Successes and Challenges of Developing Generalizable Biomarkers

Session 2: Precision Medicine Opportunities, Challenges, and Lessons Learned in Multimodal Biomarkers for Central Nervous System Disorders

Jukka-Pekka "JP" Onnela, D.Sc.
Associate Professor of Biostatistics
Department of Biostatistics
Harvard University

March 14, 2023

Successes and Challenges of Developing Generalizable Digital Behavioral Biomarkers

Session 2: Precision Medicine Opportunities, Challenges, and Lessons Learned in Multimodal Biomarkers for Central Nervous System Disorders

Jukka-Pekka "JP" Onnela, D.Sc.
Associate Professor of Biostatistics
Department of Biostatistics
Harvard University

March 14, 2023

DISCLOSURES

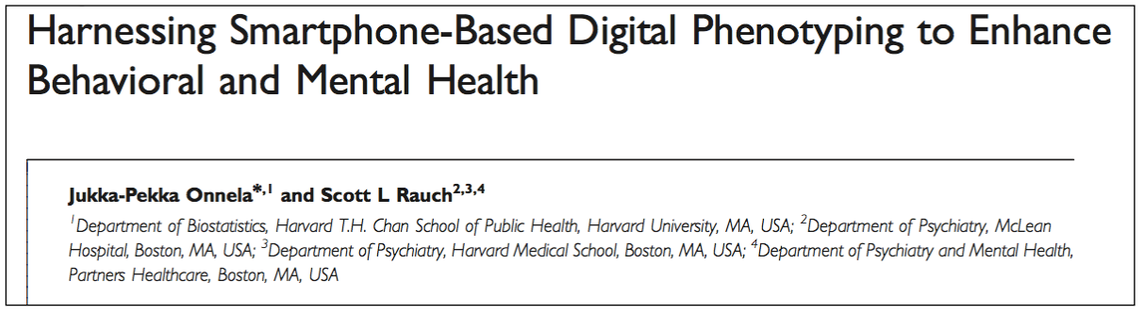
- My research at Harvard University has been sponsored by:
  - National Institutes of Health
  - Boehringer Ingelheim
  - MT Pharma
  - Otsuka Pharmaceutical
  - Siemens
  - Apple

- Consulting for Boehringer Ingelheim

- I'm a co-founder of PheBe Health, an early stage digital health startup company

DIGITAL PHENOTYPING

- Social, behavioral, and cognitive phenotypes present special challenges because of their temporal nature and context dependence
- Need a scalable way to measure markers objectively outside research labs and clinics
- Digital phenotyping*: "Moment-by-moment quantification of individual-level human phenotype using data from personal digital devices (smartphones)"
- Important advantages: large N and large T (using passive data)



Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health

Jukka-Pekka Onnela*,[1] and Scott L Rauch[2,3,4]

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, MA, USA; [2]Department of Psychiatry, McLean Hospital, Boston, MA, USA; [3]Department of Psychiatry, Harvard Medical School, Boston, MA, USA; [4]Department of Psychiatry and Mental Health, Partners Healthcare, Boston, MA, USA

New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research

John Torous[1,2], MD; Mathew V Kiang[3], MPH; Jeanette Lorme[4], BA; Jukka-Pekka Onnela[4], DSc

[1]Brigham and Women's Hospital, Department of Psychiatry, Harvard Medical School, Boston, MA, United States
[2]Beth Israel Deaconess Medical Center, Department of Psychiatry, Harvard Medical School, Boston, MA, United States
[3]Department of Social and Behavioral Sciences, Harvard TH Chan School of Public Health, Harvard University, Boston, MA, United States
[4]Department of Biostatistics, Harvard TH Chan School of Public Health, Harvard University, Boston, MA, United States

* Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health by JP Onnela and SL Rauch. Neuropsychopharmacology 41, 1691, 2016.

* New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research by J Torous, MV Kiang, J Lorme, and JP Onnela. JMIR Mental Health 3(2):e16, 2016.

DIGITAL PHENOTYPING

- Digital Phenotyping Project:
  1. Infrastructure
     - Customizable platform for high-throughput smartphone-based digital phenotyping
  2. Methods
     - Big data, noisy data, temporally dense, high-dimensional, missingness, etc.
  3. Studies
     - Several active studies in psychiatry, neurology, surgery, and other fields

- Goal: Systematize data collection and analysis in smartphone-based digital phenotyping
- Support started with 2013 NIH Director's New Innovator Award DP2MH103909

DATA COLLECTION AND DATA ANALYSIS

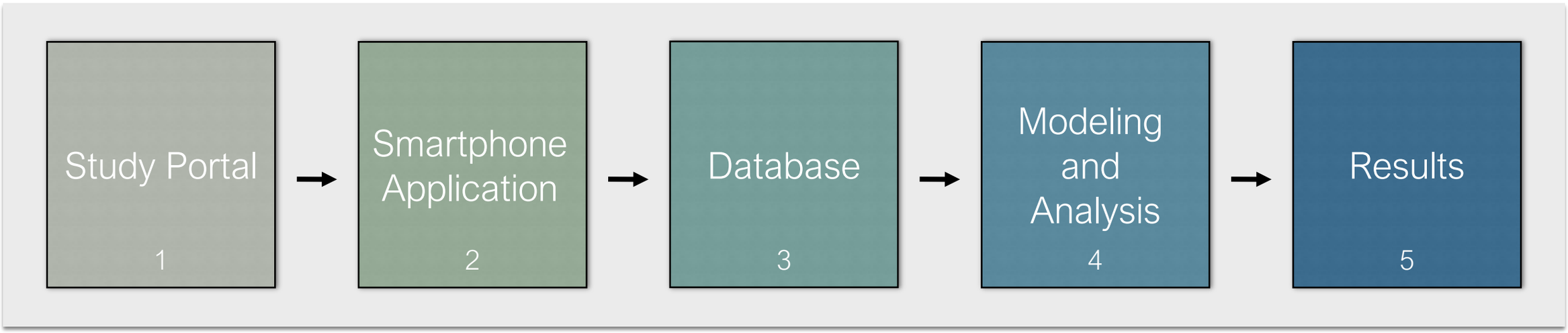
- Our **Beiwe** raw data collection platform in development and use since 2013
  - Name of a Sámi goddess of sunlight and mental health
  - Front-end: Android and iOS applications for collecting active and passive data
  - Back-end: AWS cloud computing, scalable, globally deployable

- Our **Forest** data analytics library incorporates an increasing number of our methods
  - Implemented in Python, runs locally and on AWS, integrates with Tableau

- Options for use
  - Beiwe Service Center, Harvard Service Core, under the software-as-a-service model
  - Self-deployment from open source under BSD-3 license from GitHub

WORKFLOW

- Workflow on the Beiwe research platform:
    1. Create and manage studies using web-based study portal
    2. Subjects download their app
    3. Collect and store hashed and encrypted raw data on AWS
    4. Model and analyze data using our data analysis pipeline
    5. Share results and configuration files (1-click reproducibility) with the community



STANDARD BEIWE DATA STREAMS

Active data

- Survey responses and metadata
- Audio diaries
- Ambient audio recording
- Image capture

Passive data

- GPS
- Accelerometer
- Gyroscope
- Magnetometer
- Phone and screen state
- WiFi
- Bluetooth
- Phone call logs (Android only)
- Text message logs (Android only)
- Proximity

TWO PATHS TO DATA

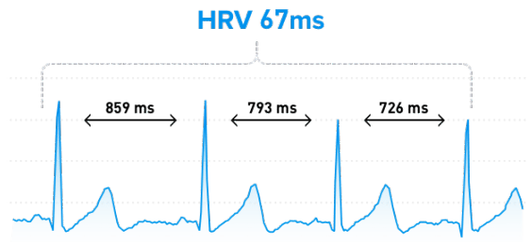
- **Using prepackaged data summaries** (Apple's and Google's SDKs)
  - Each has their own closed algorithms
  - Algorithms evolve fast, even within-person over time comparisons may not be valid
  - Low reproducibility of studies using proprietary algorithms
  - Need ways to quantify uncertainty from various sources

- **Working with raw data**
  - Can address all of the above challenges
  - Can re-analyze data using different algorithms and pool data
  - Data volumes are large and analytics are complex

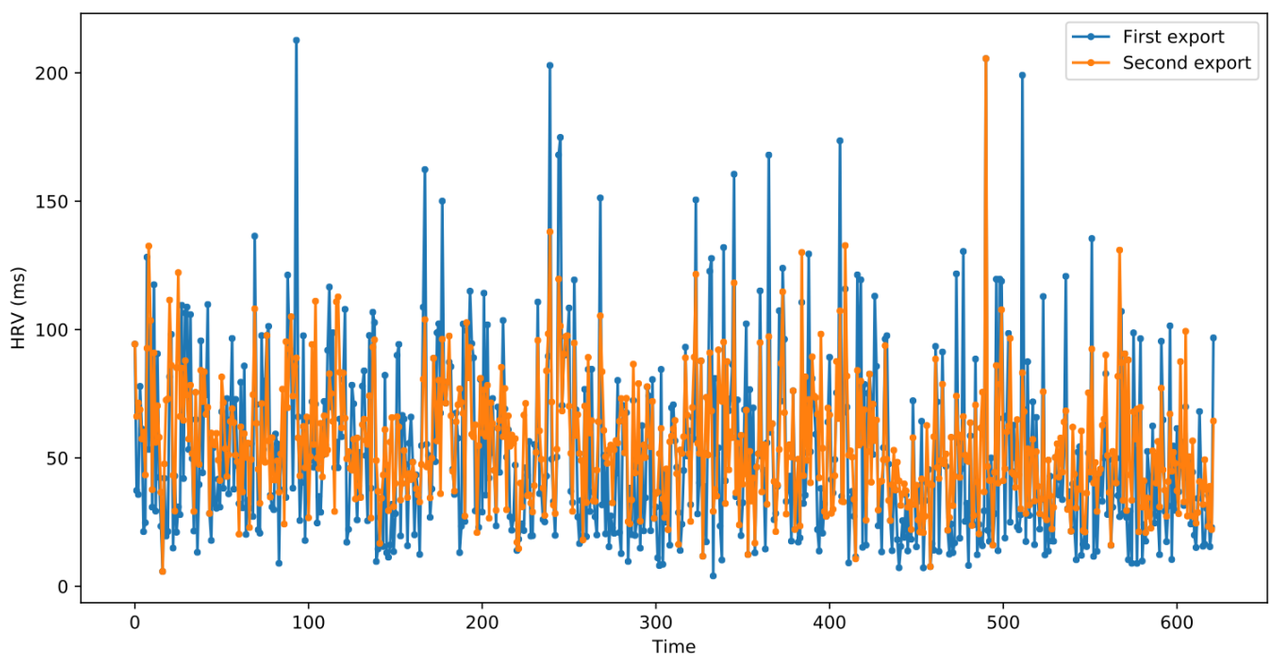
CHALLENGES WITH CLOSED ALGORITHMS

- Heart rate variability (HRV) can be measured using commercial wearables
- RR intervals are inter-beat intervals between successive heartbeats
- Time-domain characterization in terms of standard deviation of RR intervals (SDRR)
- Exported data twice from a commercial smartwatch

Start 12/2018 — Export 9/2020 — Export 4/2021



- 640 days in common across the two exports; missing data for 18 days



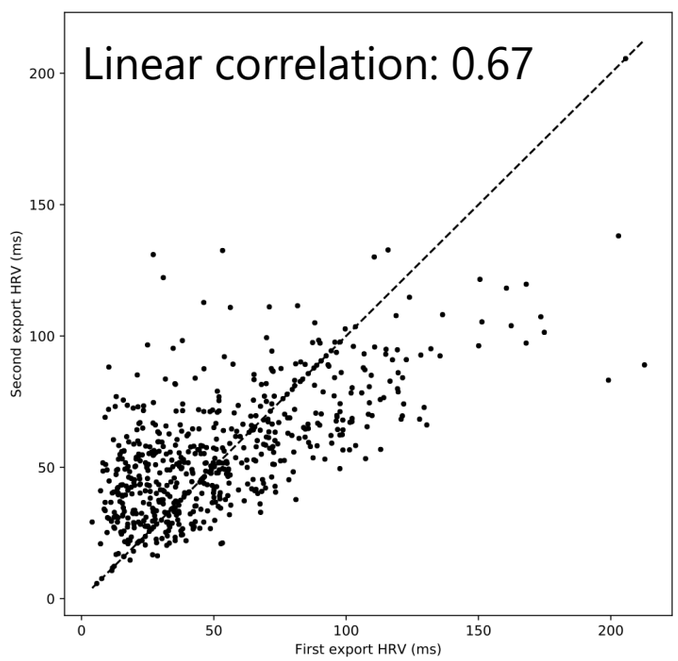
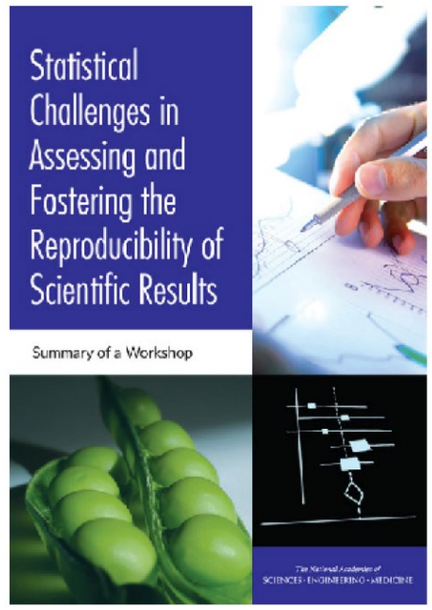


Figure shows RMSSD. Source: https://www.whoop.com/thelocker/heart-rate-variability-hrv/

REPRODUCIBILITY

- Reproducibility remains a substantial challenge (NAS report)
- All Beiwe settings (active and & passive) captured in a single JSON file
- Configuration files can be imported to and exported from Beiwe



Questions about the reproducibility of scientific research have been raised in numerous settings and have gained visibility through several high-profile journal and popular press articles. Quantitative issues contributing to reproducibility challenges have been considered (including improper data management and analysis, inadequate statistical expertise, and incomplete data, among others), but there is no clear consensus on how best to approach or to minimize these problems.

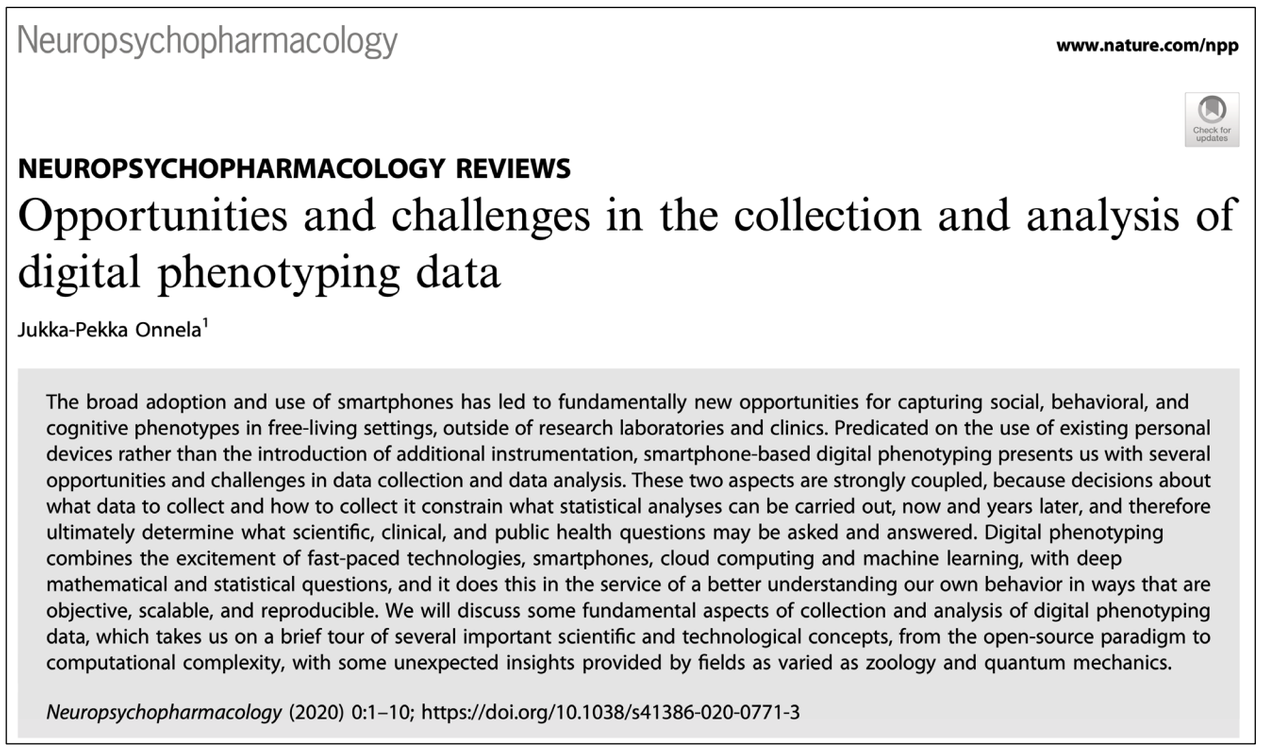
This is an issue across all scientific domains. A recent study found that 65 percent of medical studies were inconsistent when retested, and only 6 percent were completely reproducible (Prinz et al., 2011). The following year, a survey published in *Nature* found that 47 out of 53 medical research papers on the subject of cancer were irreproducible (Begley and Ellis, 2012). The Begley and Ellis *Nature* study was itself reproduced in the journal *PLOS ONE*, which confirmed that a majority of cancer researchers surveyed had been unable to reproduce a result.

A lack of reproducibility of scientific results has created some distrust in scientific findings among the general public, scientists, funding agencies, and industries. For example, the pharmaceutical and biotechnology industries depend on the validity of published findings from academic investigators prior to initiating programs to develop new diagnostic and therapeutic agents that benefit cancer patients. But that validity has come into question recently as investigators from companies have noted poor reproducibility of published results from academic laboratories, which limits the ability to transfer findings from the laboratory to the clinic (Mobley et al., 2013).


```
{
  "device_settings":{
    "accelerometer":true,
    "accelerometer_on_duration_seconds":60,
    "accelerometer_off_duration_seconds":60,
    "survey_submit_success_toast_text":"Thank you!”,
    "seconds_before_auto_logout":600,
    "check_for_new_surveys_frequency_seconds":21600,
    "create_new_data_files_frequency_seconds":900,
    "upload_data_files_frequency_seconds":3600
  },
  "surveys":[
    {
      "content":[
        {
          “question_text”:"Has it been difficult to fall asleep?”,
          "question_type":"slider",
          "min":"0",
          "max":"3",
          "question_id":"5d5f81b5-6d03-4d1e-ccca-a358b7d1b66e"
        },
```

CHALLENGES FOR NEW METHODS

- Our main focus is development of new statistical methods for DP data
- Main challenge in smartphone-based digital phenotyping is data analysis
- The challenges are due to behavior and technology (**heterogeneity**)
- **Our approach: tackle data modalities one at a time**
- **Our approach: aim for transparency and interpretability in modeling**
- Once individual modalities understood, combine them to a multimodal approach



Neuropsychopharmacology

www.nature.com/npp

NEUROPSYCHOPHARMACOLOGY REVIEWS

Opportunities and challenges in the collection and analysis of digital phenotyping data

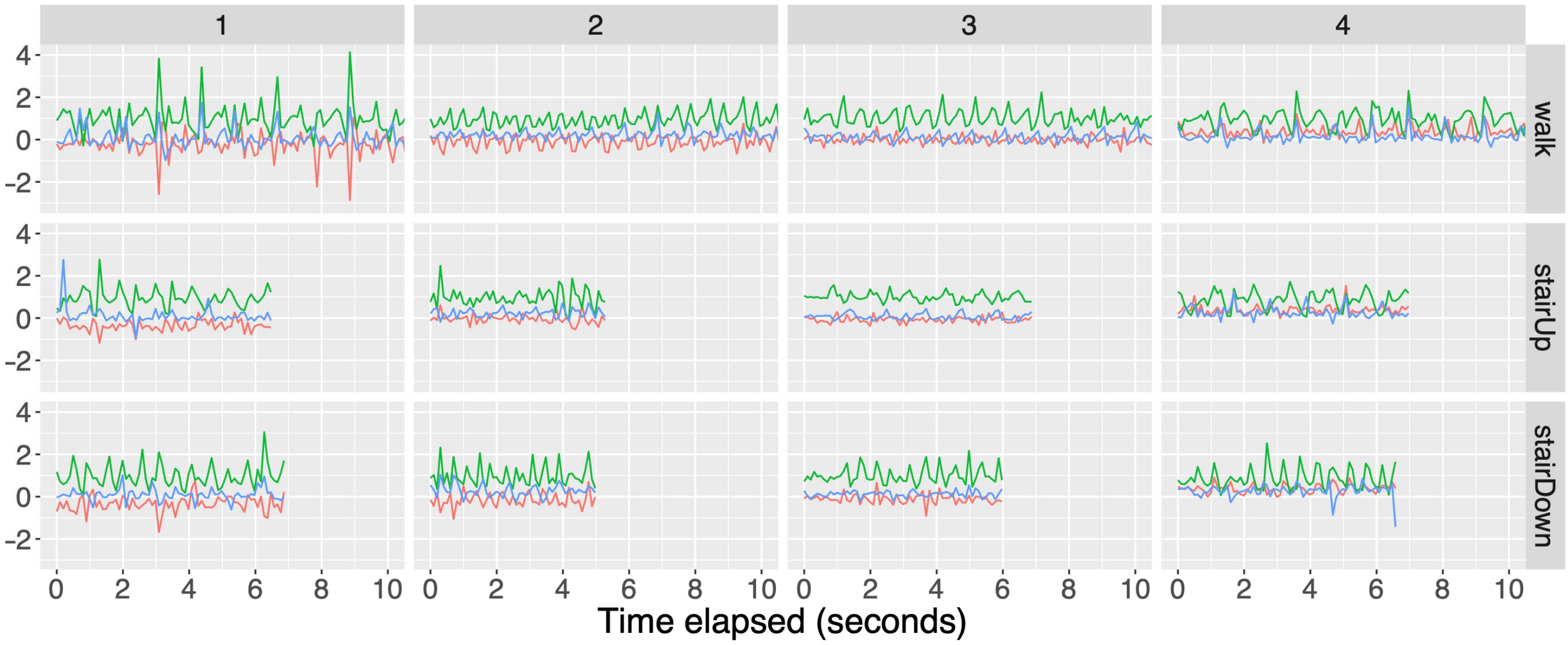
Jukka-Pekka Onnela[1]

The broad adoption and use of smartphones has led to fundamentally new opportunities for capturing social, behavioral, and cognitive phenotypes in free-living settings, outside of research laboratories and clinics. Predicated on the use of existing personal devices rather than the introduction of additional instrumentation, smartphone-based digital phenotyping presents us with several opportunities and challenges in data collection and data analysis. These two aspects are strongly coupled, because decisions about what data to collect and how to collect it constrain what statistical analyses can be carried out, now and years later, and therefore ultimately determine what scientific, clinical, and public health questions may be asked and answered. Digital phenotyping combines the excitement of fast-paced technologies, smartphones, cloud computing and machine learning, with deep mathematical and statistical questions, and it does this in the service of a better understanding our own behavior in ways that are objective, scalable, and reproducible. We will discuss some fundamental aspects of collection and analysis of digital phenotyping data, which takes us on a brief tour of several important scientific and technological concepts, from the open-source paradigm to computational complexity, with some unexpected insights provided by fields as varied as zoology and quantum mechanics.

Neuropsychopharmacology (2020) 0:1–10; https://doi.org/10.1038/s41386-020-0771-3

PHYSICAL ACTIVITY

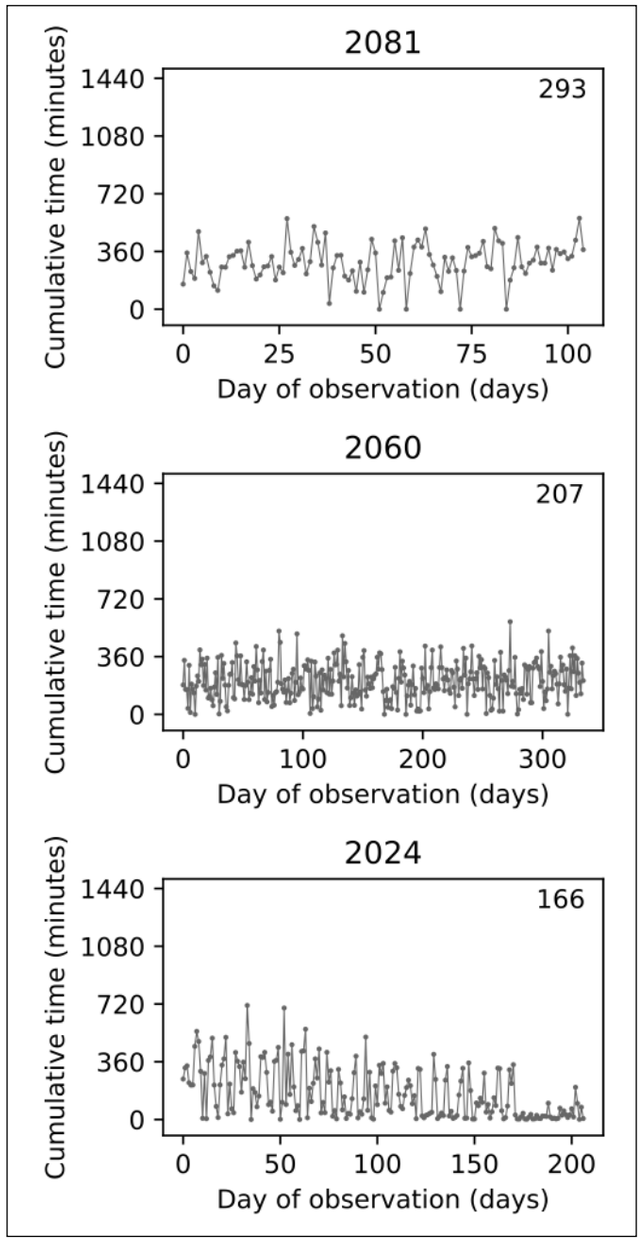
- Data from 4 subjects doing prescribed activities
- Same study phone, phone placement, phone orientation, and jeans
- **Triaxial accelerometer data** (units g = 9.81 m/s^2) over time



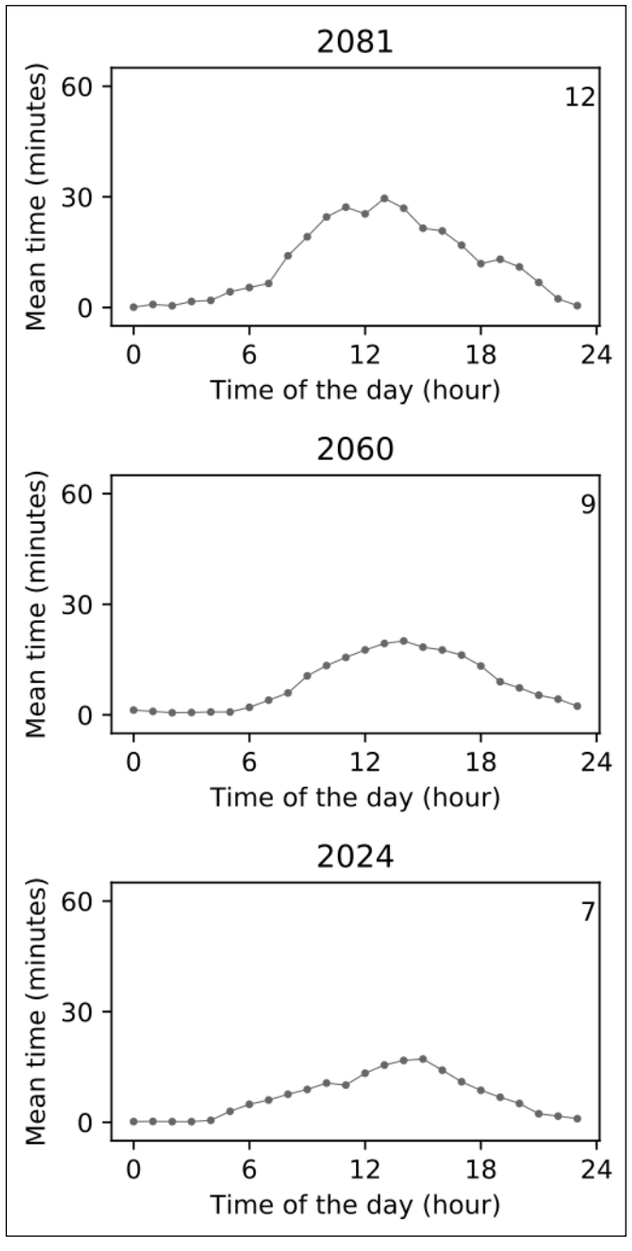
VARIABILITY IN WEAR TIME (SMARTPHONE)

- Estimated **smartphone** wear time in free-living settings

Daily cumulative wear time



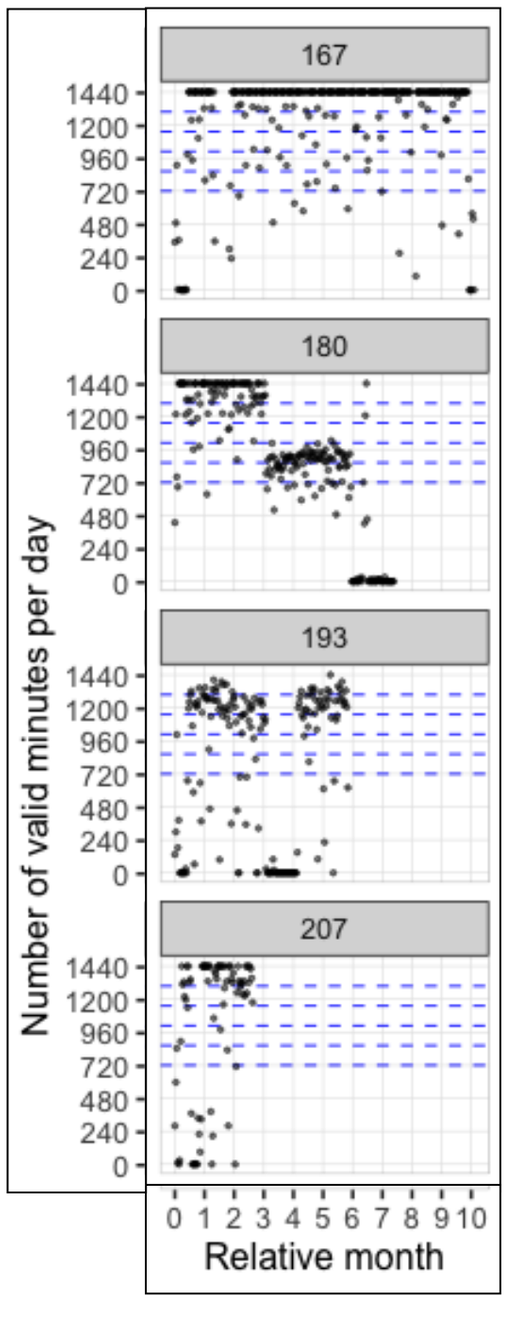
Mean hourly wear time



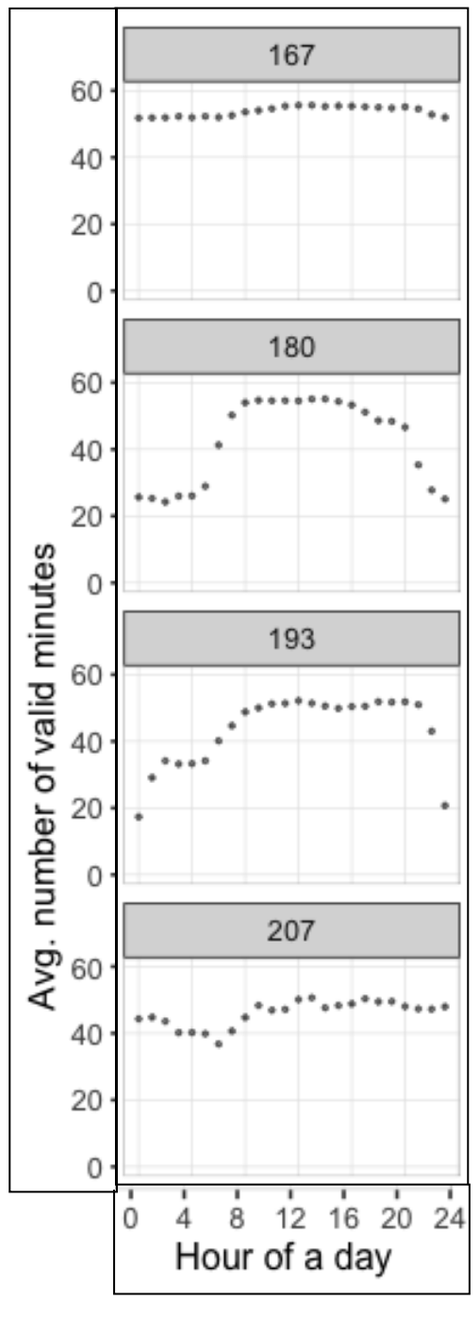
VARIABILITY IN WEAR TIME (WEARABLE)

- Estimated **wearable** wear time in an ALS clinical trial (high compliance)

Daily cumulative wear time

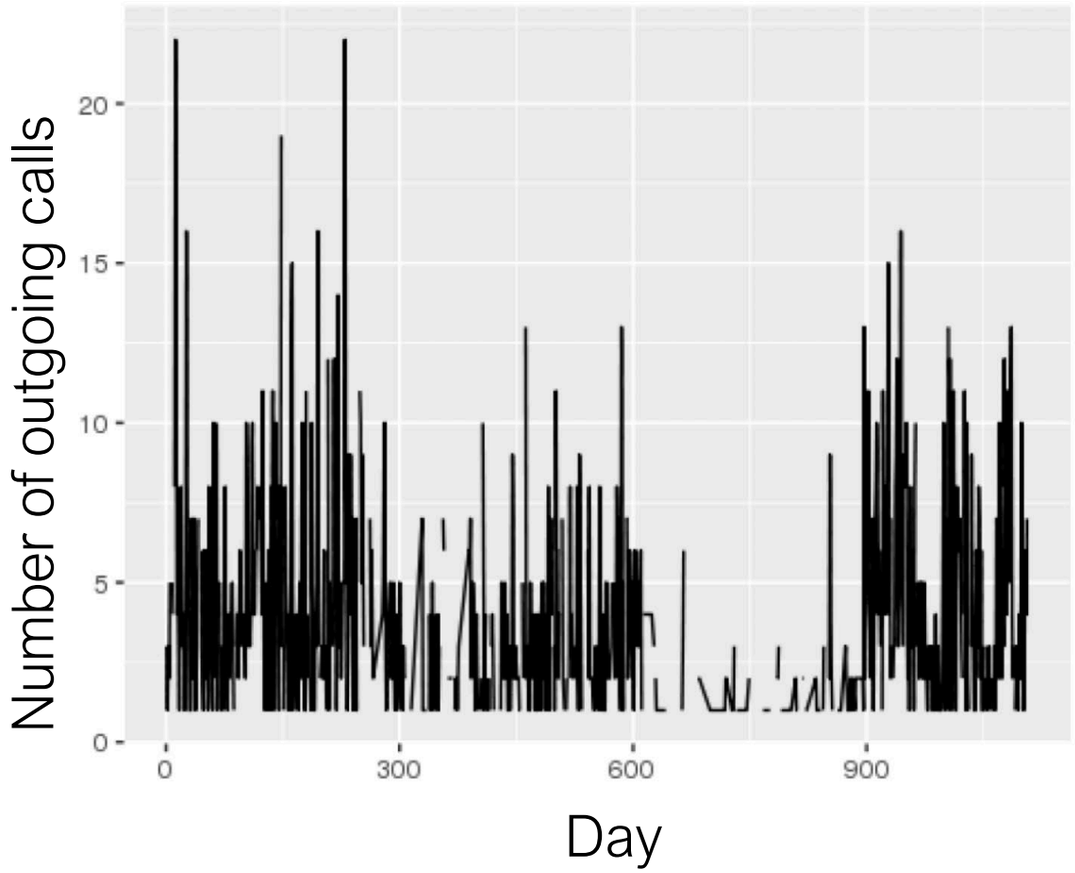


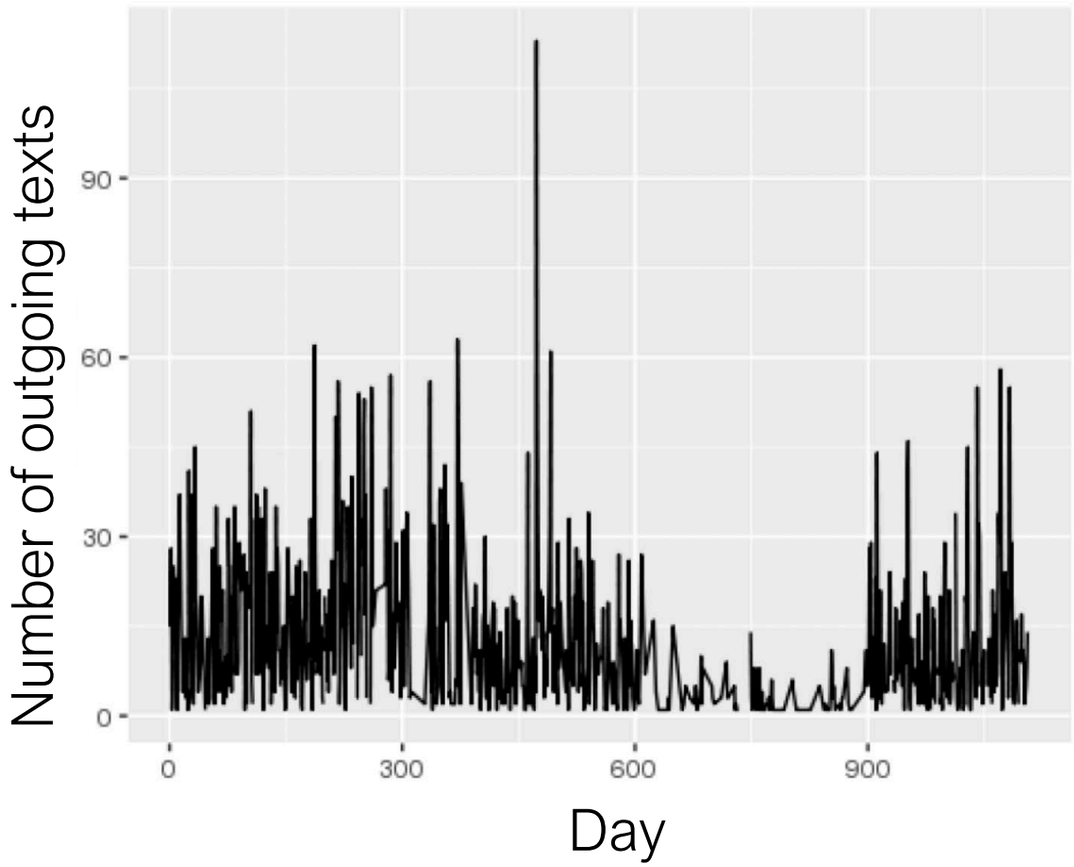
Mean hourly wear time



BIPOLAR DISORDER

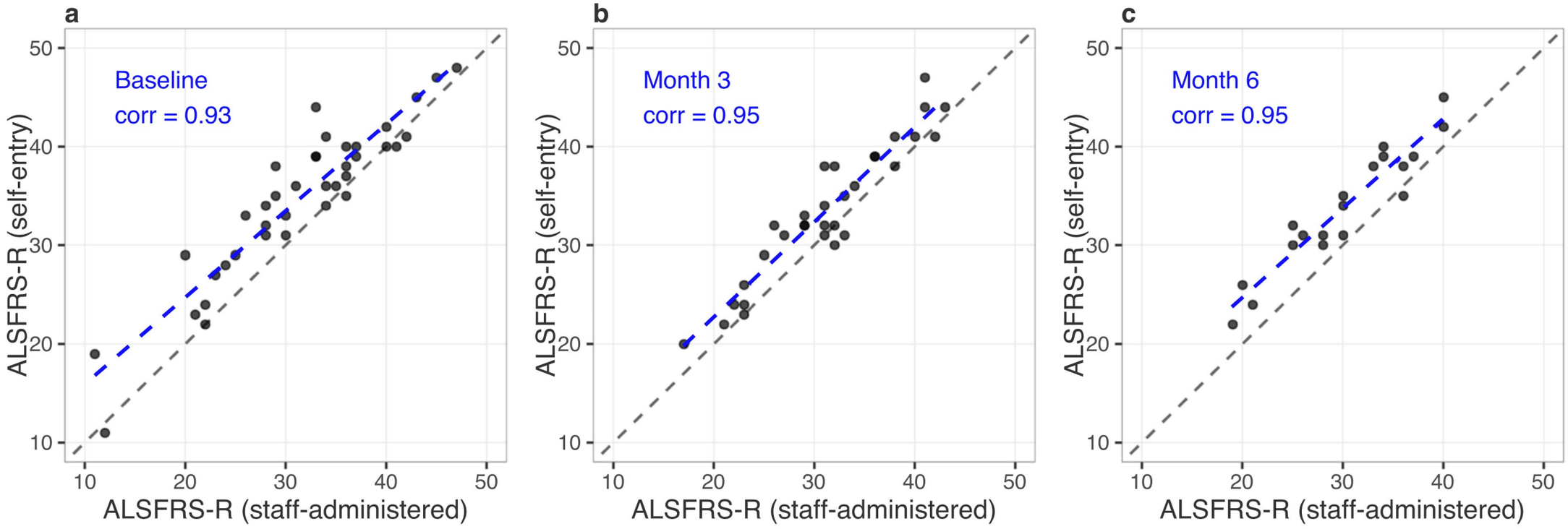
- Communication logs capture communication metadata for calls and text messages
- Bipolar disorder study at McLean Hospital using Beiwe
- Team includes Linda Valeri, Justin Baker (co-PI), JP Onnela (co-PI)





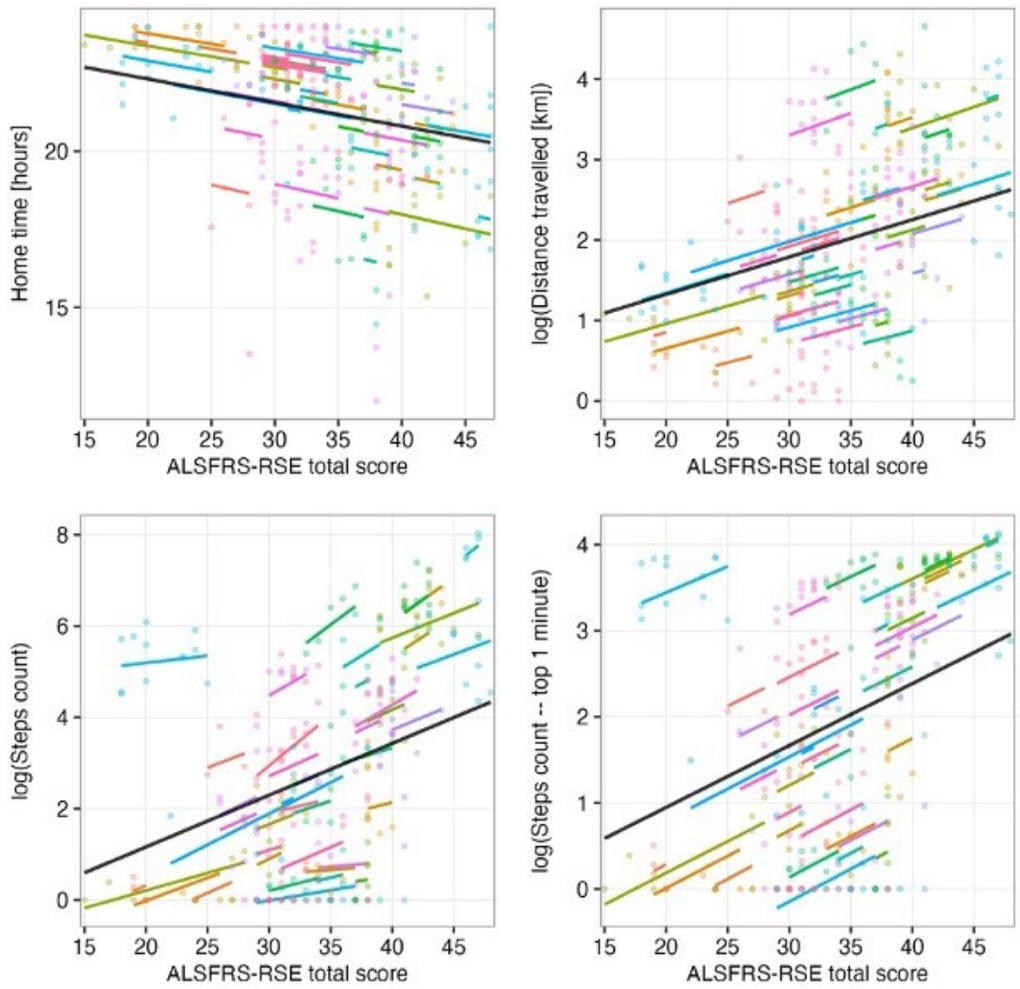
AMYOTROPHIC LATERAL SCLEROSIS (ALS)

- ALS is neurodegenerative motor neuron disease leading to progressive muscle weakness
- Common outcome is the revised ALS Functional Rating Scale (ALSFRS-R) score
  - Includes 12 questions, each scored 0–4, overall range 0–48 (full function)
  - The scale is typically administered by a clinician at a clinic
- Investigated phone-based self-administered (y-axis) and staff-administered (x-axis) scores



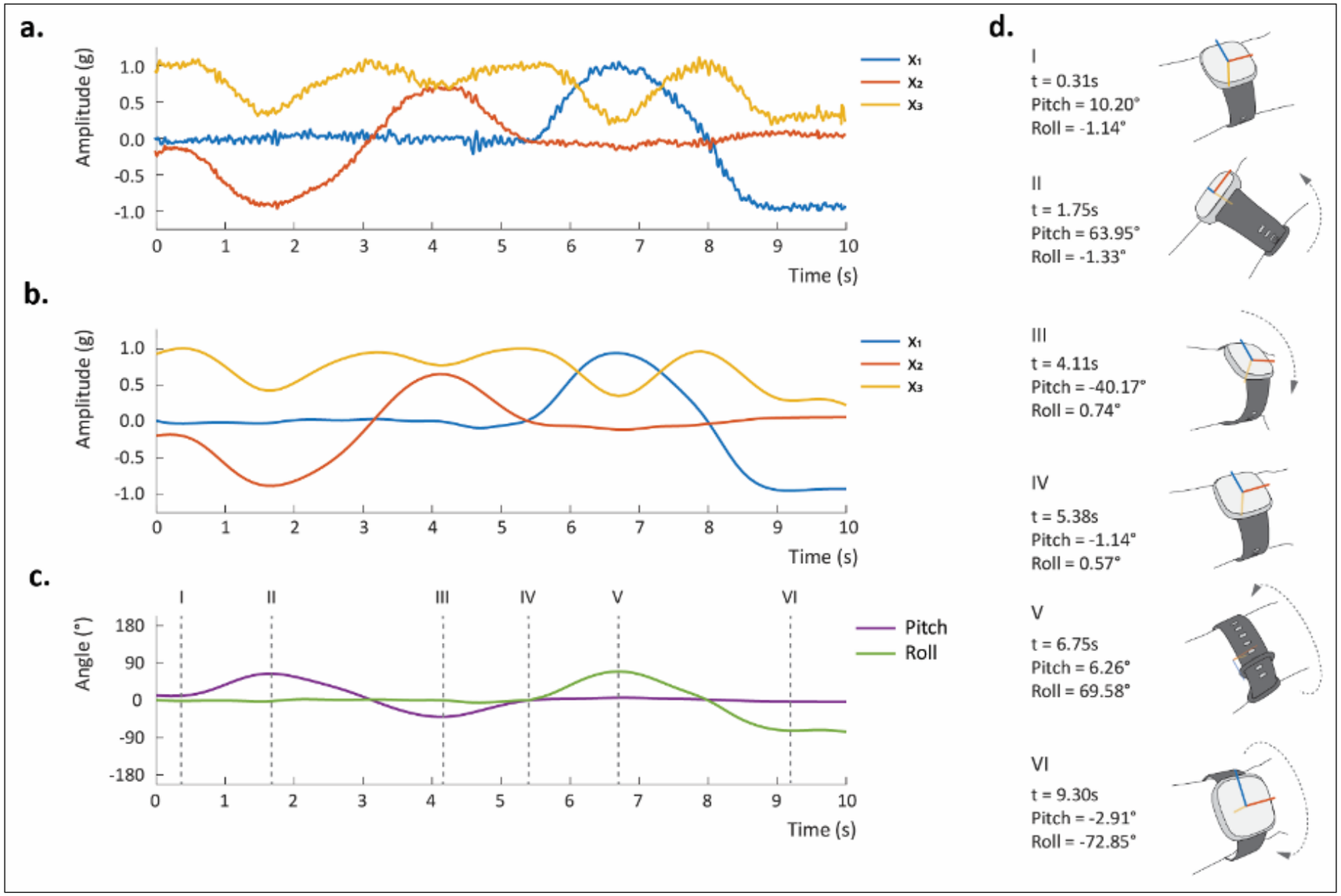
MOBILITY AND PHYSICAL ACTIVITY IN ALS

- Preliminary results from ongoing ALS trial: phones and research-grade wearables
- Self-entry ALS scale and smartphone-based outcomes
- Title: "Passively-collected smartphone data can track ALS disease progression"



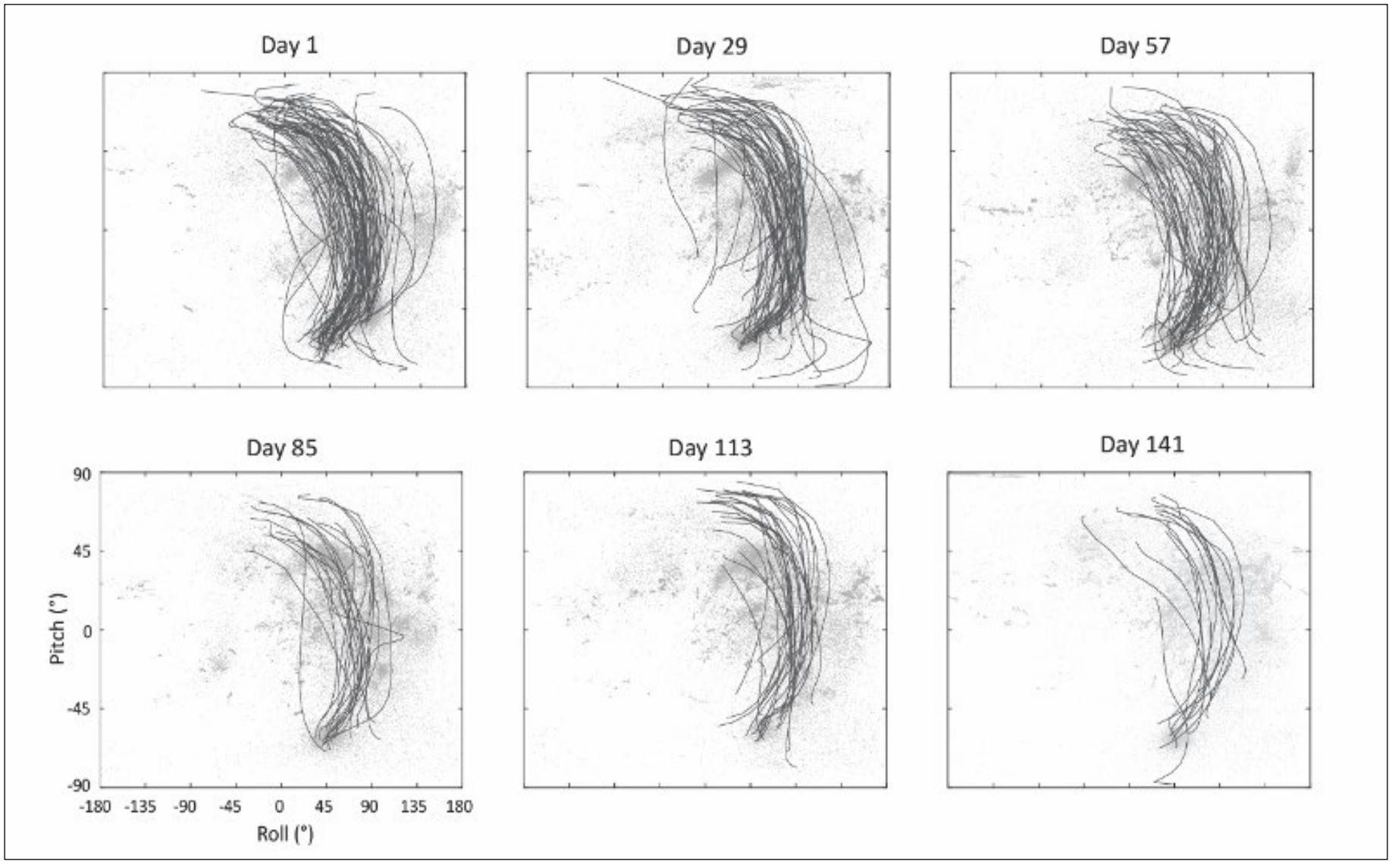
FOREARM MOVEMENT IN ALS

- Preliminary results from ongoing ALS trial: phones and research-grade wearables
- Title: "Forearm-movements as digital biomarkers in people with ALS"



FOREARM MOVEMENT IN ALS

- Preliminary results from ongoing ALS trial: phones and research-grade wearables
- Title: "Forearm-movements as digital biomarkers in people with ALS"



LESSONS LEARNED 2013–2023

1. Move from group-level averages to individual-level temporal trajectories

2. Smartphones maximize heterogeneity and minimize inequities in participation

3. Smartphone passive data minimizes participant burden and maximizes adherence

4. Wearables provide complementary data in short–medium term

5. Raw data opens up scientific opportunities, enables re-analyses, enhances reproducibility

6. Need trustworthy and transparent methods for insights, one modality at a time

7. Building collaborations with domain experts takes years but is the only way forward

ACKNOWLEDGEMENTS

- Everything is a team effort
- With gratitude to:
    - Onnela Lab members (students, postdocs, staff)
    - Clinical collaborators (MGH, McLean)
    - Funders (NIH, philanthropic funding)
    - Patients

Lab: https://www.hsph.harvard.edu/onnela-lab/

Beiwe: https://www.beiwe.org/

Twitter: jponnela