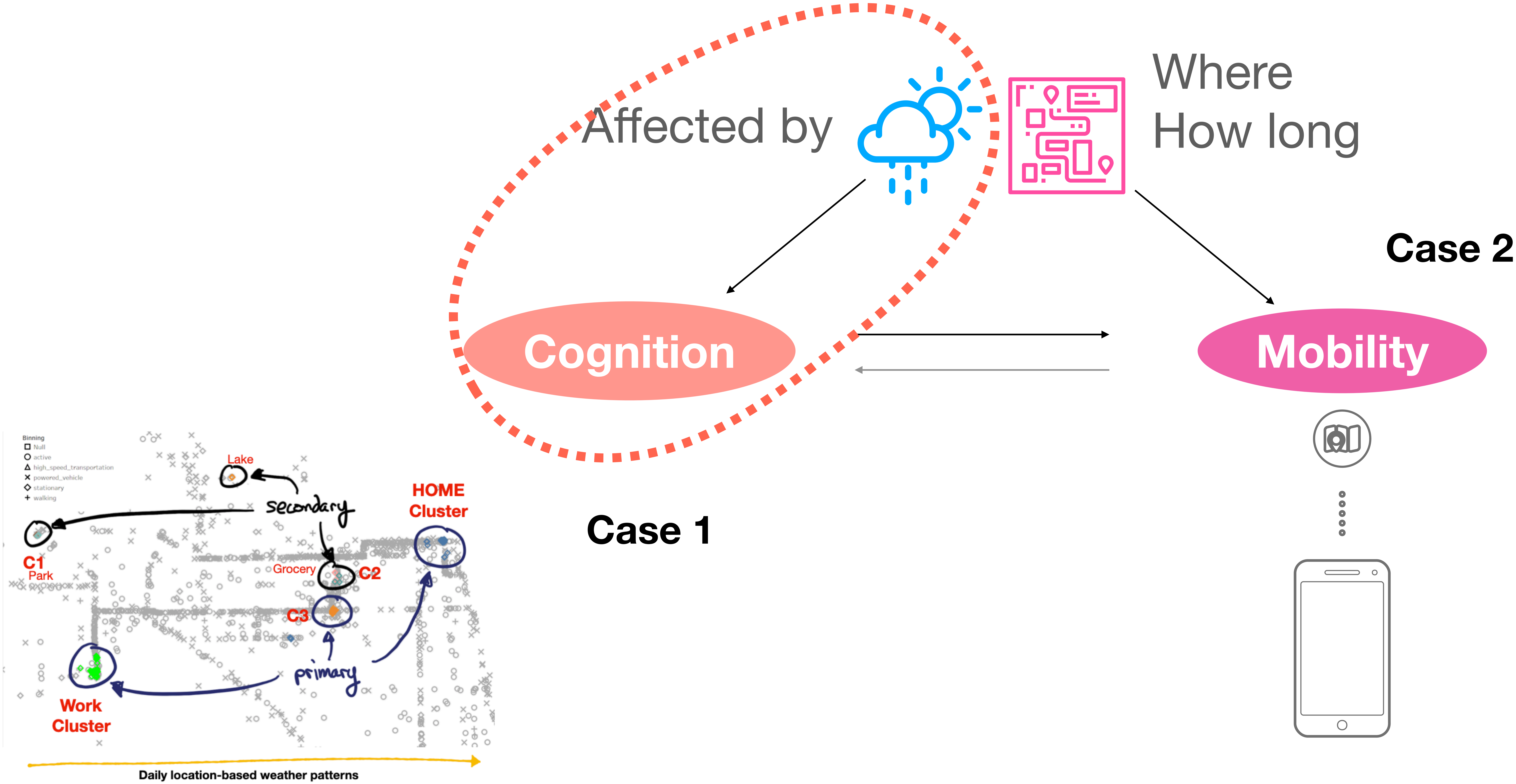
Edited by:

Abhishek Pratap[1,2], Steve Steinhubl[3], Elias Chaibub Neto[1], Stephan W. Wegerich[4], Christine Tara Peterson[5], Lizzy Weiss[6], Sheila Patel[5,7], Deepak Chopra[5,6] and Paul J. Mills[5]*

Mixed-effect model p-values < .001

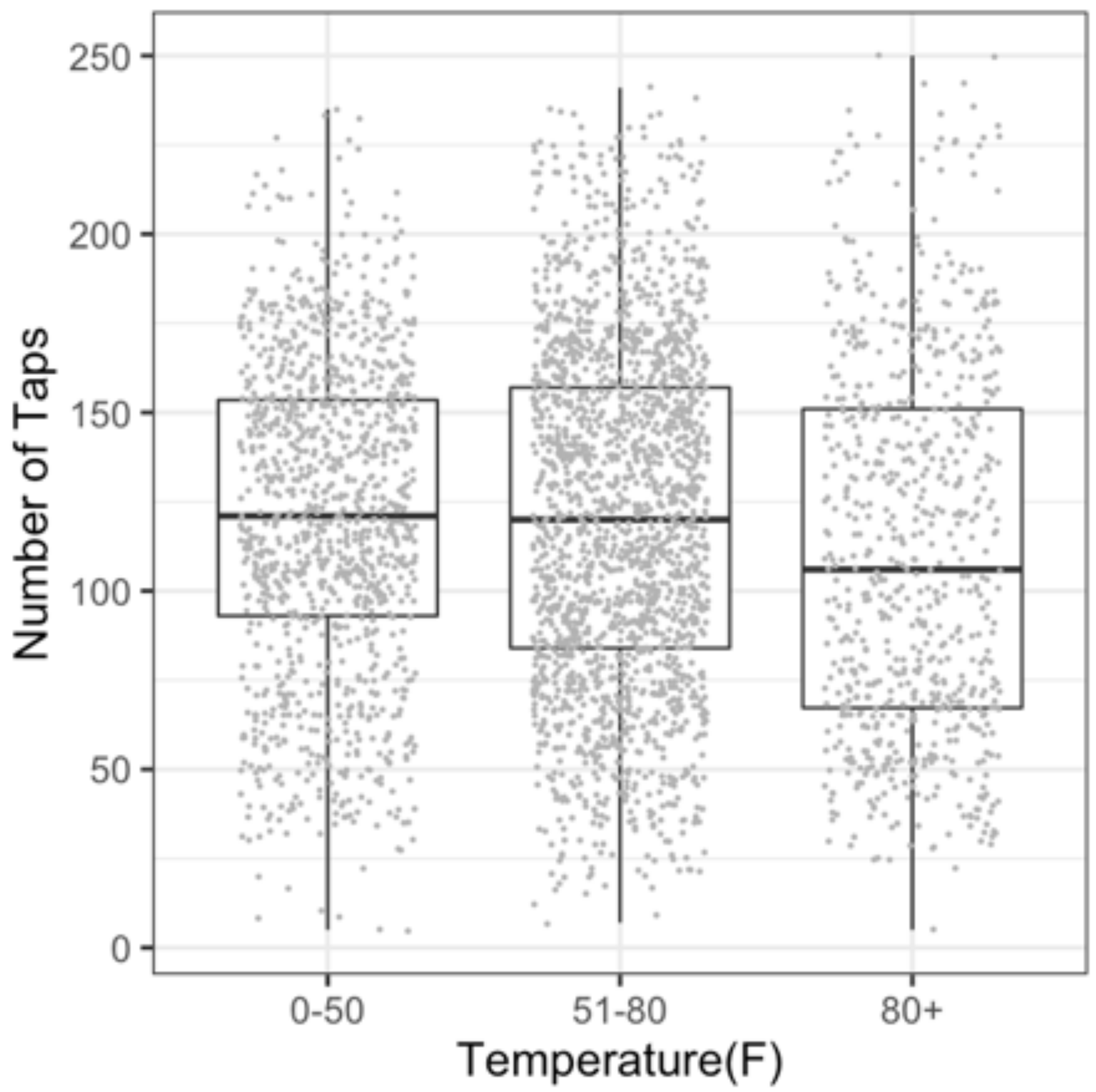
1

Context Matters

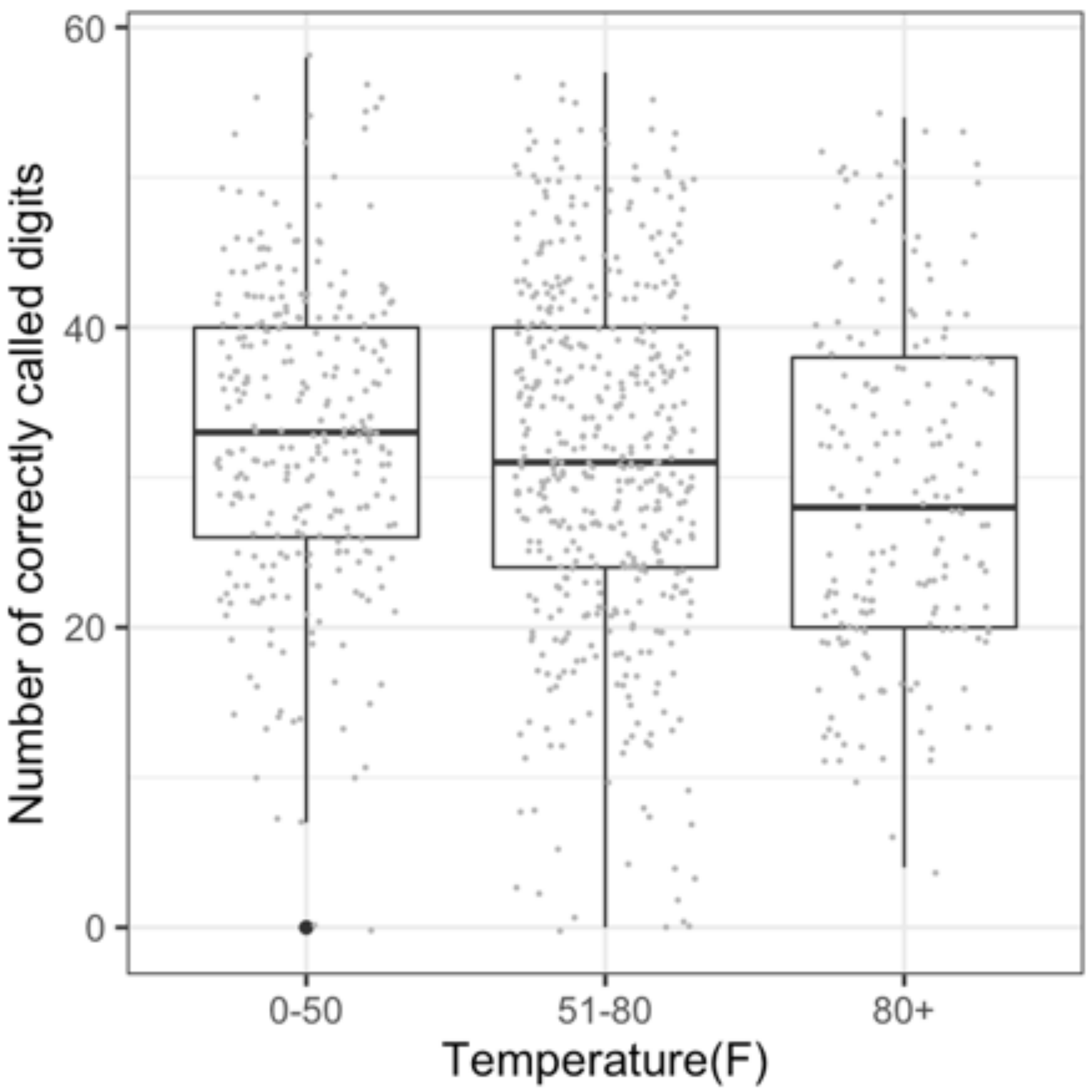


In MS external environmental factors linked to disease symptoms and triggers can impact participants performance in sensor-based active tasks

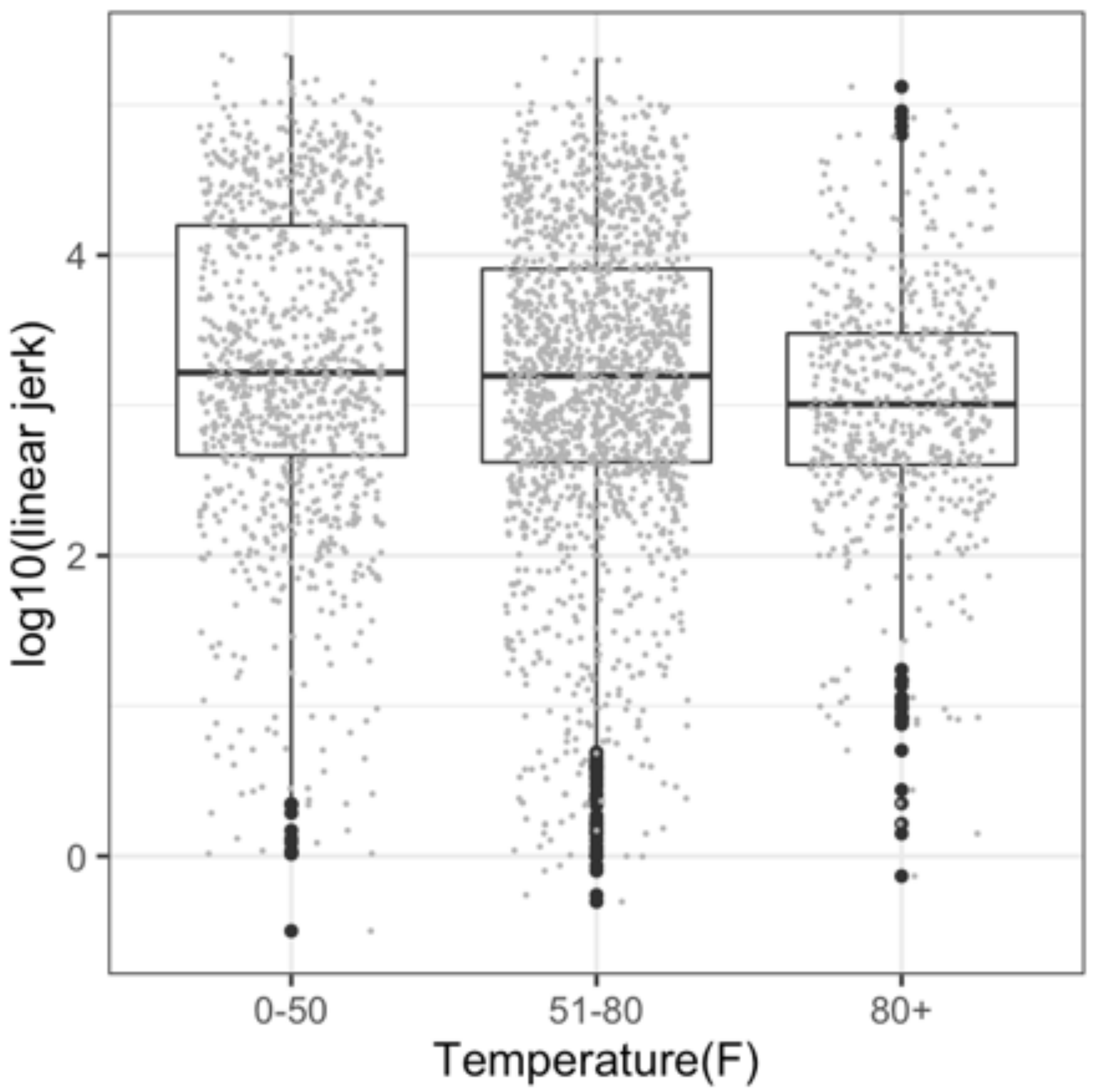
Fine motor control



Cognition
(vDSST)



Tremor



Pratap A, Grant D, Vegesna A, Tummalacherla M, Cohan S, Deshpande C, Mangravite L, Omberg L Evaluating the Utility of Smartphone-Based Sensor Assessments in Persons With Multiple Sclerosis in the Real-World Using an App (elevateMS): Observational, Prospective Pilot Digital Health Study JMIR Mhealth Uhealth

Mixed effect Models p-value < .01 or lower

Feasibility & Predictability

Example 3

Understanding real-world risk factors linked to suicidal behavior

Over 50+ years of research limited success in predicting suicide

Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research.

EXPORT Add To My List Database: PsycARTICLES Other

Franklin, Joseph C. Ribeiro, Jessica D. Fox, Kathryn R. Bentley, Kate H. Kleiman, Evan M. Huang, Xieyining
Musacchio, Katherine M. Jaroszewski, Adam C. Chang, Bernard P. Nock, Matthew K.

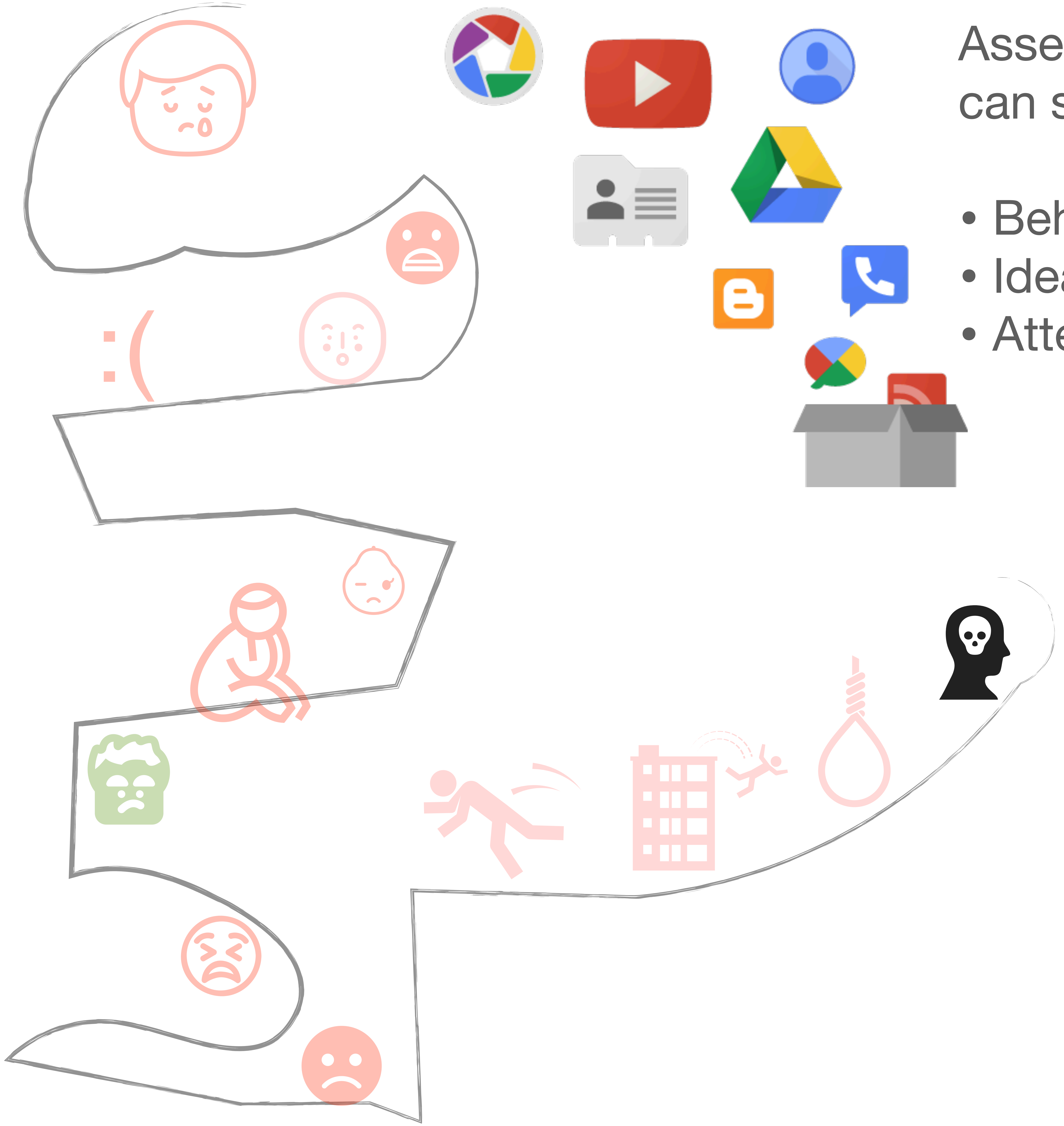
Psychological Bulletin
Editor Dolores AlbarracÃn, PhD
Journal TOC

Citation

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., . . . Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin, 143*(2), 187-232.

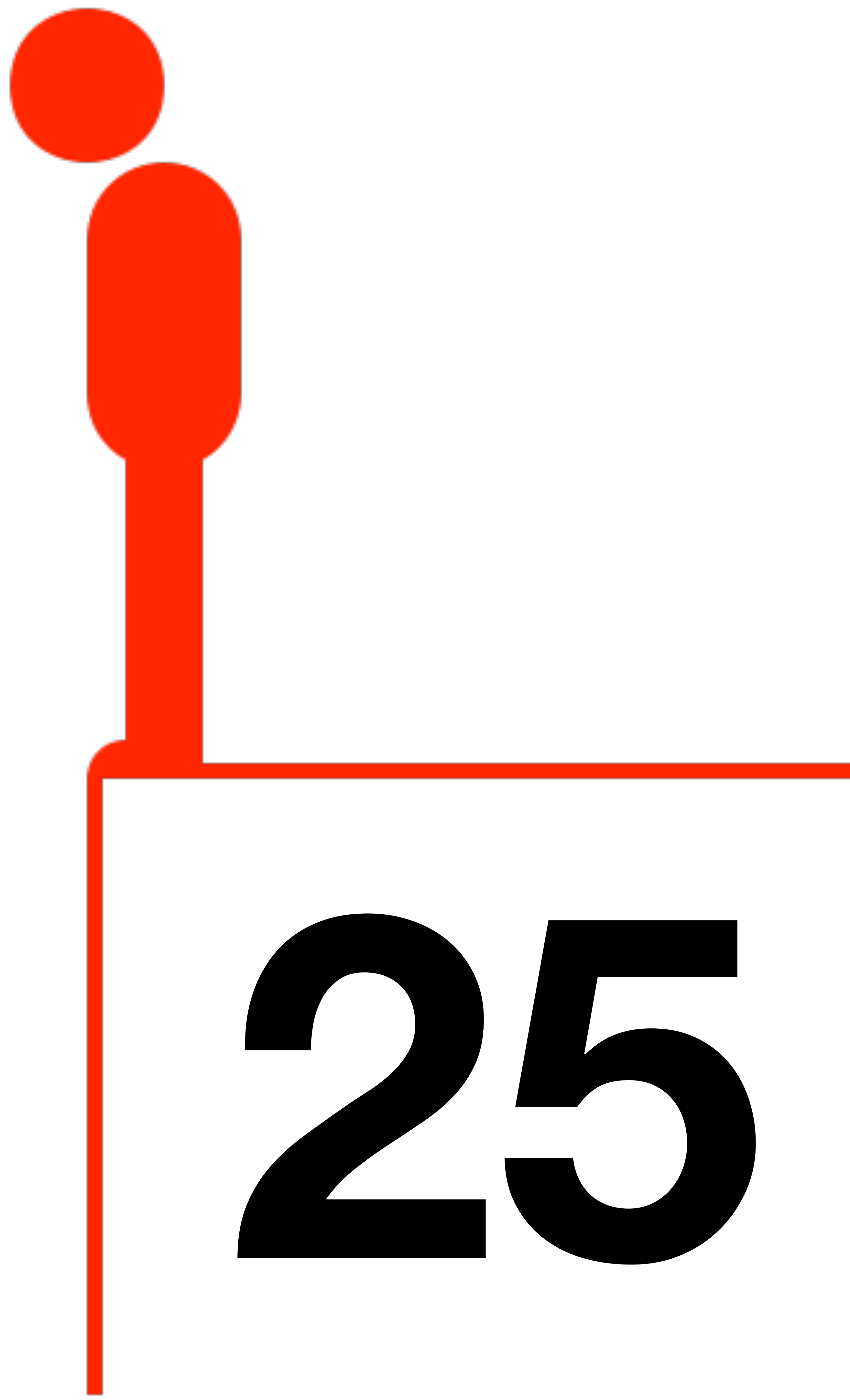
We know WHO might be at risk but not much about WHEN someone might be at highest risk of self- harm

One of the first steps to improving the prevention and treatment of STBs is to establish risk factors (i.e., longitudinal predictors). To provide a summary of current knowledge about risk factors, we conducted a meta-analysis of studies that have attempted to longitudinally predict a specific STB-related outcome. This included 365 studies (3,428 total risk factor effect sizes) from the past 50 years. The present random-effects meta-analysis produced several unexpected findings: across odds ratio, hazard ratio, and diagnostic accuracy analyses, prediction was only slightly better than chance for all outcomes; no broad category or subcategory accurately predicted far above chance levels; predictive ability has not improved across 50 years of research; studies rarely examined the combined effect of multiple risk factors; risk factors have been homogenous over time, with 5 broad categories accounting for nearly 80% of all risk factor tests; and the average study was nearly 10 years long, but longer studies did not

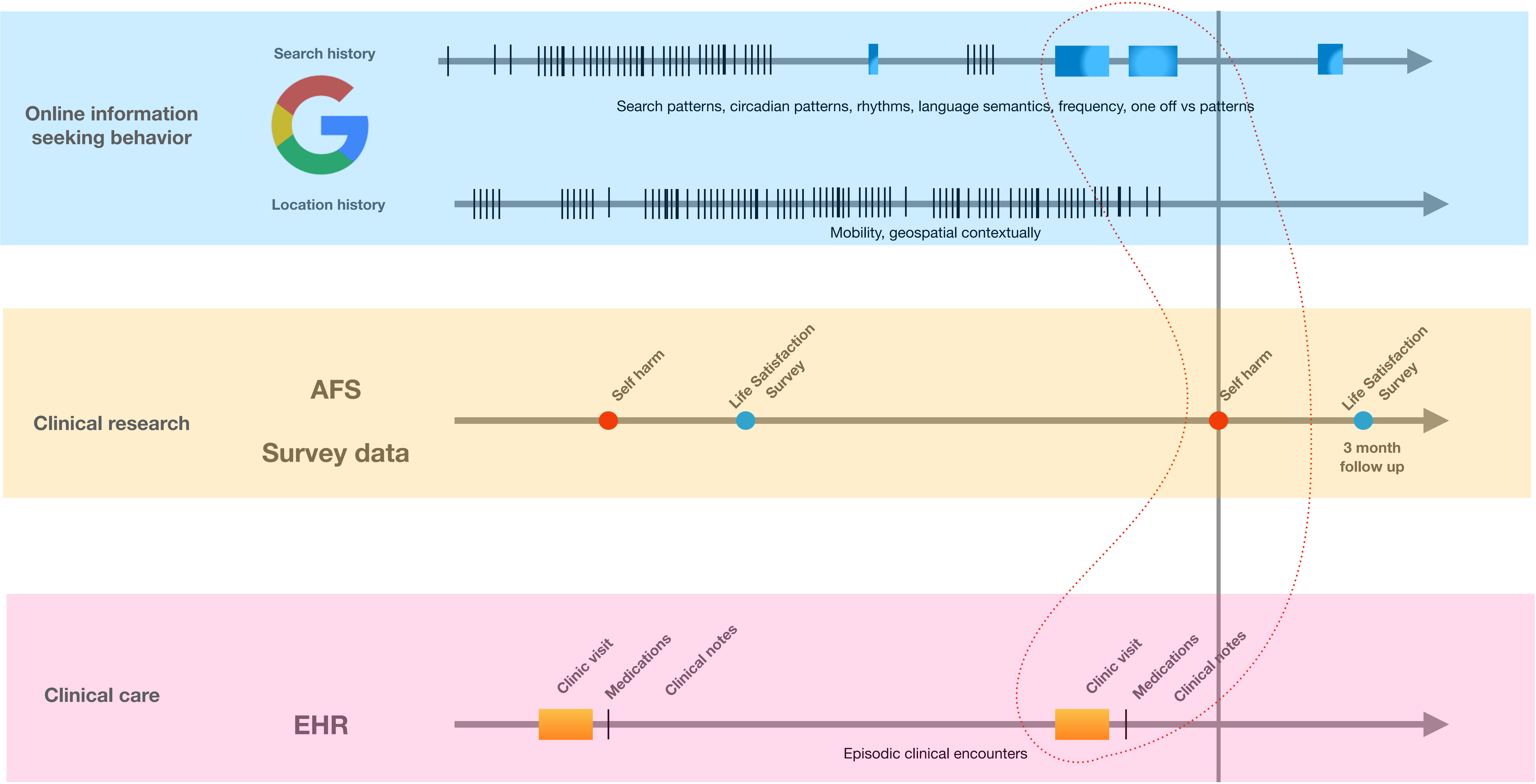


Assess if **online information seeking behavior** can surface risk factors related to suicide

- Behavior
- Ideation
- Attempt



Vertical real-world data integration



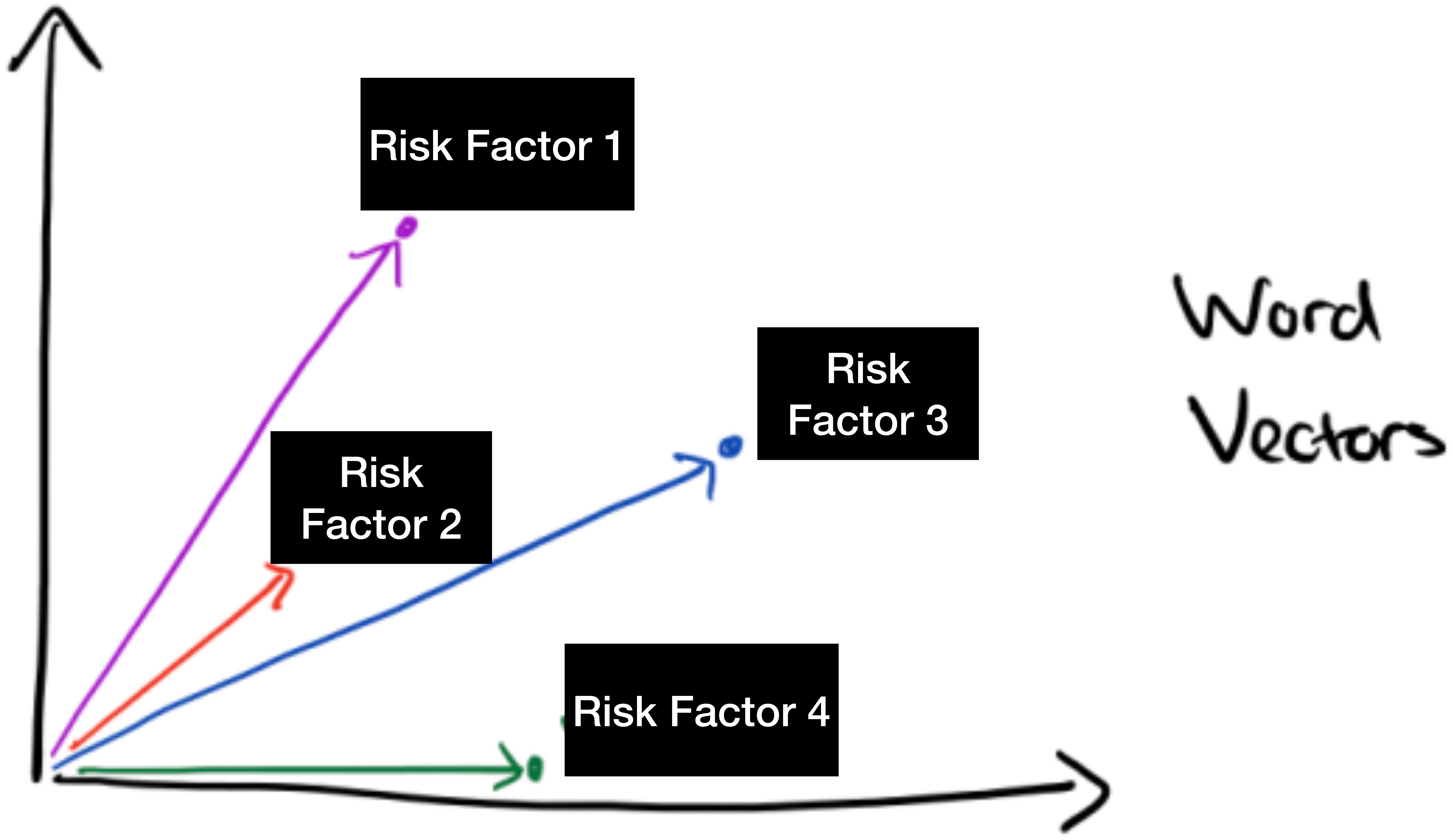
Using online searches to understand underlying thoughts

2016-11-03	NO WEB SEARCHES	SUICIDE ATTEMPT DATE
2016-11-04	how much does it cost to go to a psych ward	
2016-11-04	do you have to go to college to become a mathematician	
2016-11-04	i shouldn't be here in the psych ward	
2016-11-05	drawing flames	
2016-11-06	how to survive as a college dropout	
2016-11-06	jungle book cast	
2016-11-06	sombra info overwatch	

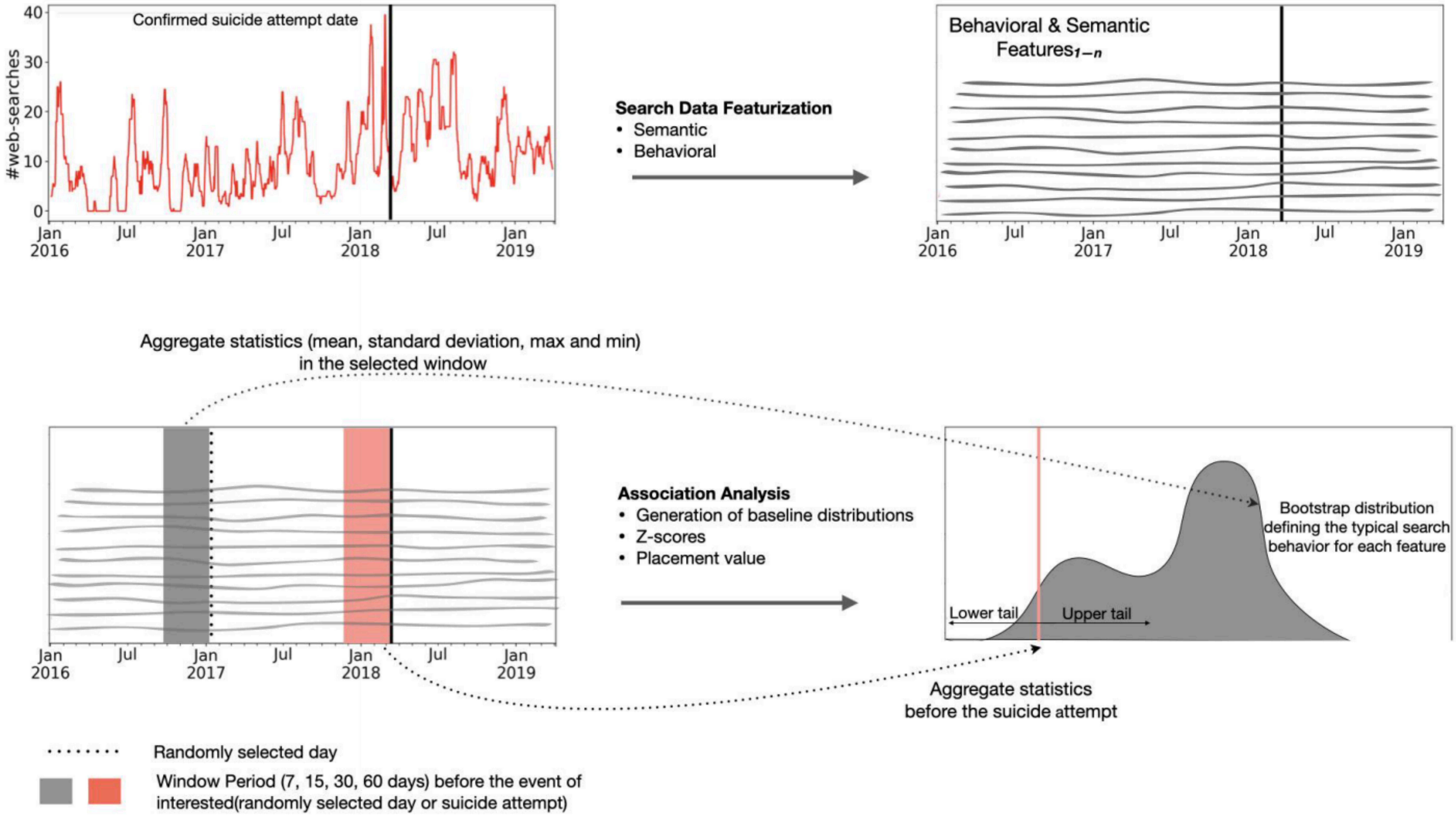
Areán PA*, **Pratap A***, Hsin H, Huppert TK, Hendricks KE, Heagerty PJ, Cohen T, Bagge C, Comtois KA Perceived Utility and Characterization of Personal Google Search Histories to Detect Data Patterns Proximal to a Suicide Attempt in Individuals Who Previously Attempted Suicide: Pilot Cohort Study **J Med Internet Res 2021;23(5):e27918**

Participants search queries were mapped to pre-identified 11 risk factors associated with suicide

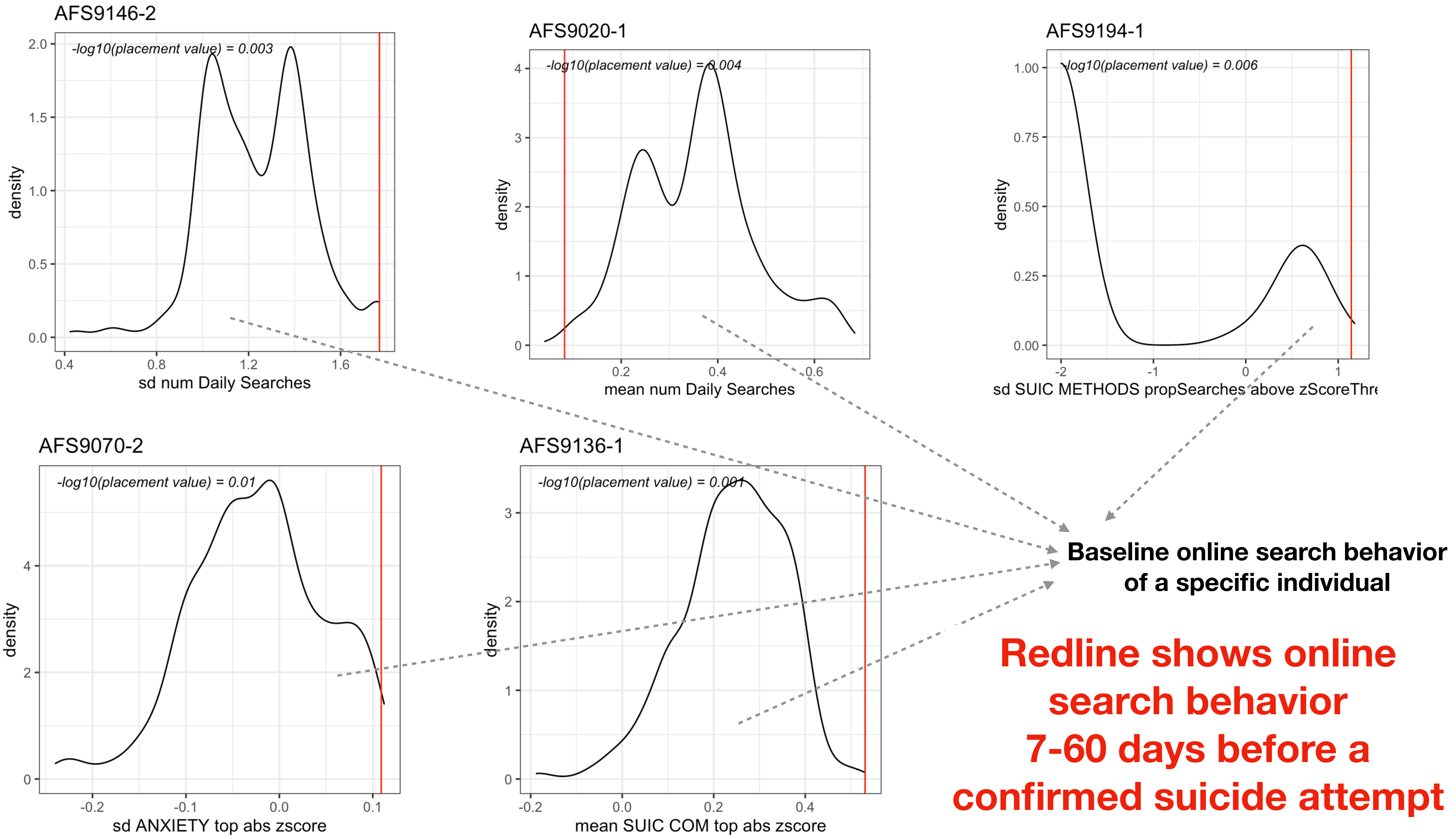
“how long does codeine take to hit”



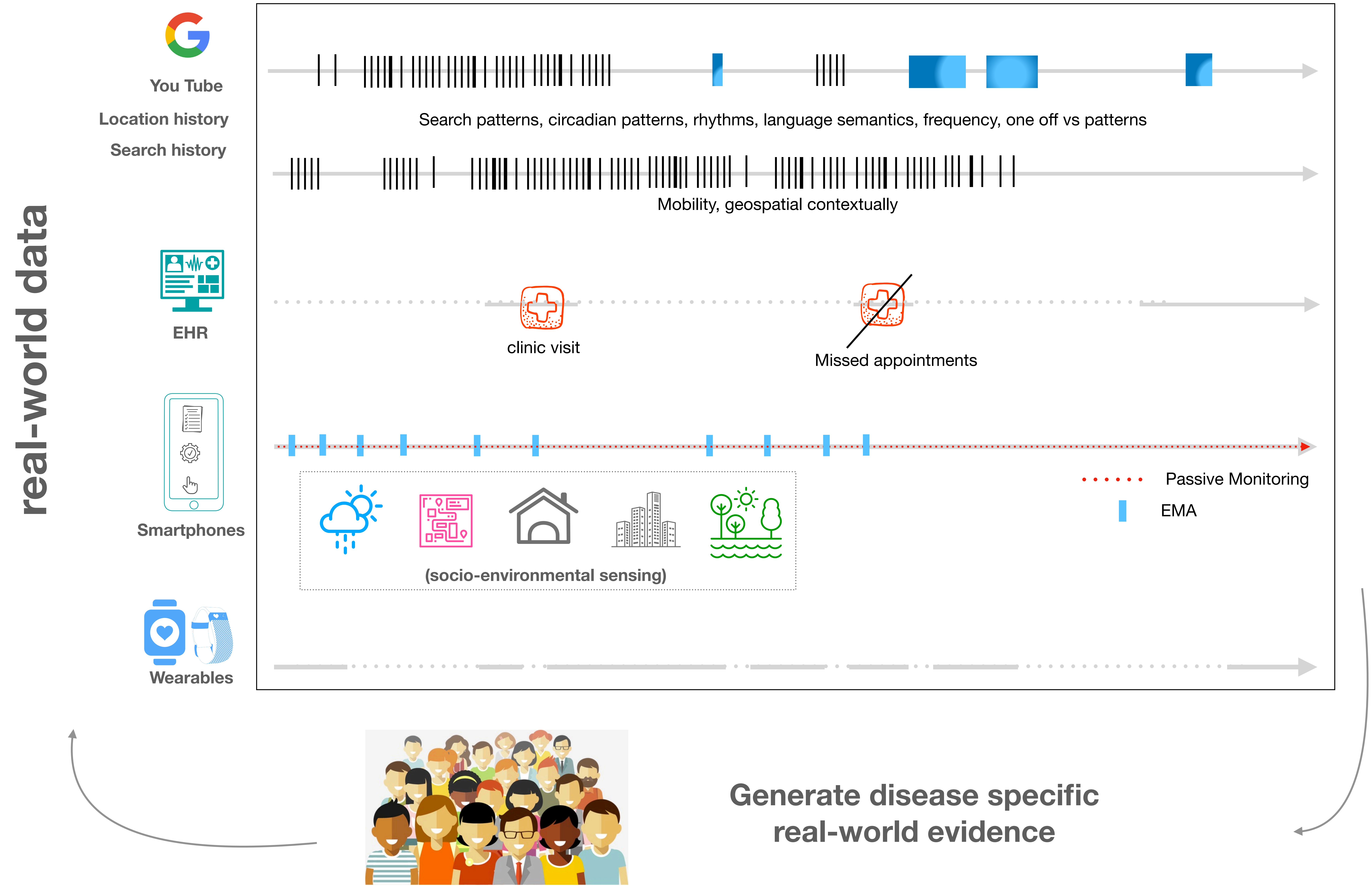
https://www.depends-on-the-definition.com/guide-to-word-vectors-with-gensim-and-keras/



Confirmed suicide attempt date
#web-searches
40
30
20
10
0
Jan 2016
Jul
Jan 2017
Jul
Jan 2018
Jul
Jan 2019
Search Data Featurization
• Semantic
• Behavioral
Behavioral & Semantic Features$_{1-n}$
Aggregate statistics (mean, standard deviation, max and min) in the selected window
Association Analysis
• Generation of baseline distributions
• Z-scores
• Placement value
Bootstrap distribution defining the typical search behavior for each feature
Lower tail
Upper tail
Aggregate statistics before the suicide attempt
Randomly selected day
Window Period (7, 15, 30, 60 days) before the event of interested(randomly selected day or suicide attempt)


AFS9146-2
-log10(placement value) = 0.003
density
sd num Daily Searches
AFS9020-1
-log10(placement value) = 0.004
density
mean num Daily Searches
AFS9194-1
-log10(placement value) = 0.006
density
sd SUIC METHODS propSearches above zScoreThre
AFS9070-2
-log10(placement value) = 0.01
density
sd ANXIETY top abs zscore
AFS9136-1
-log10(placement value) = 0.001
density
mean SUIC COM top abs zscore
Baseline online search behavior of a specific individual
Redline shows online search behavior 7-60 days before a confirmed suicide attempt

Triangulate new sources of real-world data



Why

Digital Health for Mental Health

Opportunities

How

Using digital health to assess CNS symptoms
"in the real world"

Feasibility & Predictability

When

If we build tech, communities will embrace it

Challenges & Solutions

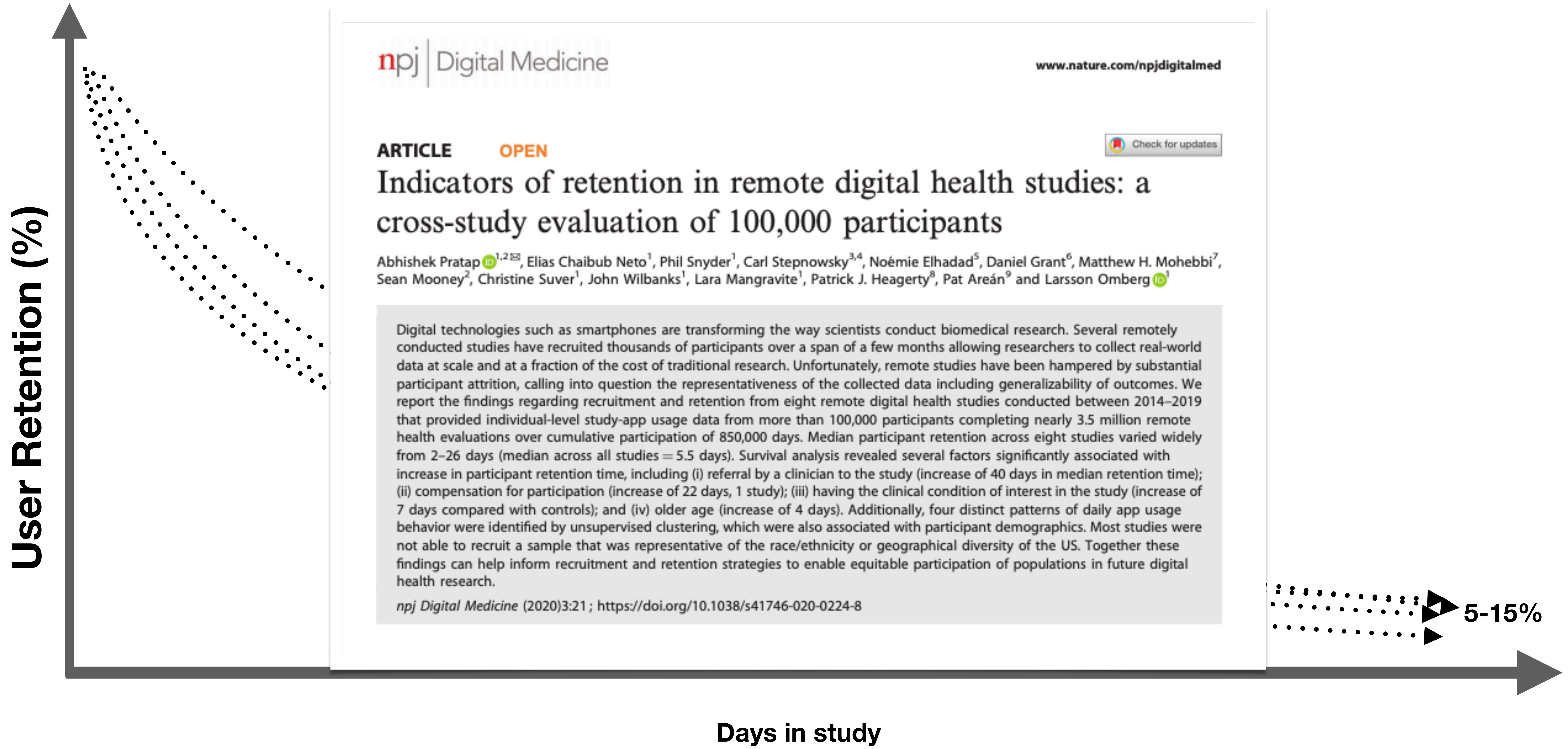
https://www.raconteur.net/smartphone-addiction-in-5-charts/

Just because you can ask, you shouldn't

Just because you asked, they might not respond

http://www.justinhallcomics.com/break-your-smartphone-addiction/

Building digital tech alone is not enough


npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

Check for updates

Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants

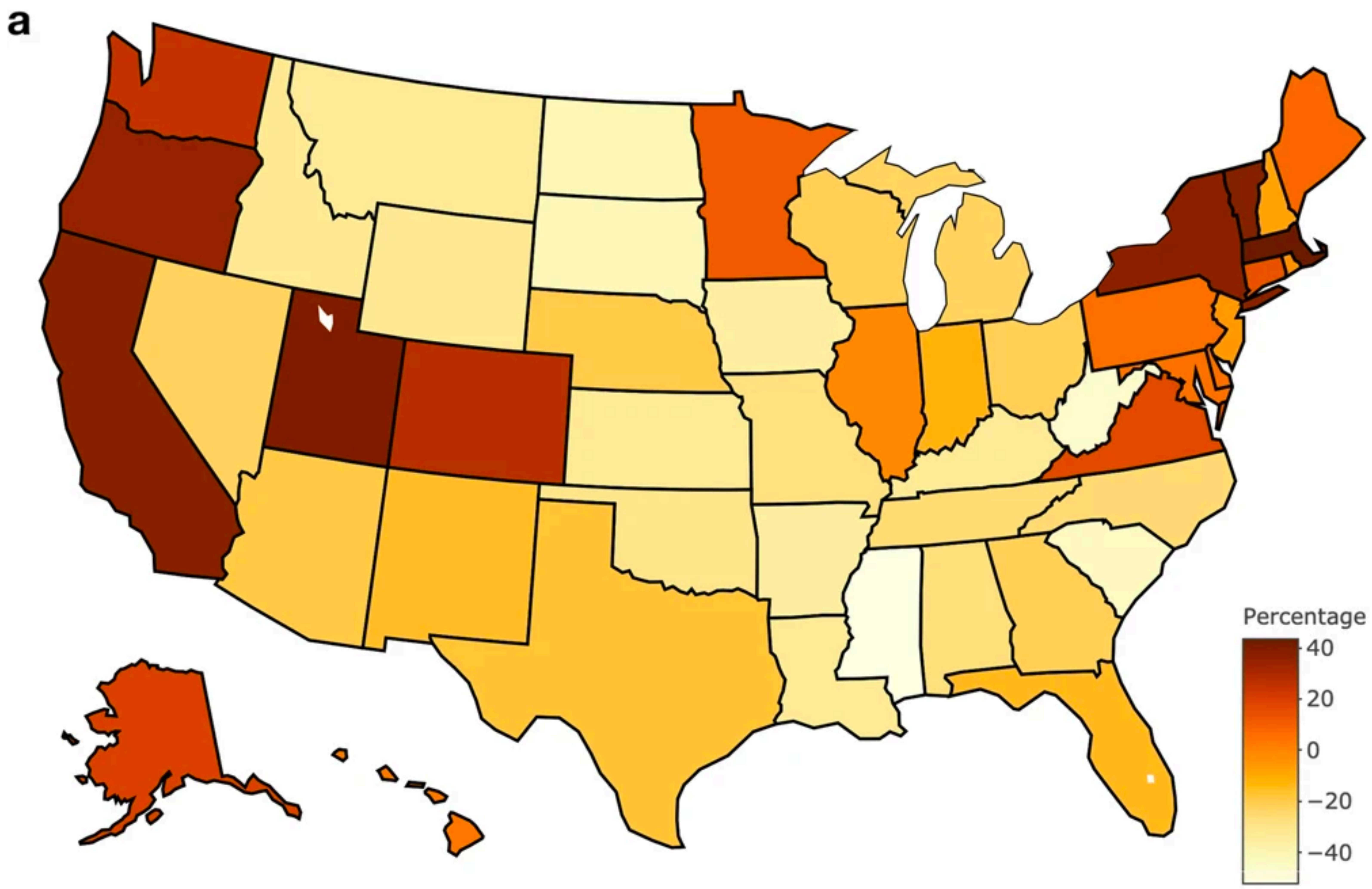
Abhishek Pratap[1,2✉], Elias Chaibub Neto[1], Phil Snyder[1], Carl Stepnowsky[3,4], Noémie Elhadad[5], Daniel Grant[6], Matthew H. Mohebbi[7], Sean Mooney[2], Christine Suver[1], John Wilbanks[1], Lara Mangravite[1], Patrick J. Heagerty[8], Pat Areán[9] and Larsson Omberg[1]

Digital technologies such as smartphones are transforming the way scientists conduct biomedical research. Several remotely conducted studies have recruited thousands of participants over a span of a few months allowing researchers to collect real-world data at scale and at a fraction of the cost of traditional research. Unfortunately, remote studies have been hampered by substantial participant attrition, calling into question the representativeness of the collected data including generalizability of outcomes. We report the findings regarding recruitment and retention from eight remote digital health studies conducted between 2014–2019 that provided individual-level study-app usage data from more than 100,000 participants completing nearly 3.5 million remote health evaluations over cumulative participation of 850,000 days. Median participant retention across eight studies varied widely from 2–26 days (median across all studies = 5.5 days). Survival analysis revealed several factors significantly associated with increase in participant retention time, including (i) referral by a clinician to the study (increase of 40 days in median retention time); (ii) compensation for participation (increase of 22 days, 1 study); (iii) having the clinical condition of interest in the study (increase of 7 days compared with controls); and (iv) older age (increase of 4 days). Additionally, four distinct patterns of daily app usage behavior were identified by unsupervised clustering, which were also associated with participant demographics. Most studies were not able to recruit a sample that was representative of the race/ethnicity or geographical diversity of the US. Together these findings can help inform recruitment and retention strategies to enable equitable participation of populations in future digital health research.

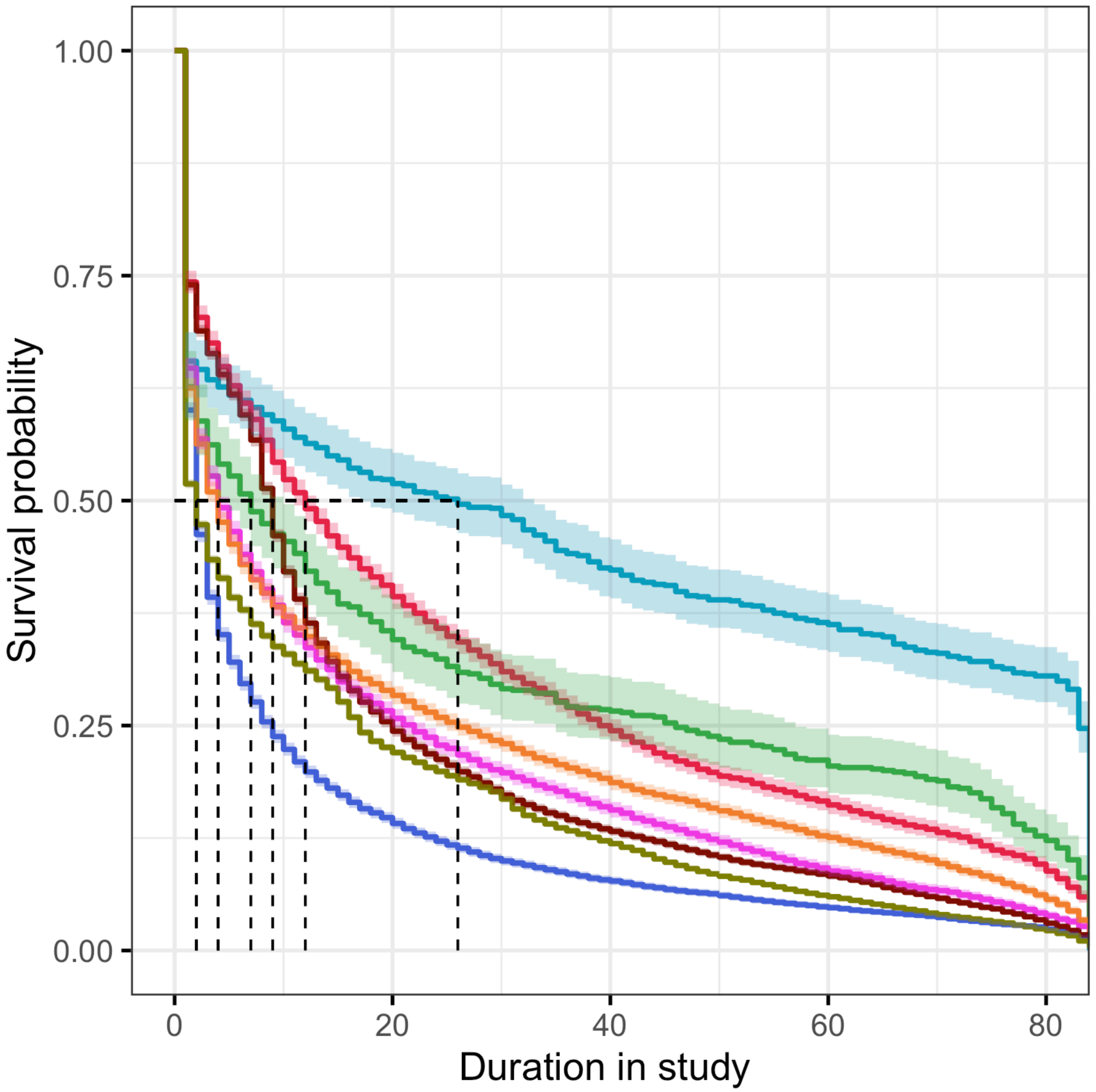
npj Digital Medicine (2020)3:21 ; https://doi.org/10.1038/s41746-020-0224-8

A Pratap, E Neto, P Snyder, C Stepnowsky, N Elhadad, D Grant, M Mohebbi, S Mooney, C Suver, J Wilbanks, L Mangravite, P Heagerty, P Arean, L Omberg - Indicators of retention in remote digital health studies: A cross-study evaluation of 100,000 participants **| Nature Digital Medicine, 2020**

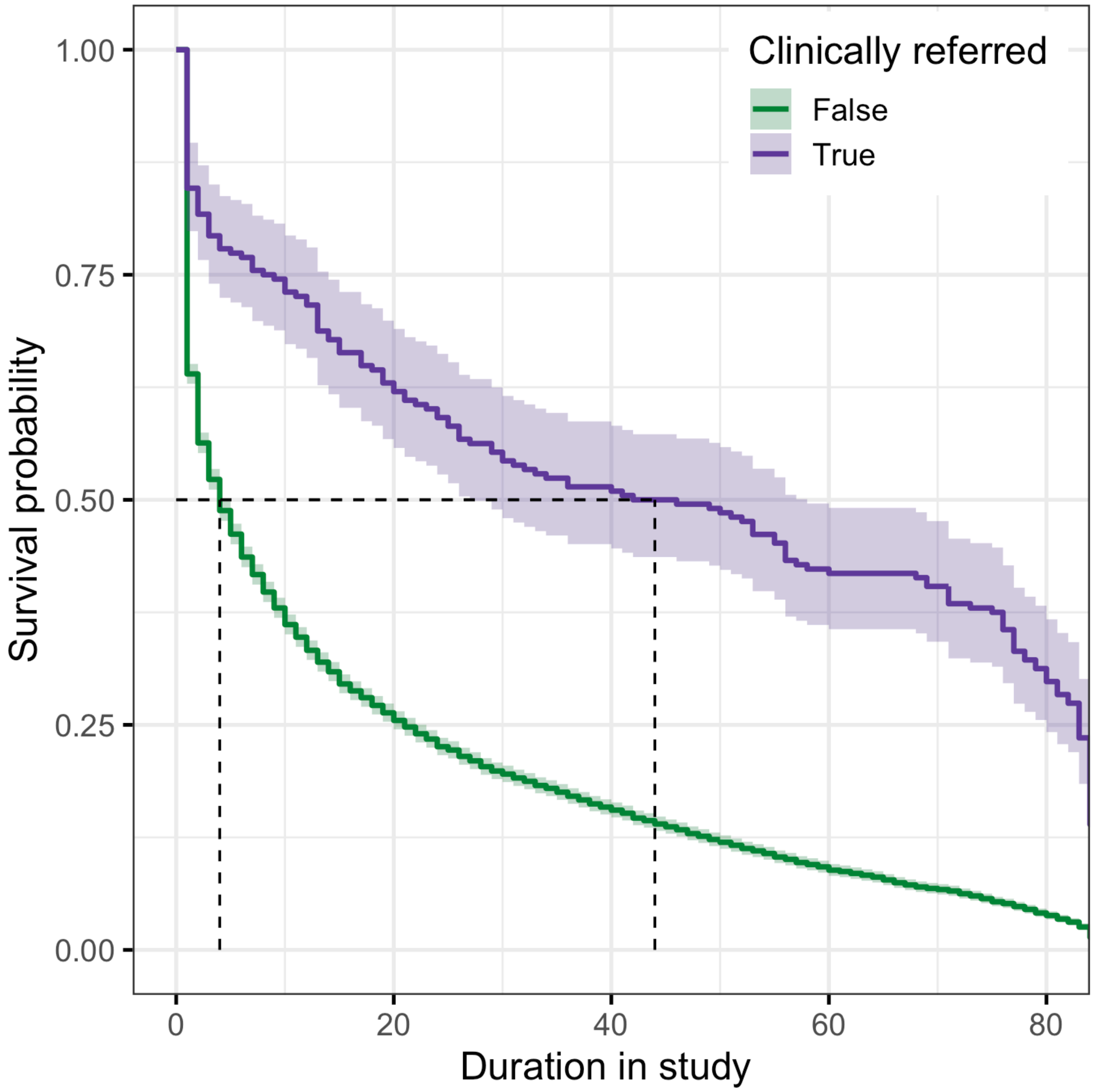
Significant Digital Divide



A Pratap, E Neto, P Snyder, C Stepnowsky, N Elhadad, D Grant, M Mohebbi, S Mooney, C Suver, J Wilbanks, L Mangravite, P Heagerty, P Arean, L Omberg - Indicators of retention in remote digital health studies: A cross-study evaluation of 100,000 participants | **Nature Digital Medicine, 2020**



- ~50% of participants leave the studies within the first 7-10 days
- Targeted Comms during the first week may help



- Participants referred by clinical sites/partners engage for significantly longer time

C1 - dedicated users

C3 - moderate users

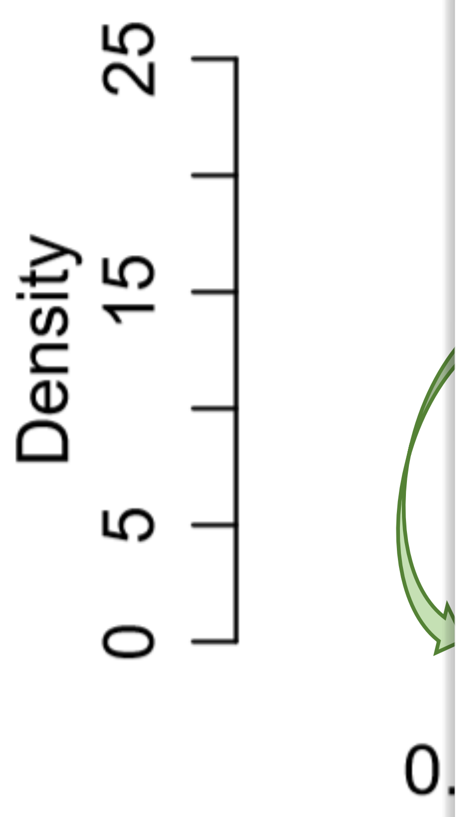
C2 - high utilizers

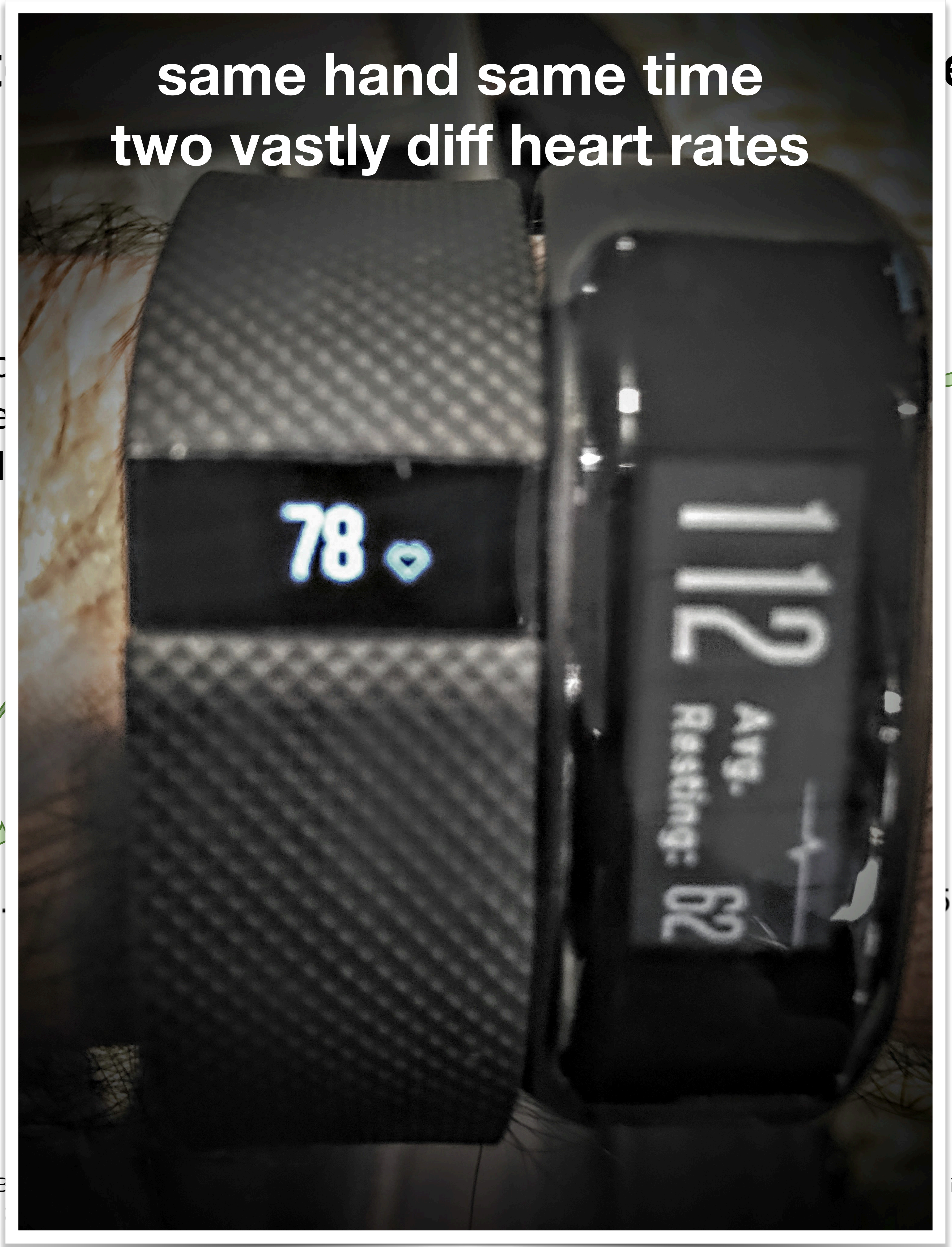
C4 - sporadic users

Confounding charact... ely impact robustness of inference from Machi...

(perfo...
to the...
signal...

m_o





*Elias Neto , **A. Pratap** , T. Perumal , M. Tummalacherla, P. Snyder, B. E... instead of the disease signal: a quantitative approach to detect identify confounding in machine learning diagnostic applications:**Nature Digital Medicine,** ...

Elias Chaibub Neto, **A. Pratap**, T. Perumal, M Tummalacherla, B Bot, L Mangravite and L Omberg - A permutation approach to assess confounding in machine learning applications for digital health. **KDD 2019**

JAMA Network | Open

Original Investigation | Health Informatics

Contemporary Views of Research Participant Willingness to Participate and Share Digital Data in Biomedical Research

Abhishek Pratap, MS; Ryan Allred, BA; Jaden Duffy, BA; Donovan Rivera, MSW; Heather Sophia Lee, PhD; Brenna N. Renn, PhD; Patricia A. Areán, PhD

Abstract

IMPORTANCE Using social media to recruit participants is a common and cost-effective practice. Willingness to participate (WTP) in biomedical research is a function of trust in the scientific team, which is closely tied to the source of funding and institutional connections.

OBJECTIVE To determine whether WTP and willingness to share social media data are associated with the type of research team and online recruitment platform.

DESIGN, SETTING, AND PARTICIPANTS This mixed-methods longitudinal survey and qualitative study was conducted over 2 points (T1 and T2) using Amazon's Mechanical Turk (MTurk) platform. Participants were US adults aged 18 years or older who use at least 1 social media platform. Recruitment was stratified to match race/ethnicity proportions of the 2010 US Census. The volunteer sample consisted of 914 participants at T1, and 655 participants completed the follow-up survey 5 months later (T2).

MAIN OUTCOMES AND MEASURES Outcomes were (1) past experience with online research and sharing social media data for research; (2) WTP in research advertised online; (3) WTP in a study sponsored by a pharmaceutical company, a university, or a federal agency; and (4) willingness to share social media data. Opinions were solicited regarding the European Union's General Data Protection Regulation statute, which came into effect between T1 and T2.

RESULTS Of 914 participants completing the first survey (T1), 604 (66.1%) were aged 18 to 39 years and 494 (54.0%) were female. Of these, 655 participants (71.7%) responded at T2. While 680 participants (74.4%) indicated WTP in biomedical research, only 454 (49.3%) were willing to share their social media data. Participants were significantly less likely to participate in federally sponsored

Key Points

Question Are people willing participate in research adve internet, and is willingness t associated with type of stud

Findings This mixed-metho and qualitative study of 914 respondents indicated that more likely to participate an social media data with resea university-led research stud studies conducted by the U government or pharmaceuti companies. However, only 4 indicated they would share media data at all.

Meaning These findings in researchers may face challe recruiting representative sa recruiting from internet plat

+ Supplemental content

Author affiliations and article infor listed at the end of this article.

Understanding Participant Needs for Engagement and Attitudes towards Passive Sensing in Remote Digital Health Studies

Samantha Kolovson
Human Centered Design & Engineering, University of Washington
kolovson@uw.edu

Abhishek Pratap
Biomedical Informatics & Medical Education, University of Washington
Sage Bionetworks
apratap@sagebase.org

Ryan Allred
Psychiatry & Behavioral Sciences, University of Washington
rallred@uw.edu

Sean A. Munson
Human Centered Design & Engineering, University of Washington
smunson@uw.edu

ABSTRACT

Digital psychiatry is a rapidly growing area of research. Mobile assessment, including passive sensing, could improve research into human behavior and may afford opportunities for rapid treatment delivery. However, retention is poor in remote studies of depressed populations in which frequent assessment and passive monitoring are required. To improve engagement and understanding participant needs overall, we conducted semi-structured interviews with 20 people representative of a depressed population in a major metropolitan area. These interviews elicited feedback on strategies for long-term remote research engagement and attitudes towards passive data collection. Our results found participants were uncomfortable sharing vocal samples, need researchers to take a more active role in supporting their understanding

based on the c sider a more e a choice abou digital health design to imp

CCS CONC

• Human-ce uation meth

KEYWORDS

mental healt mote study; i

ACM Referen
Samantha Kolo

ORIGINAL RESEARCH ARTICLE

Using Real-world Data for Decision Support: Recommendations from a Primary Care Provider Survey

Patricia A Areán, PhD[1,2]; Emily C Friedman, MID, CPE[2]; Abhishek Pratap, PhD[3]; Ryan Allred, BA[1]; Jaden Duffy, BA[1]; Sara Gille, MPH[4]; Shelley Reetz, BS[4]; Erin Keast, MPH[4]; Gregory Clarke, PhD[4]

Perm J 2021;25:20.213

E-pub: 03/01/2021

https://doi.org/10.7812/TPP/20.213

ABSTRACT

Introduction: The use of data from wearable sensors, smartphones, and apps holds promise as a clinical decision-making tool in health and mental health in primary care medicine. The aim of this study was to determine provider perspectives about the utility of these data for building digitally based decision-making tools.

Methods: This mixed quantitative and qualitative cross-sectional survey of a convenience sample of primary-care clinicians at Kaiser Permanente Northwest was conducted between April and July 2019 online via Institute for Translational Health Sciences' Research Electronic Data Capture. Study outcomes were 1) attitudes toward digital data, 2) willingness to use digital data to support clinical decision making, and 3) concerns and recommendations about implementing a digital tool for clinical decision making.

Results: This sample of 131 clinicians was largely white (n = 98) female (n = 91) physicians (n = 86). Although respondents (75.7%, n = 87) had a positive attitude toward using digital tools in their practice, 88 respondents (67.3%) voiced concerns about the possible lack of clinical utility, suspected difficulty in integration with clinical workflows, and worried about the potential burden placed on patients. Participants indicated that the accuracy of the data in detecting the need for treatment adjustments would need to be high and the tool should be clinically tested.

as the Patient Health Questionnaire 9,[3] which is based on retrospective self-report of symptoms and is collected only sporadically. Indeed, there is a marked decrease in the number of follow-up depression assessments in primary care medicine in people who screen positive for depression and receive treatment for it.[4] Self-reports also are not informative about when treatment should be augmented or switched, or if a patient needs to be seen immediately for emergency reasons. Although patients find these measures somewhat informative, they also find that the questions asked do not assess important measures of improvement, such as activity, social connectedness, and work productivity.[5] Decision support and access to expert opinion on the delivery of depression care is limited and impacts the quality of care substantially.[6] This problem is recognized by many large health care systems that want to support the use of decision support tools.[7,8]

To mitigate this problem, recent efforts have turned out the use of Clinical Decision Support Systems (CDSS)—data analytic tools embedded in electronic health records that compile patient information, and synthesize and visualize the information to support clinicians in making

ge

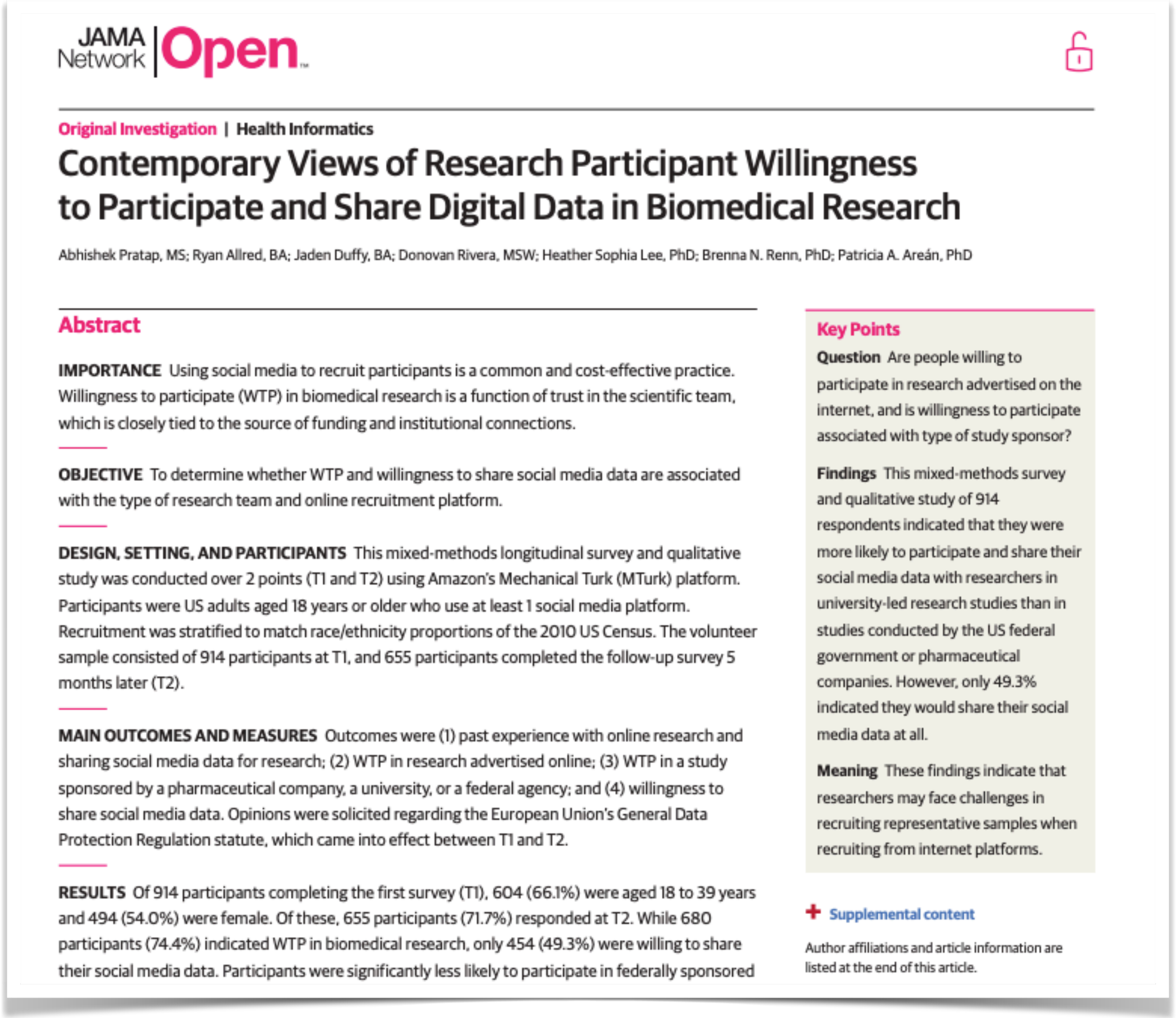
levels

background noise levels

Understanding real-world behavior using mediums that are acceptable to the target population remains critical and often (missed)

https://www.consultancy.uk/news/13430/britains-digital-families-six-archetypes-of-internet-and-tech-users

Assess Individuals’ willingness to participate and share data in online biomedical research


JAMA Network | Open

Original Investigation | Health Informatics

Contemporary Views of Research Participant Willingness to Participate and Share Digital Data in Biomedical Research

Abhishek Pratap, MS; Ryan Allred, BA; Jaden Duffy, BA; Donovan Rivera, MSW; Heather Sophia Lee, PhD; Brenna N. Renn, PhD; Patricia A. Areán, PhD

Abstract

IMPORTANCE Using social media to recruit participants is a common and cost-effective practice. Willingness to participate (WTP) in biomedical research is a function of trust in the scientific team, which is closely tied to the source of funding and institutional connections.

OBJECTIVE To determine whether WTP and willingness to share social media data are associated with the type of research team and online recruitment platform.

DESIGN, SETTING, AND PARTICIPANTS This mixed-methods longitudinal survey and qualitative study was conducted over 2 points (T1 and T2) using Amazon's Mechanical Turk (MTurk) platform. Participants were US adults aged 18 years or older who use at least 1 social media platform. Recruitment was stratified to match race/ethnicity proportions of the 2010 US Census. The volunteer sample consisted of 914 participants at T1, and 655 participants completed the follow-up survey 5 months later (T2).

MAIN OUTCOMES AND MEASURES Outcomes were (1) past experience with online research and sharing social media data for research; (2) WTP in research advertised online; (3) WTP in a study sponsored by a pharmaceutical company, a university, or a federal agency; and (4) willingness to share social media data. Opinions were solicited regarding the European Union's General Data Protection Regulation statute, which came into effect between T1 and T2.

RESULTS Of 914 participants completing the first survey (T1), 604 (66.1%) were aged 18 to 39 years and 494 (54.0%) were female. Of these, 655 participants (71.7%) responded at T2. While 680 participants (74.4%) indicated WTP in biomedical research, only 454 (49.3%) were willing to share their social media data. Participants were significantly less likely to participate in federally sponsored

Key Points

Question Are people willing to participate in research advertised on the internet, and is willingness to participate associated with type of study sponsor?

Findings This mixed-methods survey and qualitative study of 914 respondents indicated that they were more likely to participate and share their social media data with researchers in university-led research studies than in studies conducted by the US federal government or pharmaceutical companies. However, only 49.3% indicated they would share their social media data at all.

Meaning These findings indicate that researchers may face challenges in recruiting representative samples when recruiting from internet platforms.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

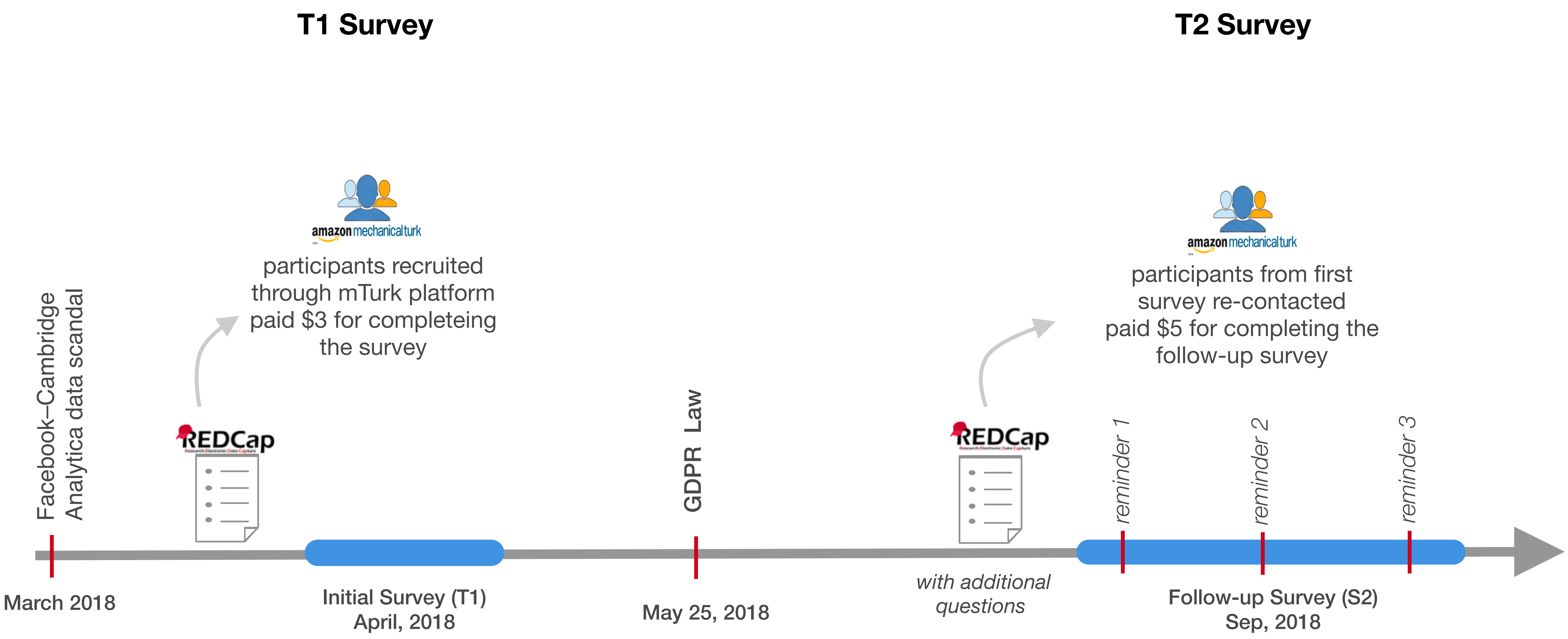
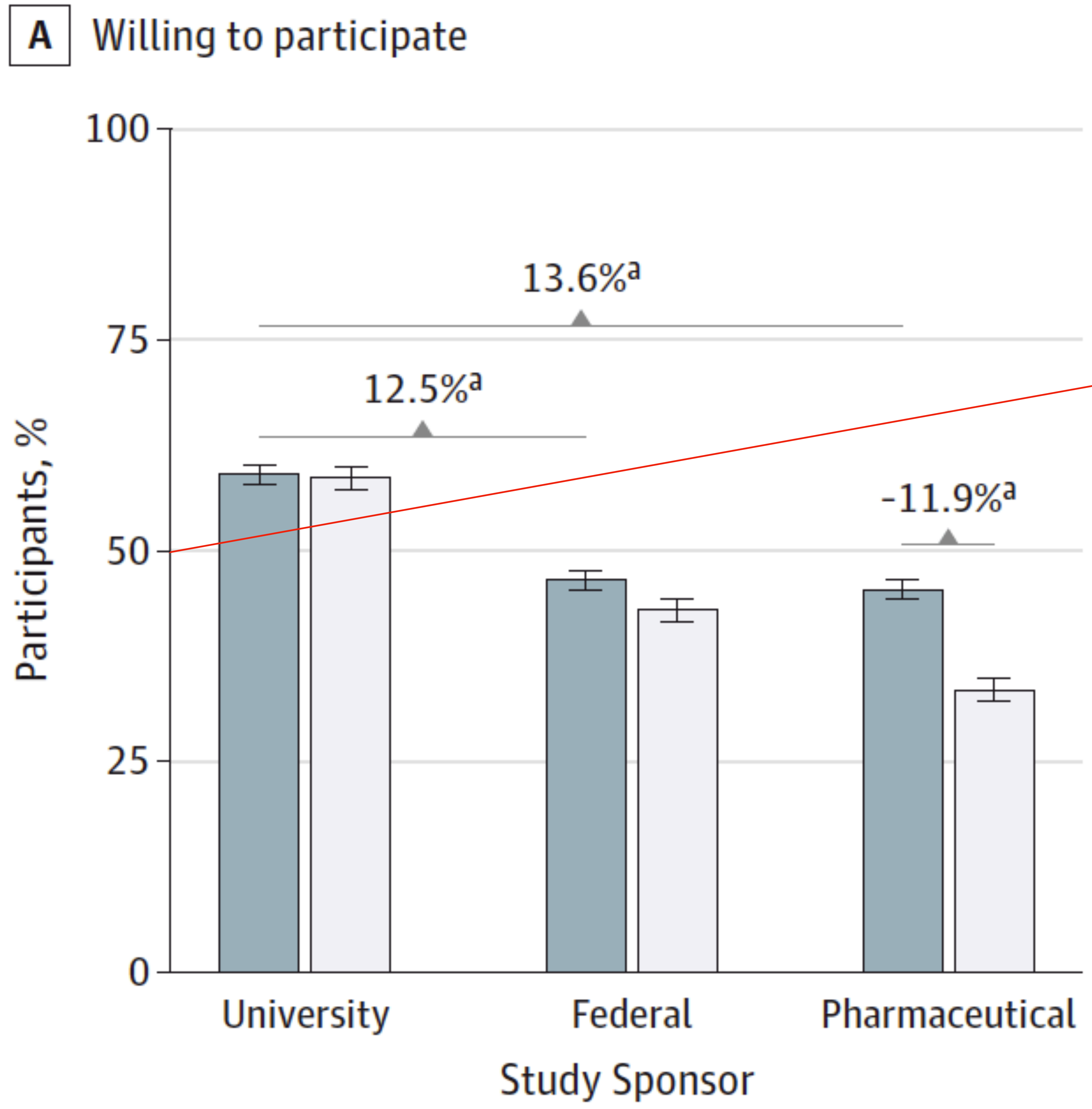

T1 Survey
T2 Survey
amazon mechanical turk
participants recruited through mTurk platform paid $3 for completeing the survey
amazon mechanical turk
participants from first survey re-contacted paid $5 for completing the follow-up survey
Facebook–Cambridge Analytica data scandal
REDCap
GDPR Law
REDCap
reminder 1
reminder 2
reminder 3
March 2018
Initial Survey (T1)
April, 2018
May 25, 2018
with additional questions
Follow-up Survey (S2)
Sep, 2018

Figure 2. Proportion of Participants Willing to Participate and Share Their Social Media Data



[a] Statistically significant at false discovery rate–corrected $P < .001$.

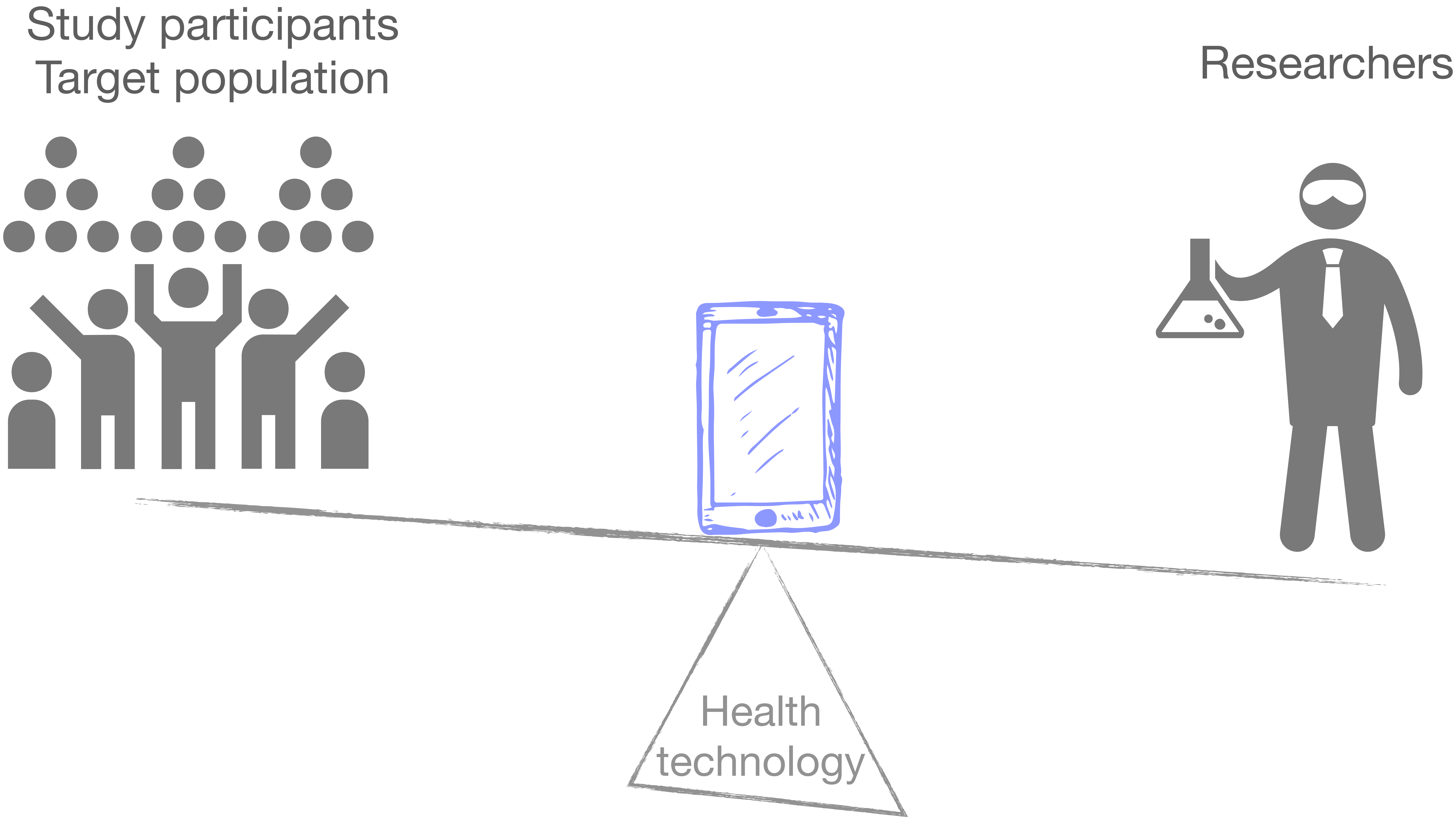
[b] Statistically significant at false discovery rate–corrected $P < .05$.

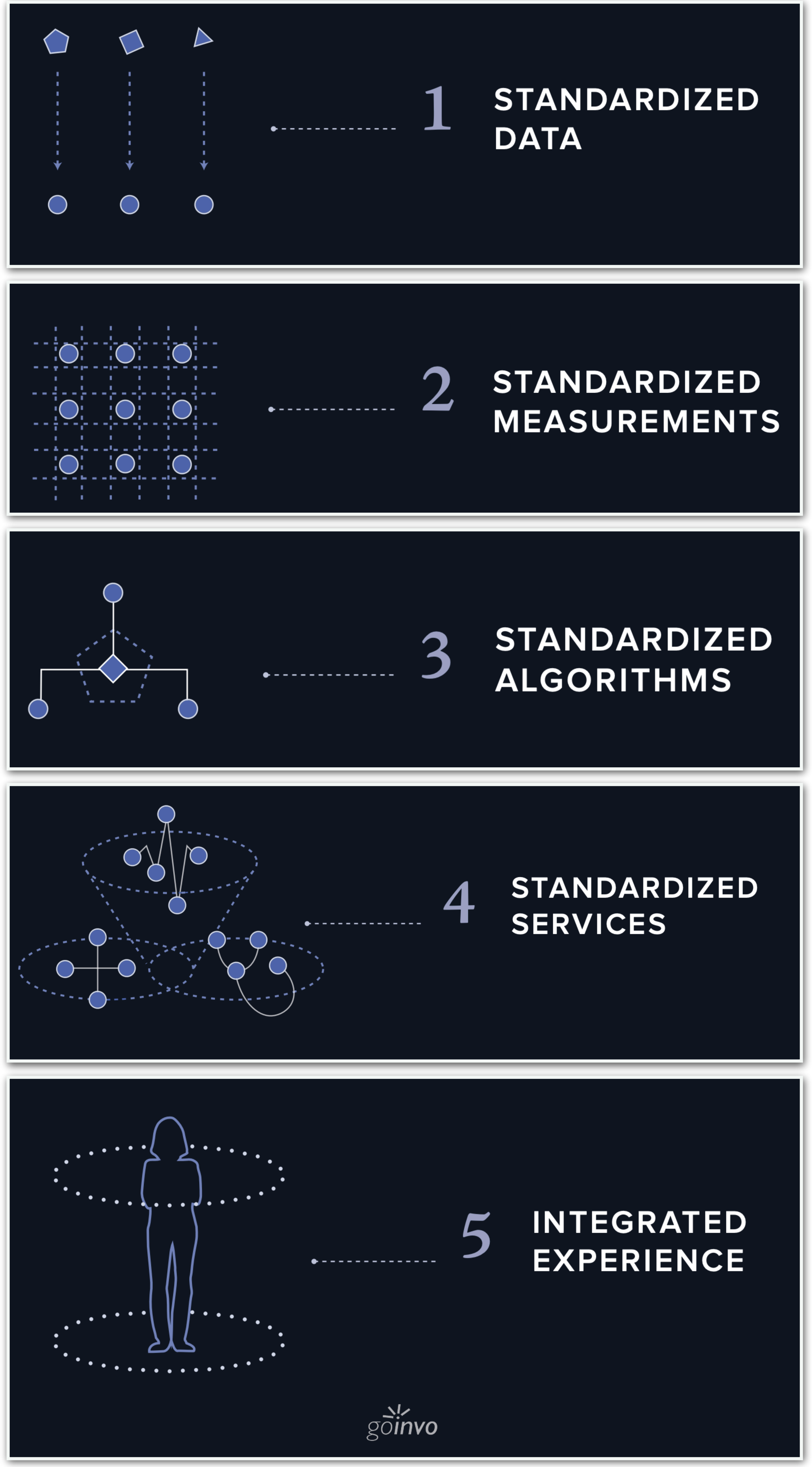
Why behind not willing to Participate/Share

"I do not trust nor do I respect ……… companies. I believe there only interest is profit for themselves and not in the best interests of the public in general."

"I just don't share any social media data especially after the Cambridge Analytica fiasco"

"I think the ads are just aimed at fixing a public relations problem. They still make their money from collecting our data and selling it and they aren't going to stop."


Study participants
Target population
Researchers
Health
technology



Future proofing data collection in health research

critical to develop transparent systems even if tech development is out-sourced

John Torous, Gerhard Andersson, Andrew Bertagnoli, Helen Christensen, Pim Cuijpers, Joseph Firth, Adam Haim, Honor Hsin, Chris Hollis, Shôn Lewis, David C Mohr, **A. Pratap**, Spencer Roux, Joel Sherrill, Patricia A. Arean Towards a consensus around standar
digital mental health **World Psychiatry. 2019**

John Torous, Joseph Firth, Kit Huckvale, Mark E Larsen, Theodore D Cosco, Rebekah Carney, Steven Chan, **A. Pratap**, Peter Yellowlees, Til Wykes, Matcheri Keshavan & Helen Christensen The Emerging Imperative for a Consensus Approach Toward the Rating a
of Mental Health Apps. **The Journal of nervous and mental disease, 2018**

https://yes.goinvo.com/articles/a-path-towards-standardized-health

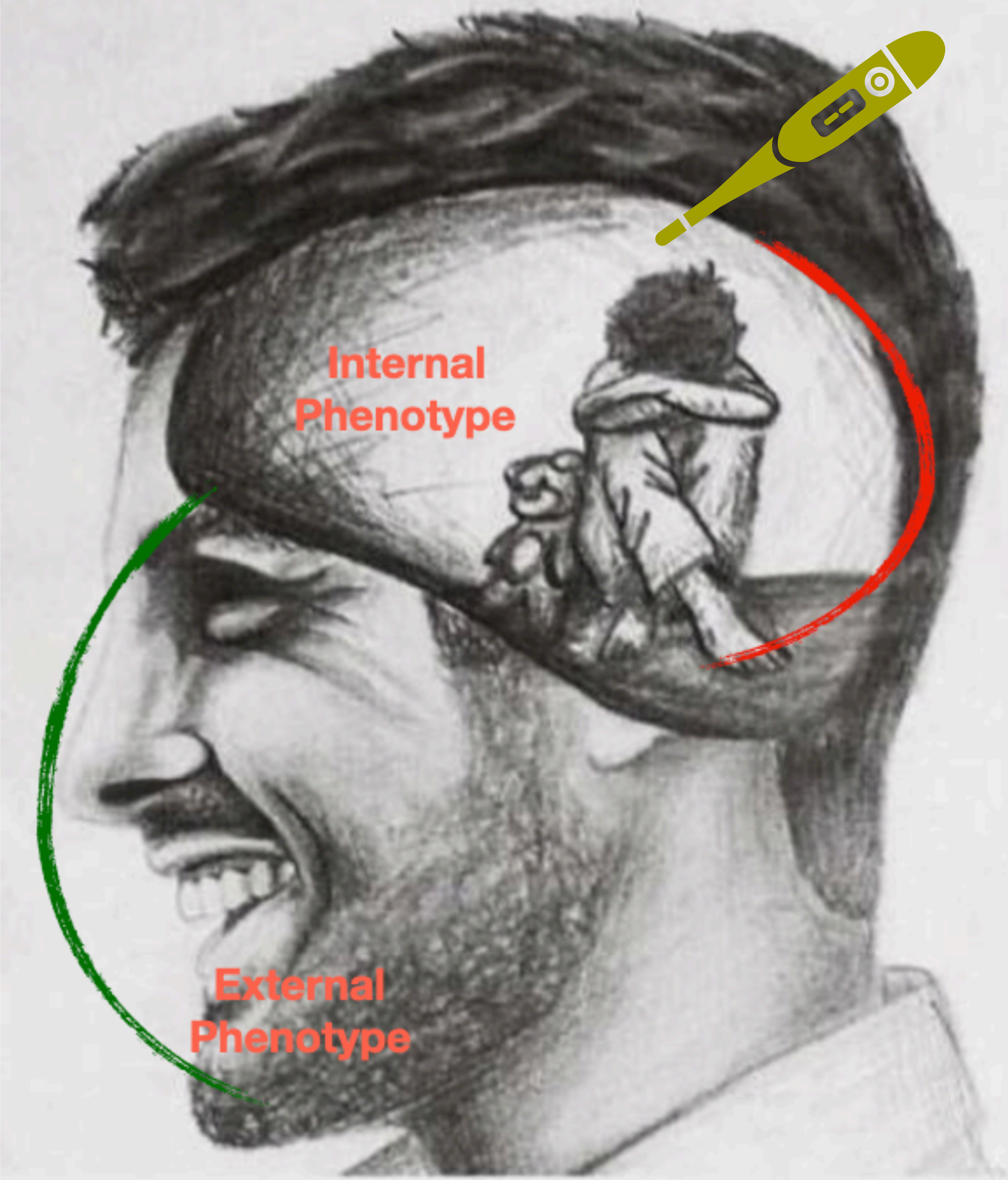
Digitally augmented clinical research & care

Subjective → **Objective & holistic**

Episodic → **Continuous**

Reactive → **Proactive**

Cohort → **Personalized**



https://www.facebook.com/photo.php?fbid=2777513792274642&set=a.290652084294171&type=3&theater

Thank You

New Digital Health & AI Research Group @ KCNI

Hiring soon

- Postdocs
- Research Analysts
- Grad students
- Designers
- Data Scientists/Engineers

abhishek.pratap@camh.ca

Skills

- Digital health apps
- Data viz & management
- Scripting - Python, R
- Statistical data analysis
- Machine learning, NLP
- User centered design
- Health communication

Who, **When** and for **How long** in the real world ?

Next

camh
Krembil Centre for Neuroinformatics

Day 7

Digital Health
&
Population-based data resources

9:00 am - 10:30 am	Digital Health for Mental Health - Opportunities & Challenges Dr. Abhi Pratap
10:45 am- 12:15 pm	Population-based resources and the BrainHealth Databank Drs. Daniel Felsky, Joanna Yu & Abhi Pratap
1:00 pm - 2:30 pm	Workshop/Demo: Reproducible analysis using Synapse as part of an integrated workflow Dr. Abhi Pratap
2:45 pm - 4:15 pm	Workshop: Introduction to interactive methods Dr. Daniel Felsky

Housekeeping

Krembil Centre for Neuroinformatics

1:00 pm - 2:30 pm	Workshop/Demo: Reproducible analysis using Synapse as part of an integrated workflow Dr. Abhi Pratap