

Challenges in deriving health insights from Real World Sensor Data

SageBionetworks and mobile health research

Non-profit research organization in Seattle; “*Better Science Together*”

Involved with mHealth/Digital Health research since the inception of the ResearchKit (iOS) and ResearchStack (Android) frameworks

> 215,000 consented study participants
40+ studies



Parkinsons



Cardiovascular



MoleMapper



Asthma



Multiple
Sclerosis



AllofUS
CRF



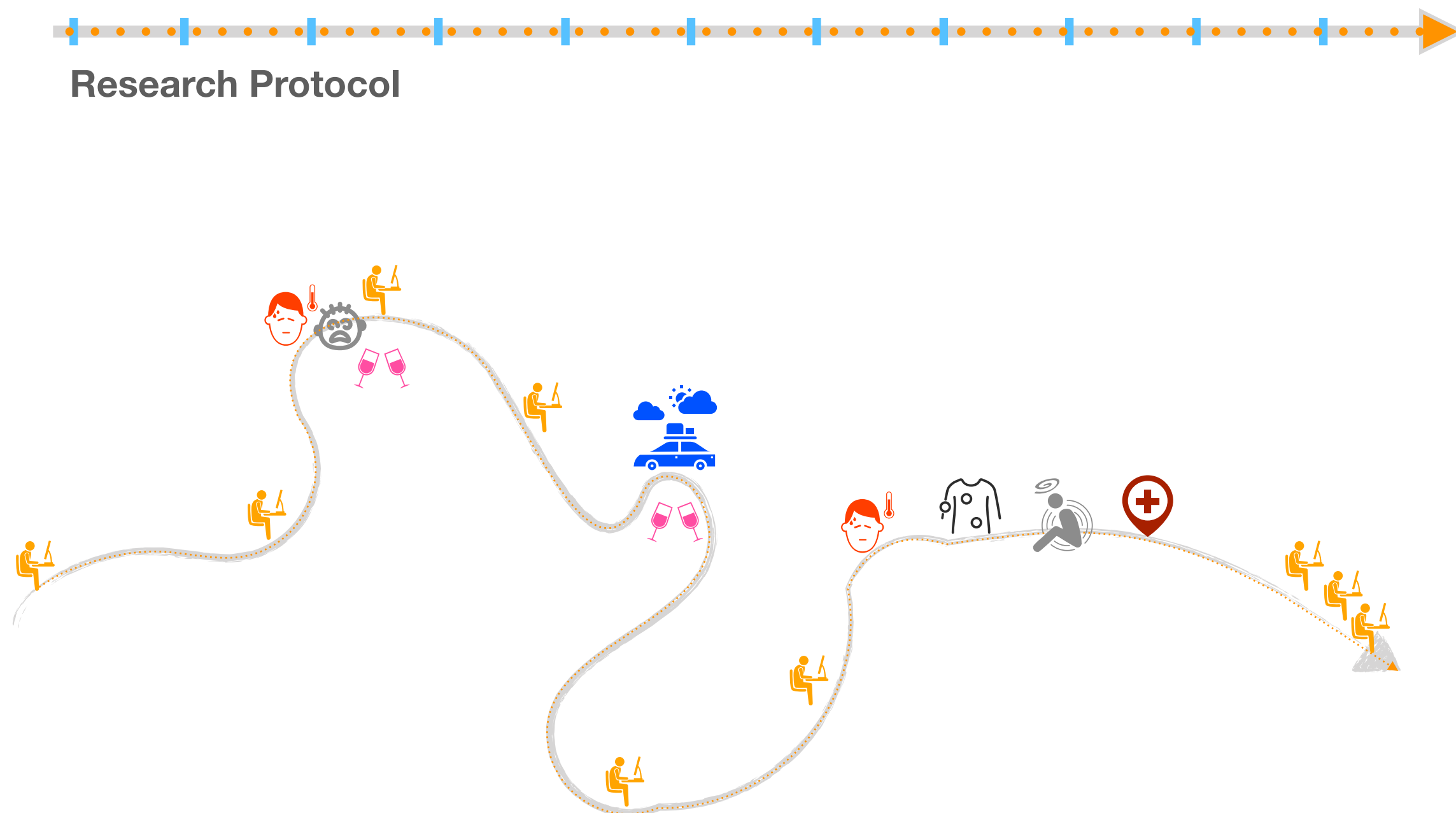
Samsung
my BP lab

<http://synapse.org/digitalhealth>



Studies in the real world

- Continual Passive Monitoring
- Frequent Tasks

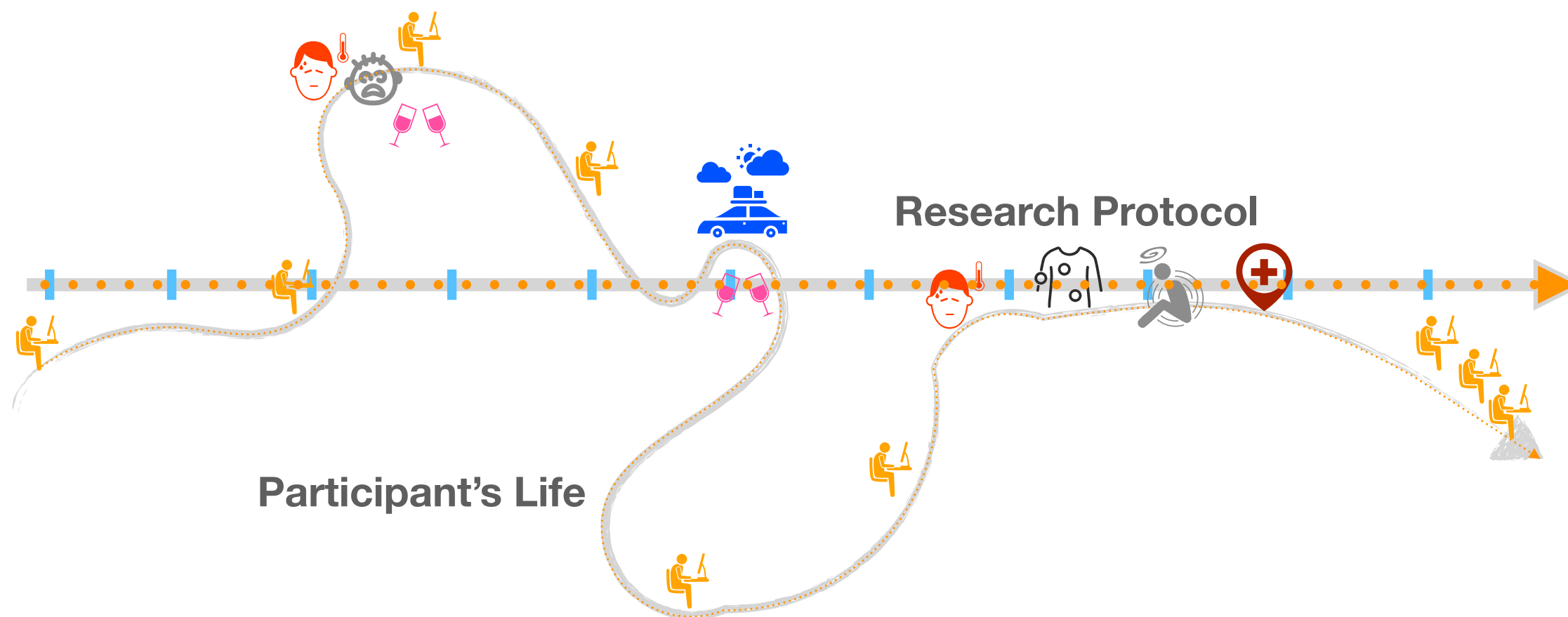


Participant's Life



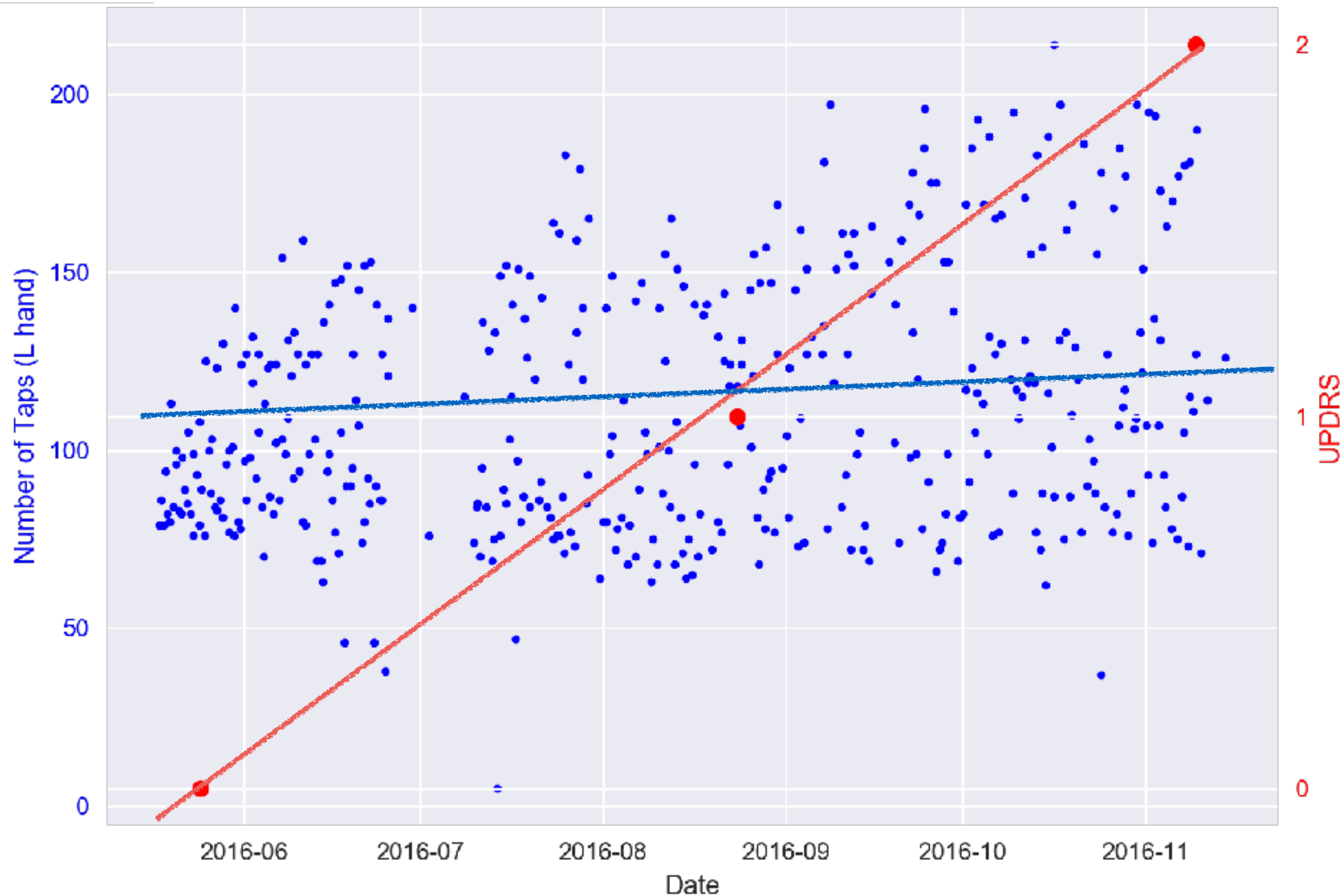
Continual Passive Monitoring

Frequent Tasks



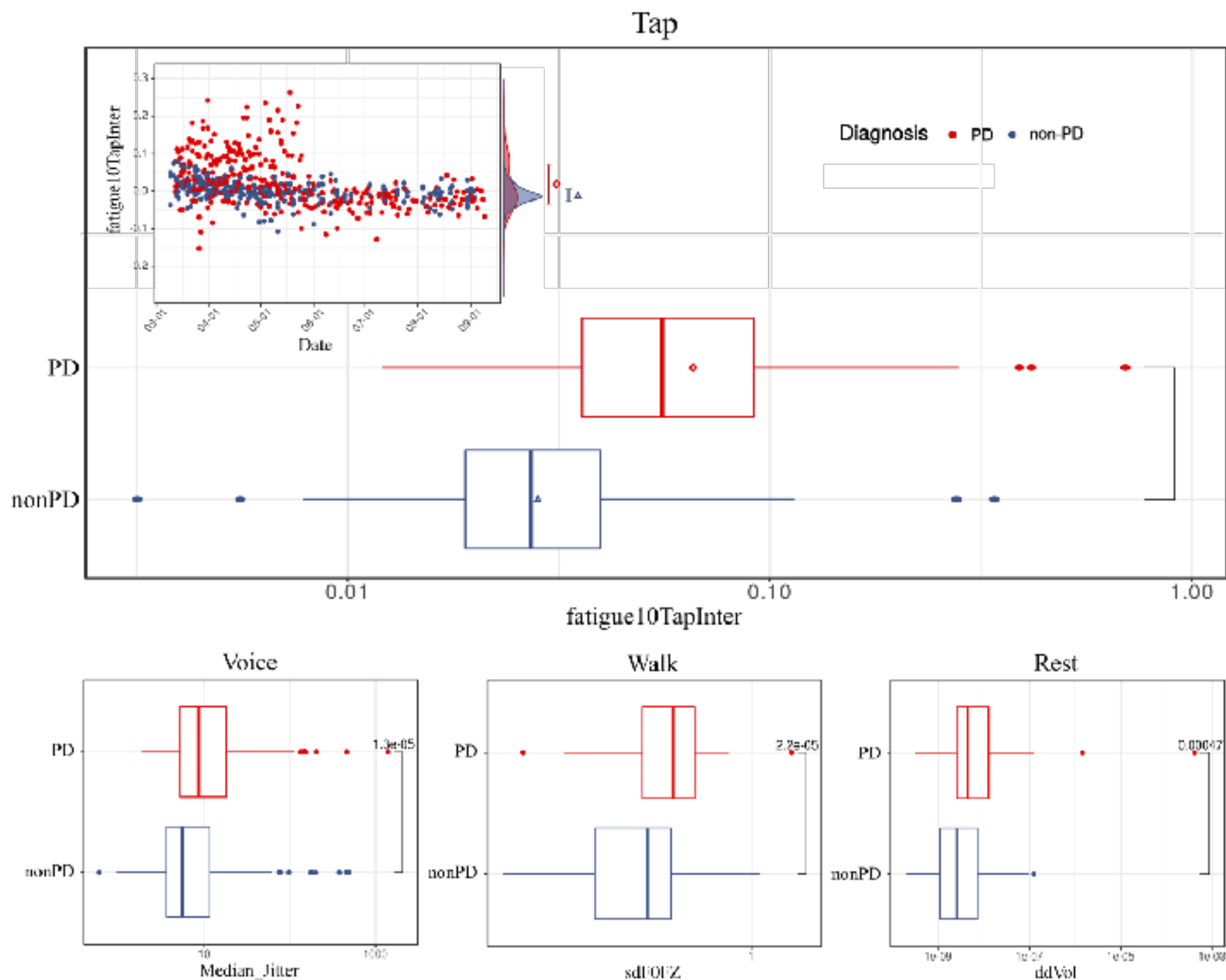


In the wild measurements: **what is happening in life?**

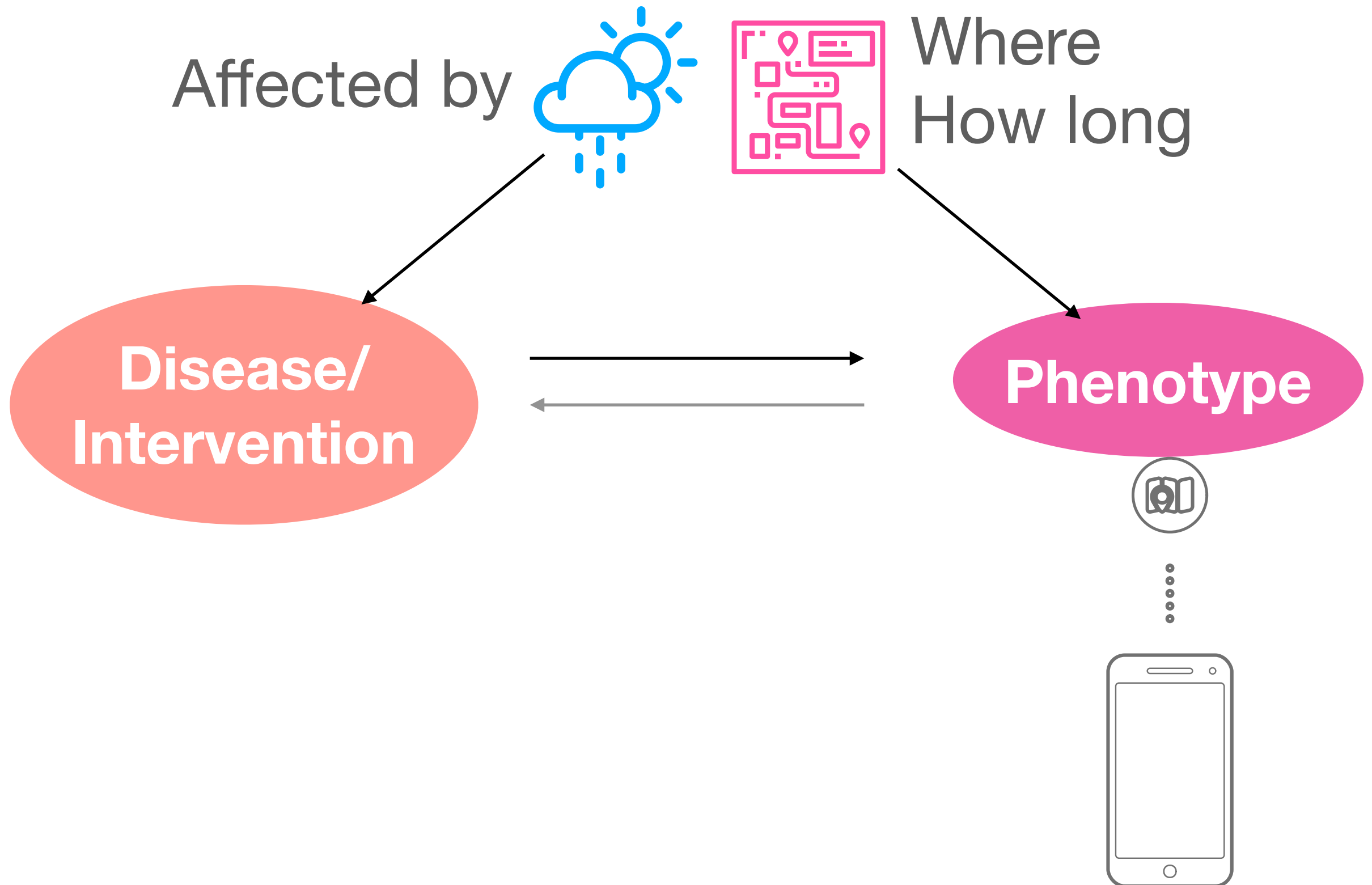




Ah but there is signal in the noise: Variation is a sign of disease in PD



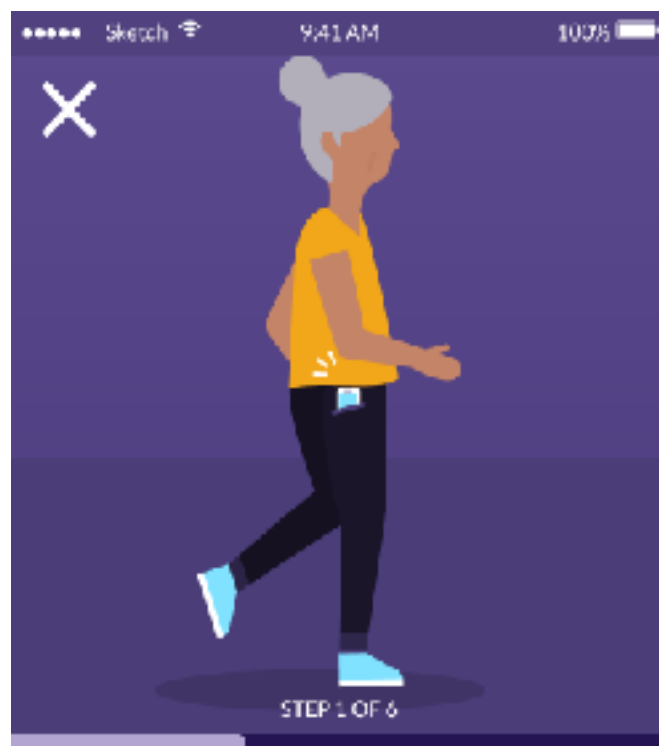
Context Matters: The problem of confounders



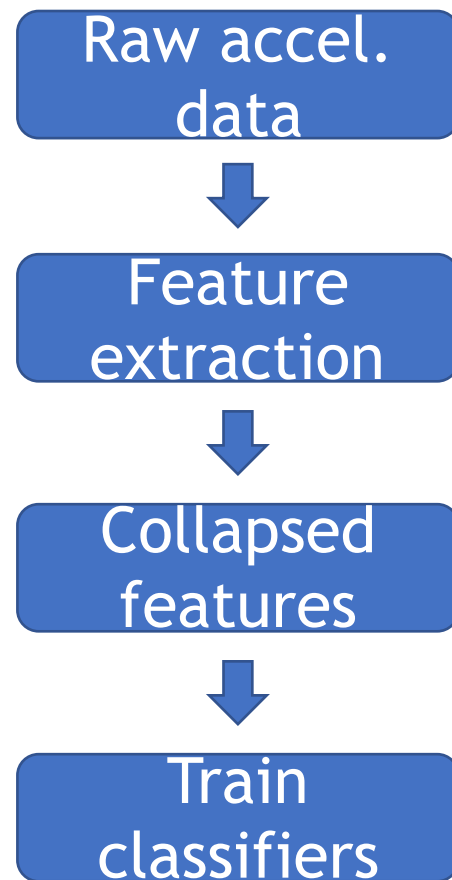
Context Matters: The problem of confounders

The ever so popular observational study

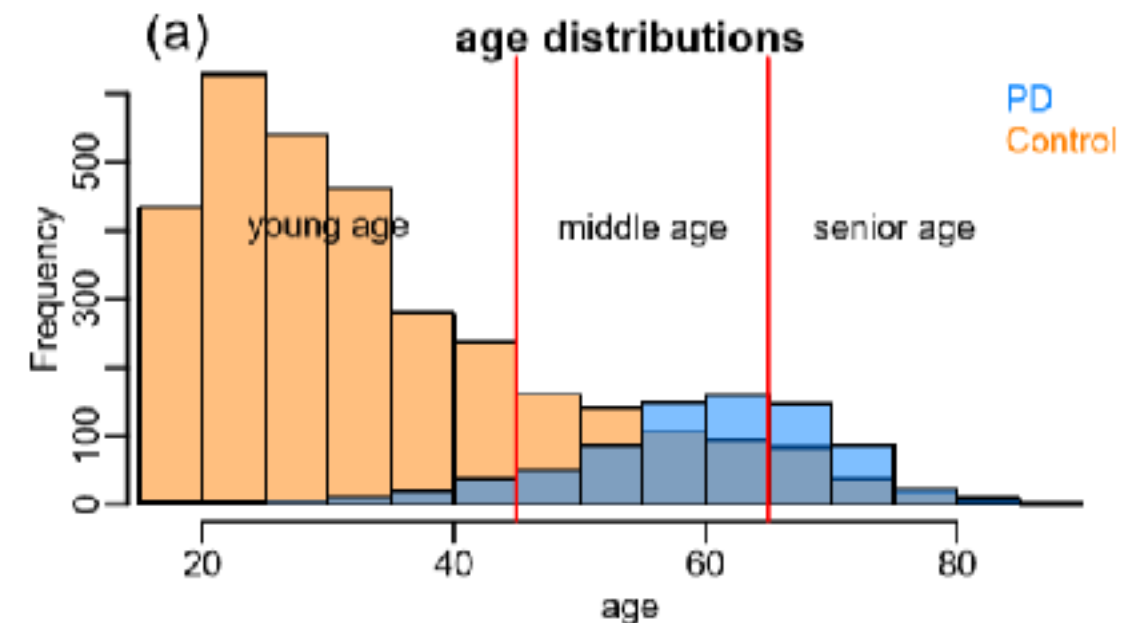
Goal: build a classifier of disease status



First, walk for 30 seconds



mPower: classifier of Parkinson's Disease patients vs healthy Controls



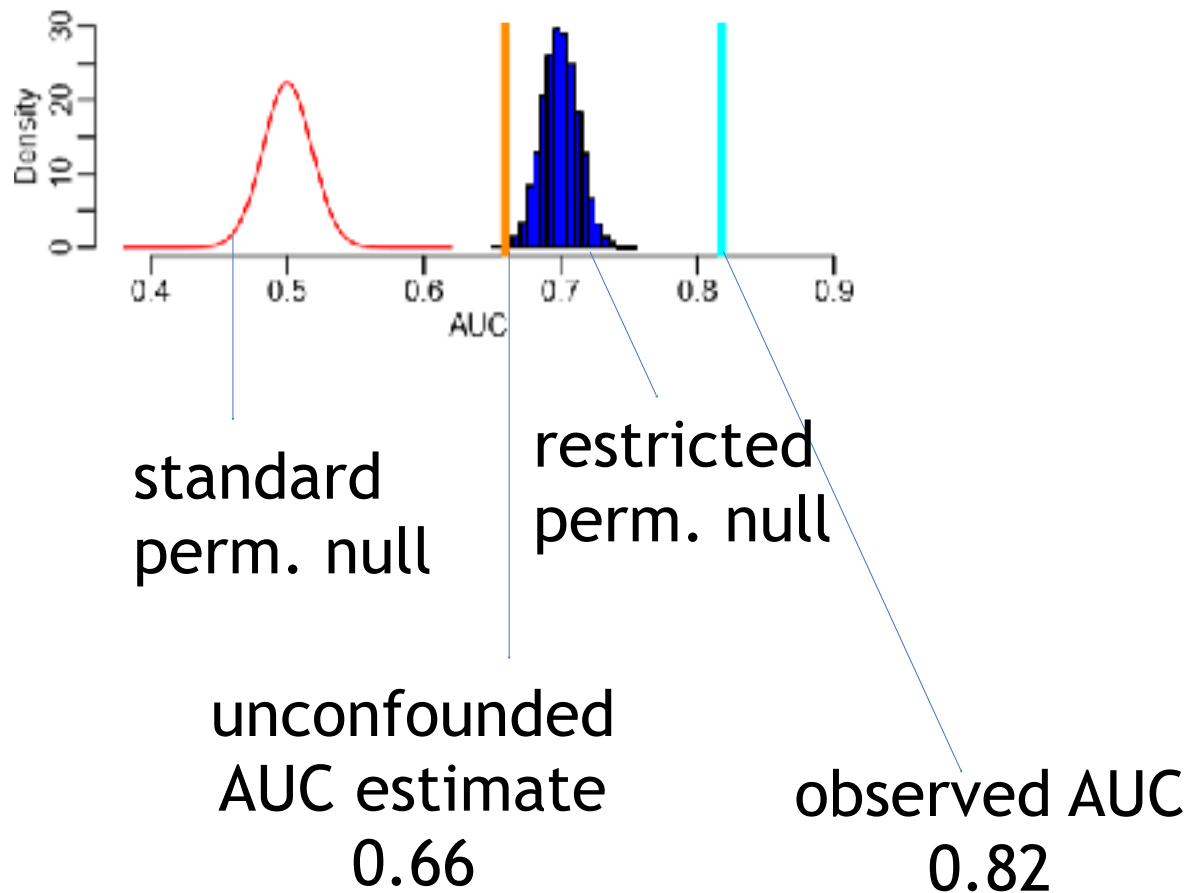
Question:

Is my classifier learning age related signals on top of the disease signal?



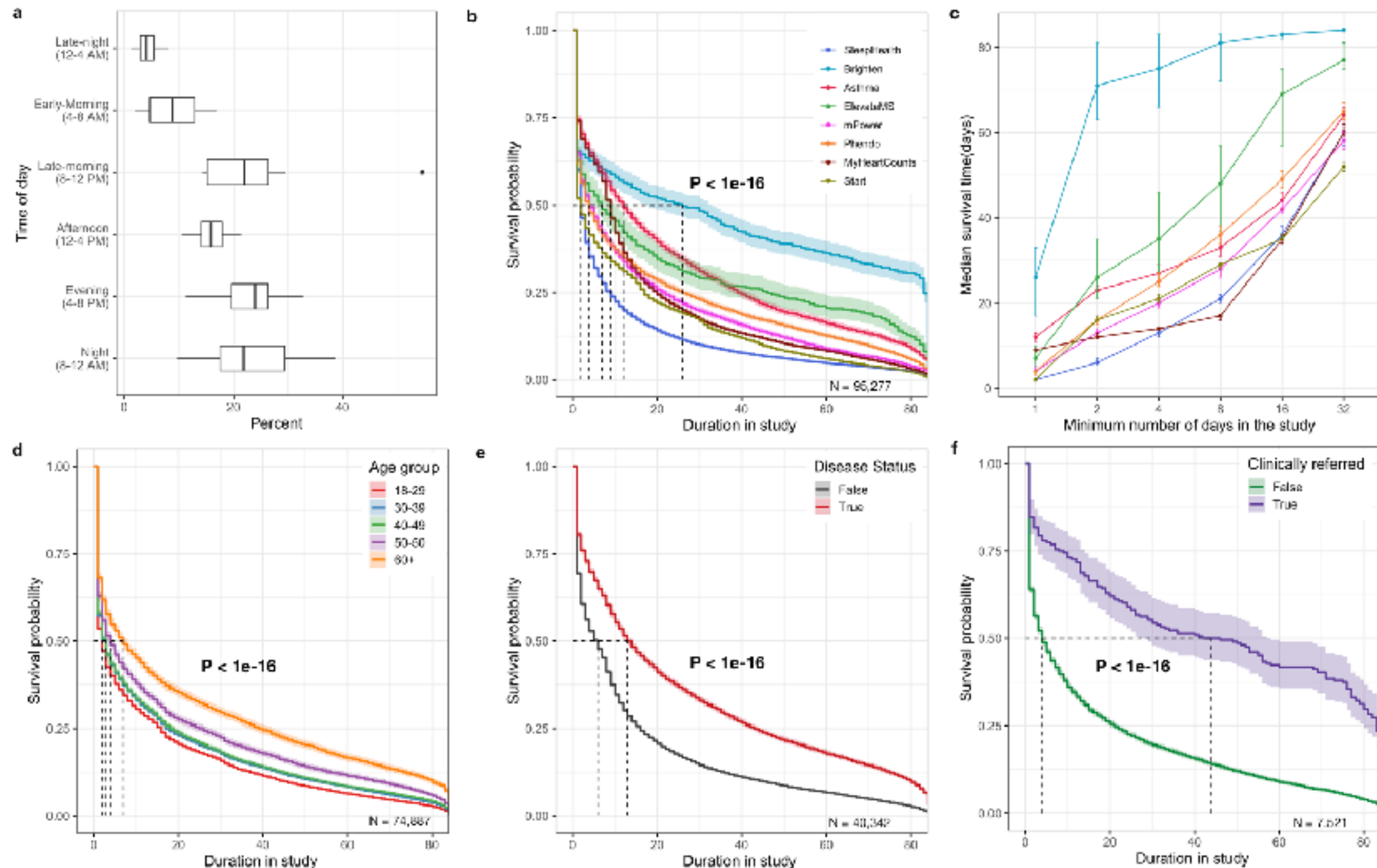
If you are not careful you will overestimate your signal

no adjustment



Chaibub Neto et al (2018) Using permutations to assess confounding in machine learning applications for digital health. ML4H Workshop at NIPS 2018.

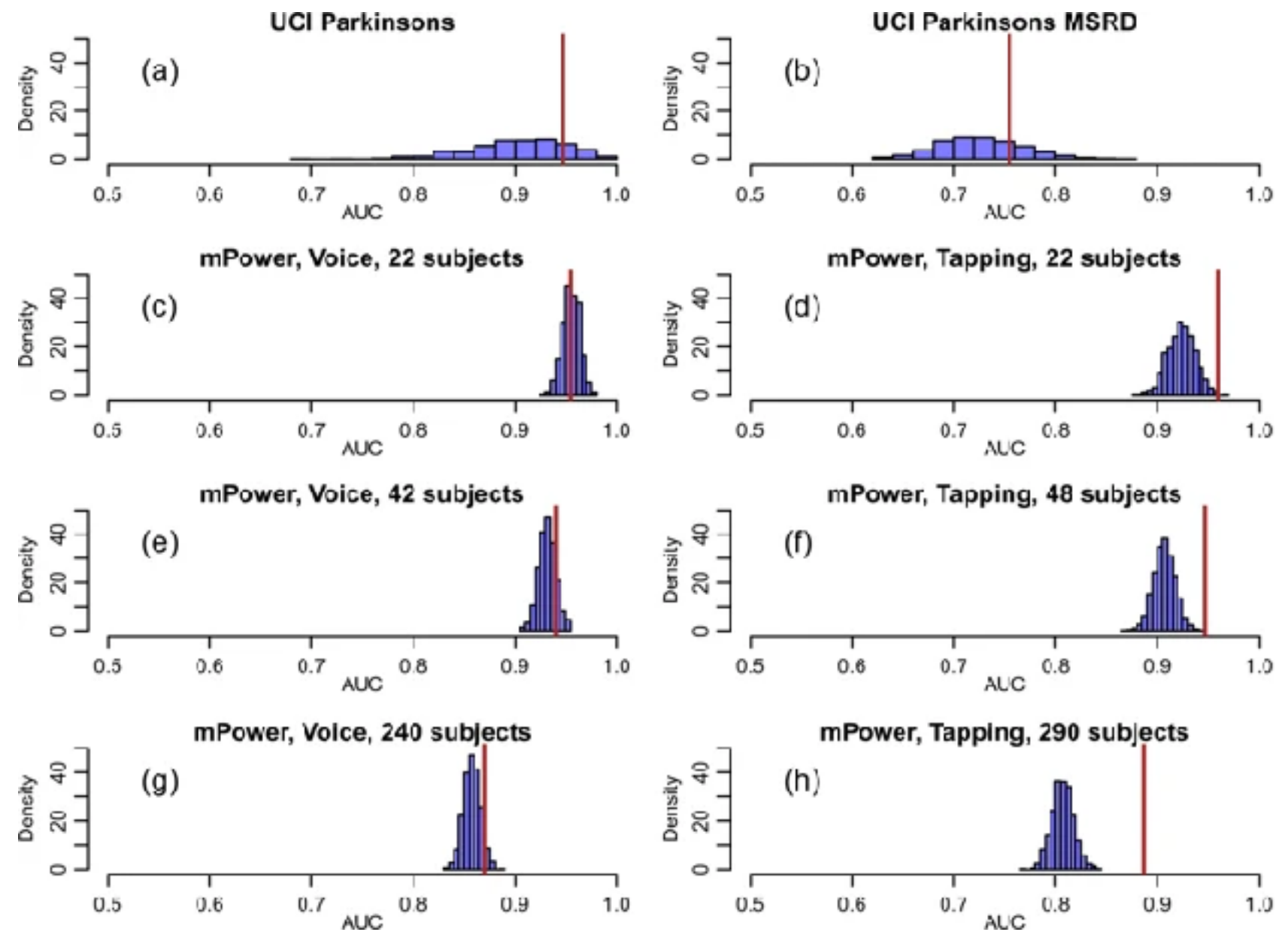
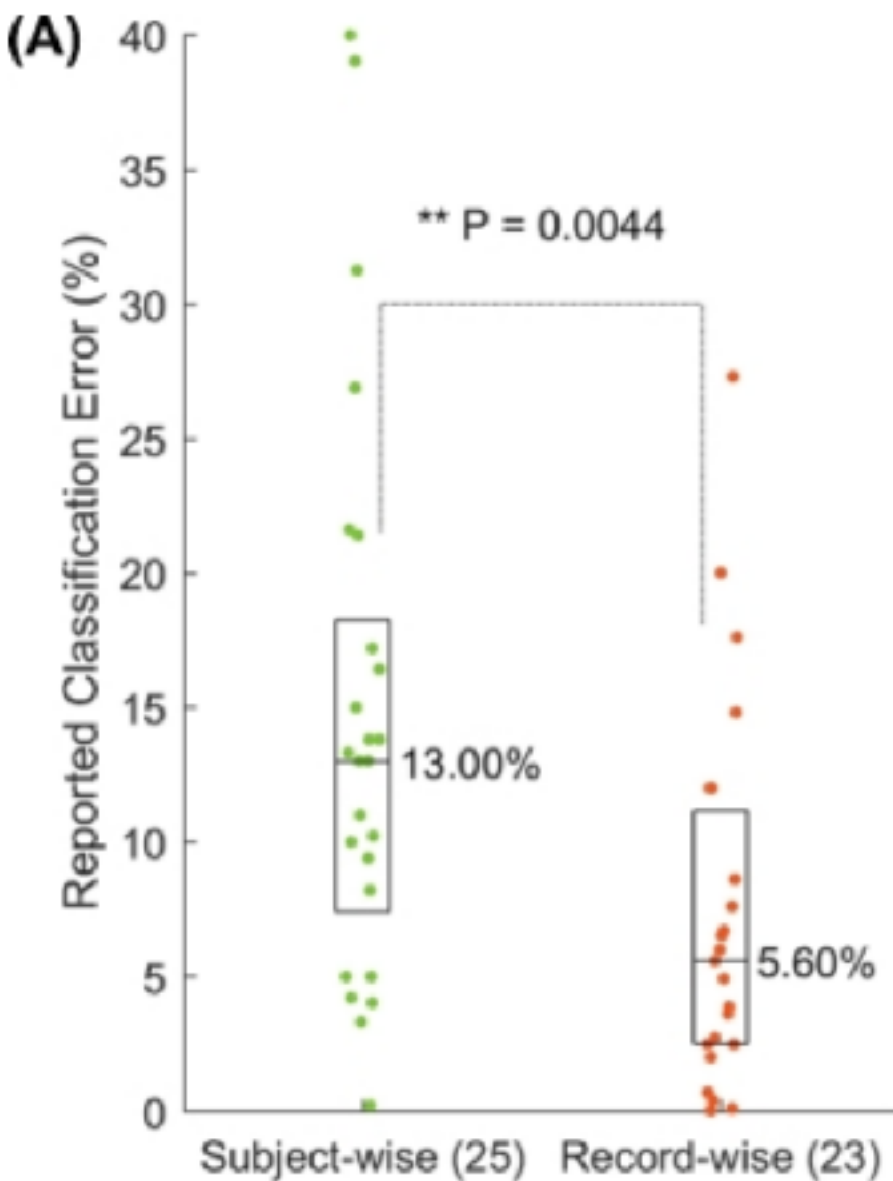
A meta study across 8 studies and 109,000 participants reveals differences in retention



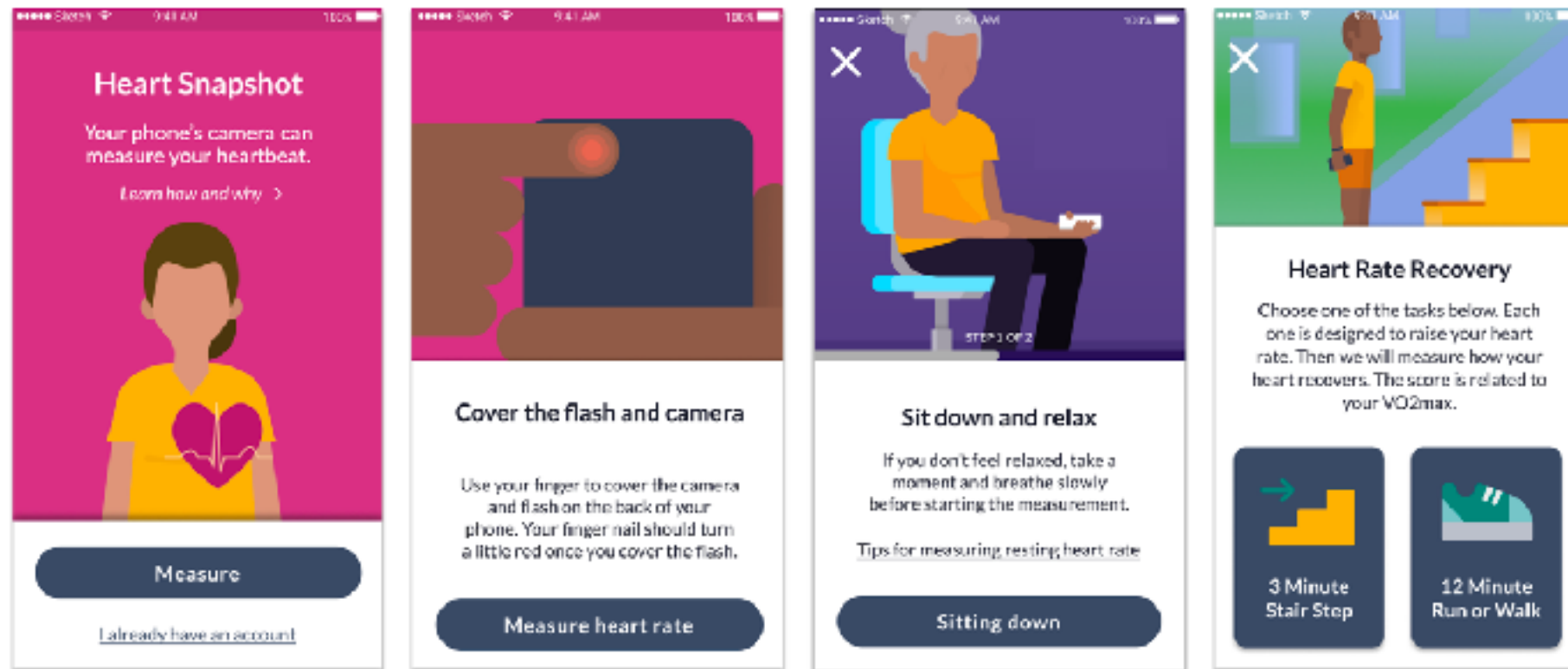
Pratap, A. *et al.* (2020) Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *npj Digit. Med.* 3, 21.



Success are overestimated: the identity confounder

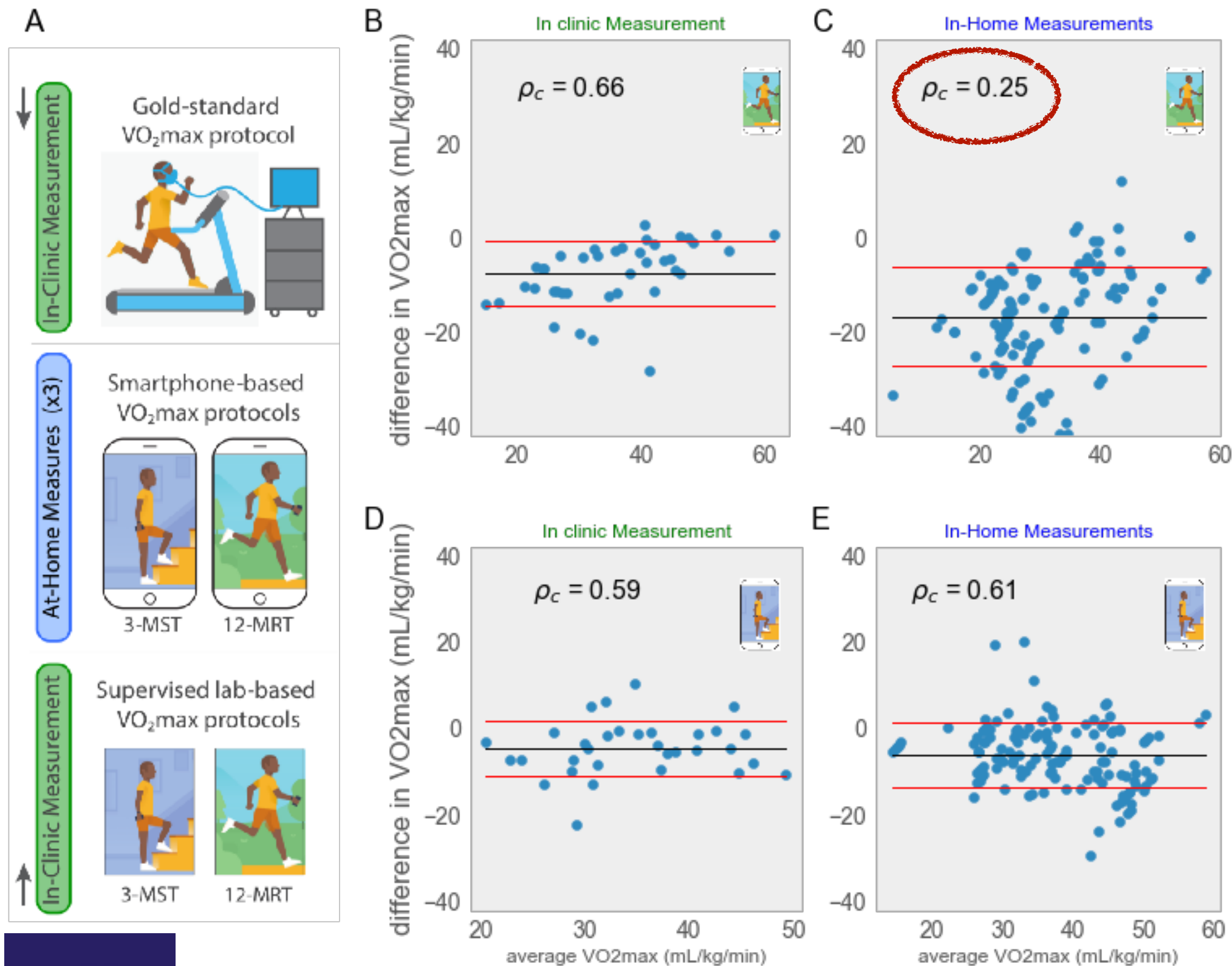


An example validation study for Cardiorespiratory fitness



Context of validation matters

Multiple avenues of validation - at home vs clinic vs lab



Digital Health Research Projects

Synapse is used as a repository for mobile Health resources including data being collected using smartphones, and wearables analytical tools for processing and analyzing digital health data and other resources.

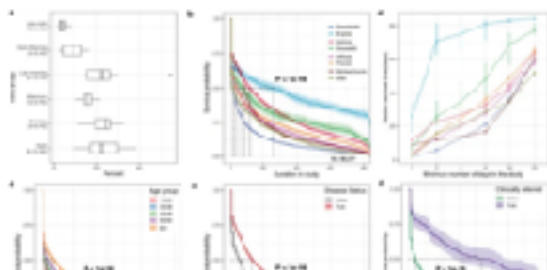
Available Projects

17

Last updated on
02/25/2020 7:31 AM

Digital Health Study Retention Data

Repository	App Code	Participants	Records	Analytical Tools
Portal	not applicable	109,914	3.5 million	github



This data is an aggregation of 8 different digital health studies used to evaluate factors that drive long term engagement and retention. This is a compendium to the paper:

Abhishek Pratap, Elias Chaibub Neto, Phil Snyder, Carl Stepnowsky, Noémie Elhadad, Daniel Grant, Matthew H. Mohebbi, Sean Mooney, Christine Suver, John Wilbanks, Lara Mangravite, Patrick J. Heagerty, Pat Arean, Larsson Omberg "Indicators of



mPower - Researcher Community

mPower Public Researcher Portal ☆

Search



Larsson Omberg (larsson)



Help



Synapse ID: syn4993293

DOI: doi:10.7303/syn4993293

Storage Location: Synapse Storage

Share

Annotations

Tools

Wiki

Files

Tables

Discussion

mPower Public Researcher Portal

- 1 - Accessing the mPower data
- 2 - Data description
- 3 - mPower research community
- 4 - Publications
- 5 - FAQs
- 6 - Contact us



mPower: Mobile Parkinson Disease Study

Data description

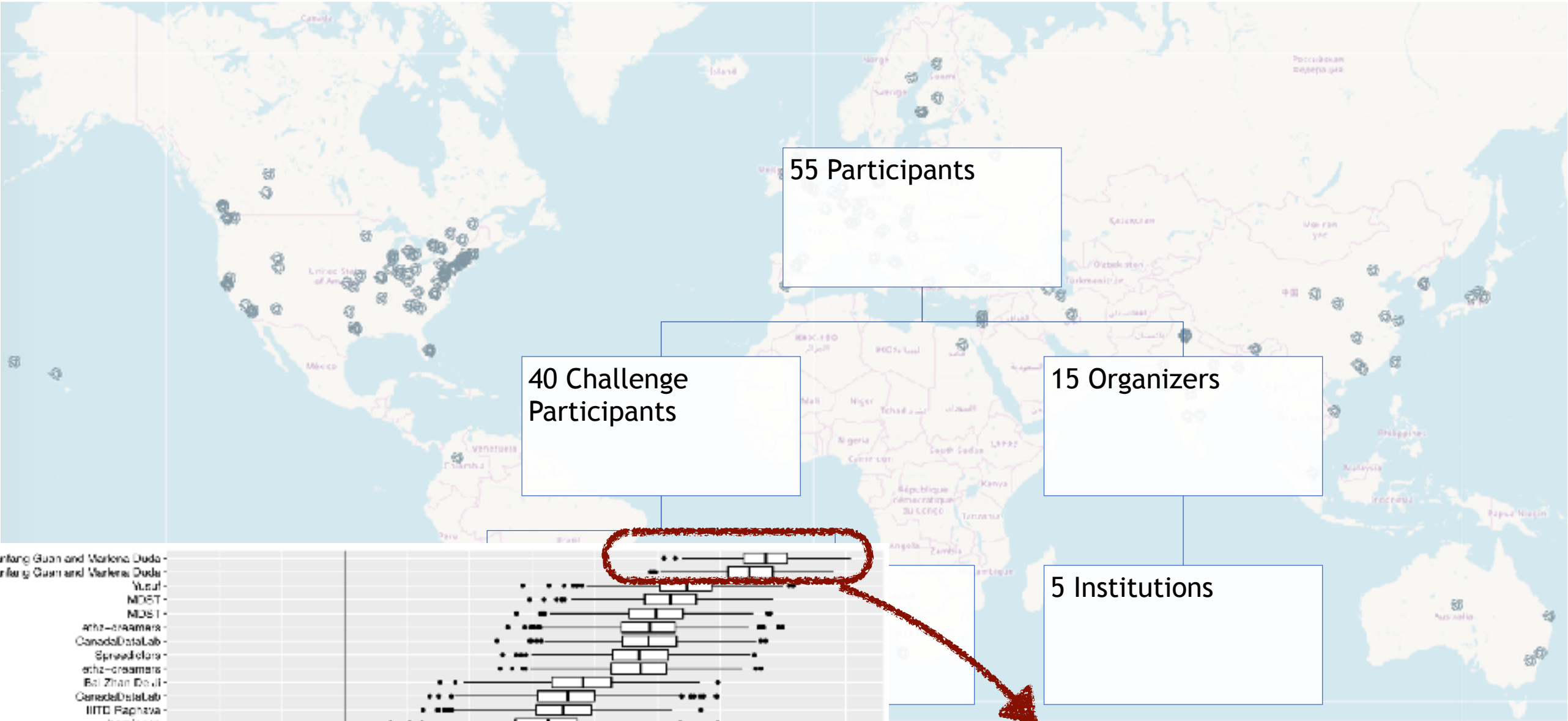
Accessing the data

Research community

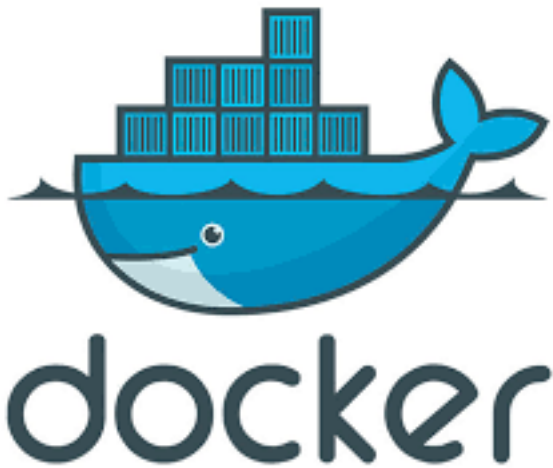
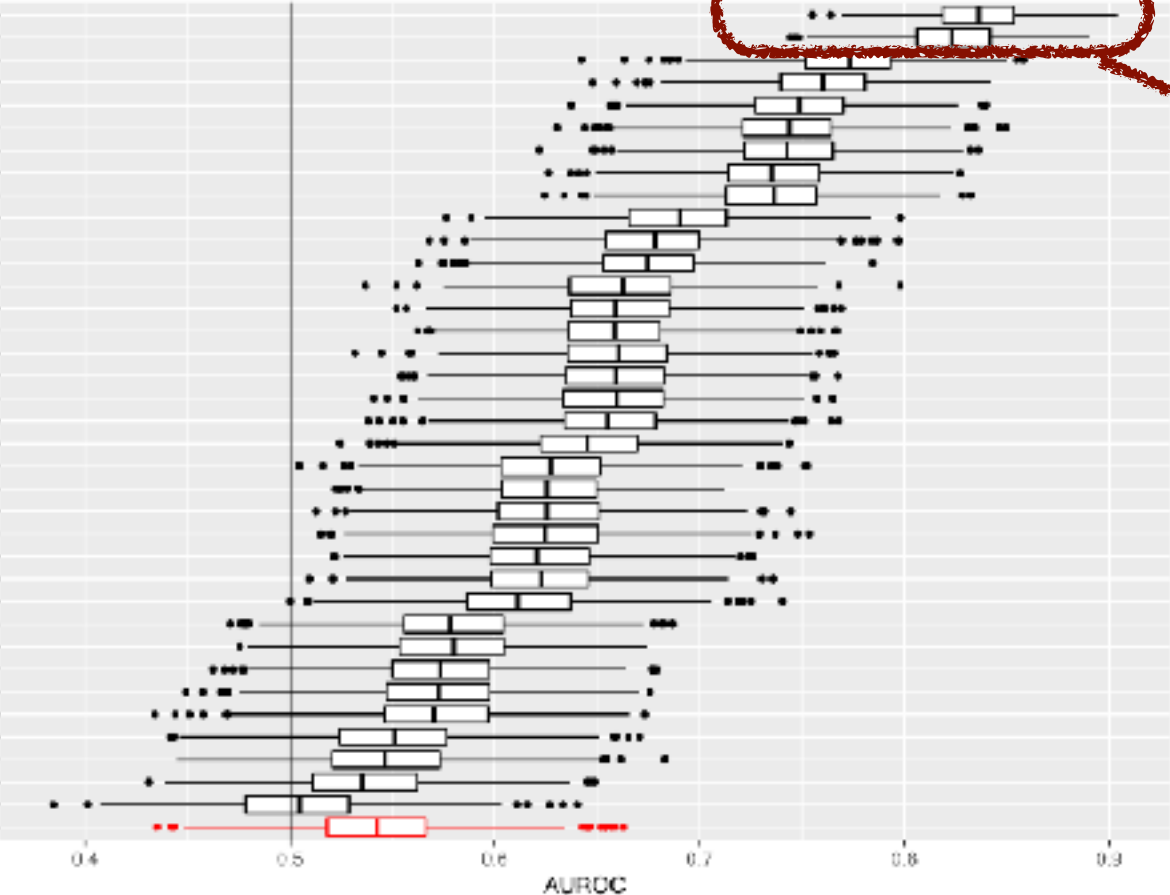
FAQs

Publications

Create impartial Benchmarking



Yuanfang Guan and Marketa Duda-
Yuanfang Guan and Marketa Duda-
Yusuf
MOST
MOST
arch-dreamers
CanadaDataLab
SpreadEars
arch-dreamers
Bei Zhan De Ji
CanadaDataLab
IITD Raghava
bennigan
Bei Zhan De Ji
maxwg
vmorozov
IITD VLBiomed
vmorozov
maxwg
bennigan
Yusuf
001
MonteShaller
MonteShaller
wangs@1990
SpreadEars
OniesEngineering
Abcensis III
team dem
DMIS_PO
001
Abcensis III
OniesEngineering
CCA Pegasus
CCA Pegasus
IITD VLBiomed
Clinical:Age+Gender



Develop Reusable Code and Pipelines

open collaborative methodologies



reusable



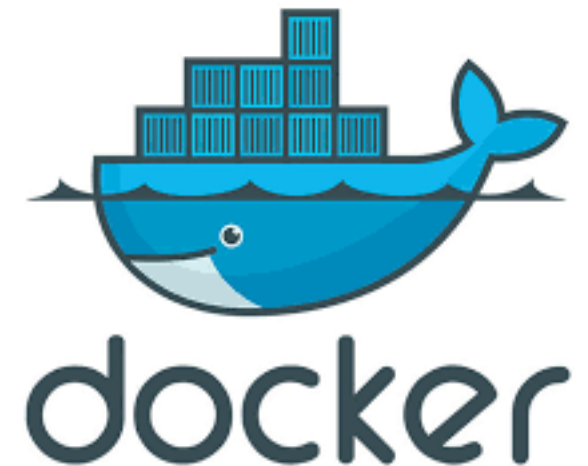
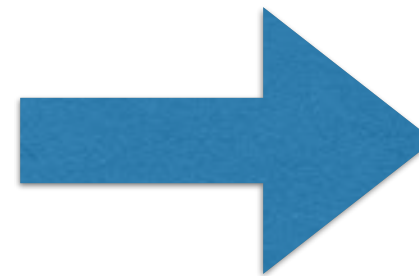
<https://github>



[syn12144677](#)

 Sage-Bionetworks / [mpowertools](#)

```
10 process_medicationChoiceAnswers <- function
11   tryCatch({
12     d <- jsonlite::fromJSON(json_file)
13     data.frame(medication = paste(unique
14       collapse = "+"), medicationTime
15       collapse = "+"))
16   }, error = function(err) {
17     data.frame(medication = "NA", medic
18   })
```



Cloud enabled workflow

<https://github.com/Sage-Bionetworks/mhealthtools>





The Open Wearables Initiative (OWEAR) is a collaboration designed to promote the effective use of high-quality, sensor-generated measures of health in clinical research through the open sharing of algorithms and data.

- Community hub for digital health and registry for
 - Algorithms
 - Data
- Benchmarking program to evaluate algorithms



SageBionetworks

<https://www.owear.org/>

Thank You



Brian M. Bot,
Elias Chaibub Neto,
Thanneer Malai Perumal,
Abhishek Prathap,
Solly Sieberts,
Phil Snyder,
Meghasyam
Tummalacheria
Lara Mangravite