

Session I:

Towards More Precise and Meaningful Measurement

Jukka-Pekka “JP” Onnela

Associate Professor

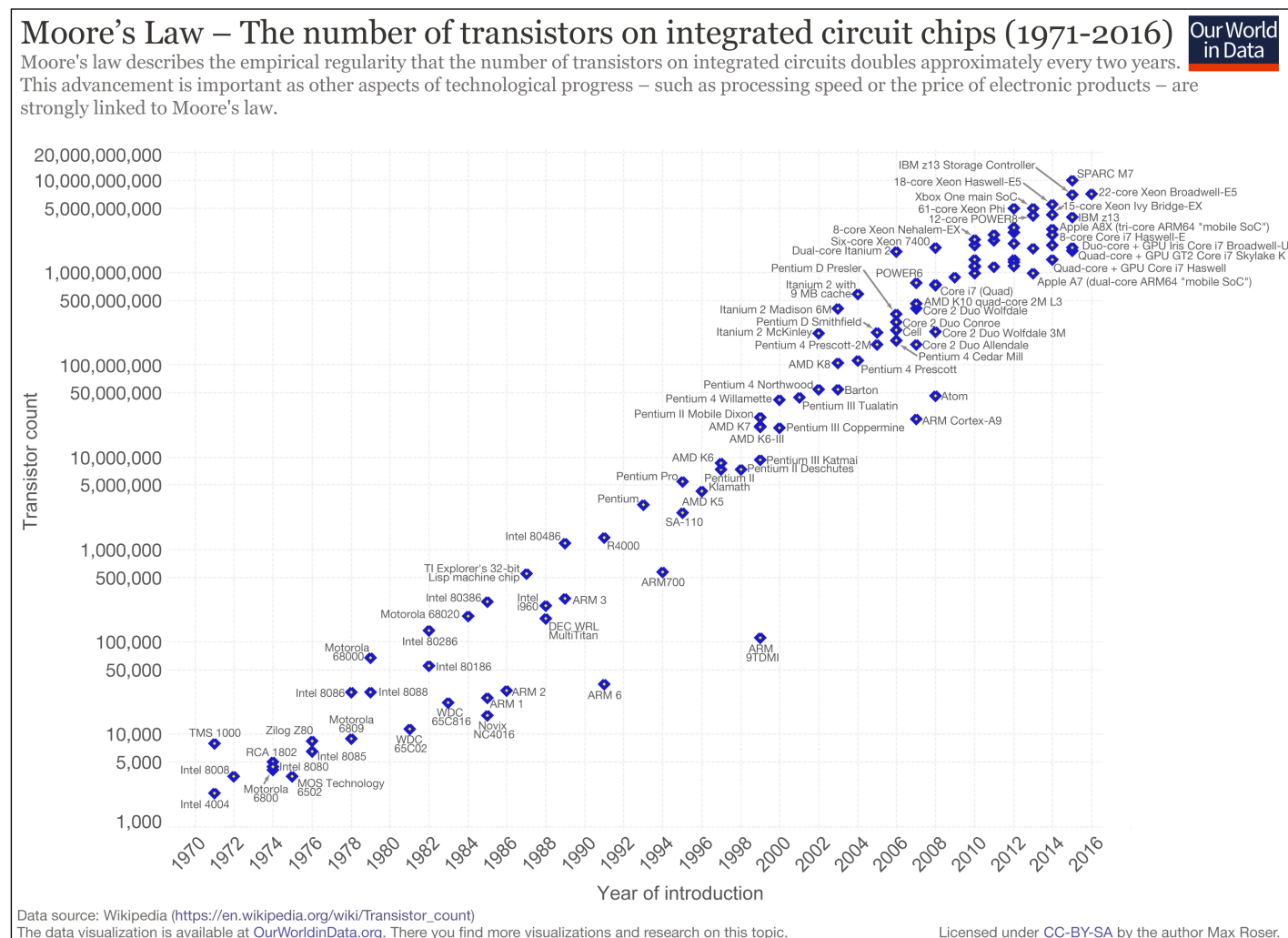
Department of Biostatistics

Harvard University

June 5, 2018

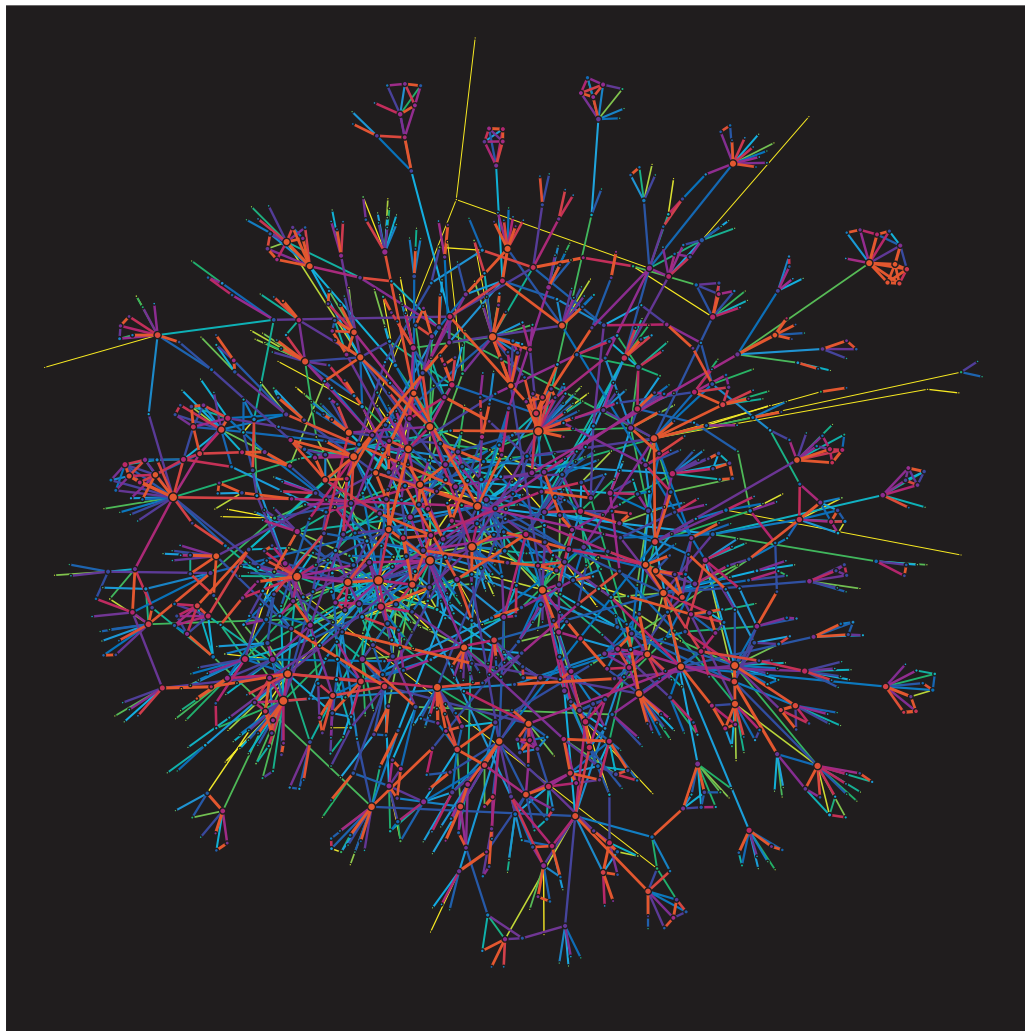
SENSORS, DATA, COMPUTATION

- Progress in science always driven by data
- Dramatic increase in data volume and types in the past 10 years, especially about human behavior
- Origin in Moore's law: transistor density doubles about every 18 months
- **Data:** Sensors are small, cheap, everywhere
- **Analysis:** Computational capacity, increasingly cloud-based

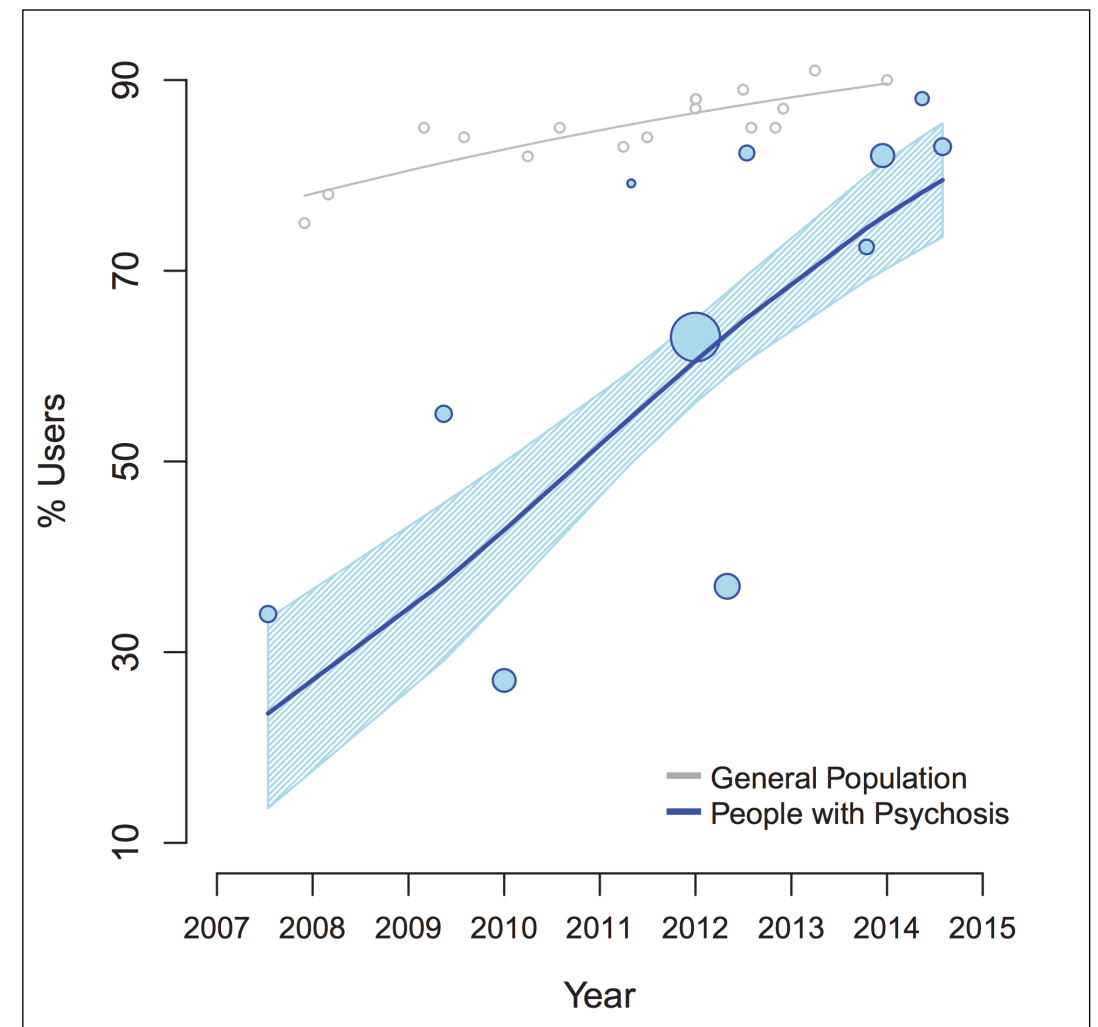


CELL PHONES AND SMARTPHONES

- More mobile phones in use in 2015 than people in the world
- There were 3.2 billion smartphone subscriptions in 2015 and 6.3 billion are expected by 2022
- US adults with smartphones: 35% in 2011 and 77% in 2017
- **Increasingly rich data in real-world settings about social, cognitive, and behavioral functioning**



Onnela, Saramäki, Hyvönen, Szabó, Lazer, Kaski, Kertész, Barabási; Structure and tie strengths in mobile communication networks, PNAS 104, 7332, 2007.



Firth et al. "Mobile phone ownership and endorsement of "mHealth" among people with psychosis: a meta-analysis of cross-sectional studies." Schizophrenia Bulletin (2015): sbv132.

PHENOTYPING CHALLENGE

- Over the past 20 years, many investigators have proposed phenomics as the natural complement to genome sequencing as a route to advances in the biomedical sciences
- **Behavior presents special challenges because of its temporal nature and context dependence**
- Need a **scalable** way to **measure** social, behavioral and cognitive markers **objectively in the wild**
- **Digital phenotyping**: “Moment-by-moment quantification of the individual-level human phenotype in situ using data from personal digital devices, in particular smartphones”
- Important advantages (passive data): **(1) large N, (2) large T, (3) pre and post data**

Phenomics: the next challenge

David Houle*, Diddahally R. Govindaraju¹ and Stig Omholt^{2,3}

Abstract | A key goal of biology is to understand phenotypic characteristics, such as health, disease and evolutionary fitness. Phenotypic variation is produced through a complex web of interactions between genotype and environment, and such a ‘genotype–phenotype’ map is inaccessible without the detailed phenotypic data that allow these interactions to be studied. Despite this need, our ability to characterize phenomes — the full set of phenotypes of an individual — lags behind our ability to characterize genomes. Phenomics should be recognized and pursued as an independent discipline to enable the development and adoption of high-throughput and high-dimensional phenotyping.

Neuroscience 164 (2009) 30–42

PHENOMICS: THE SYSTEMATIC STUDY OF PHENOTYPES ON A GENOME-WIDE SCALE

R. M. BILDER,^{a,b,*} F. W. SABB,^{a,*} T. D. CANNON,^{a,b,*}
E. D. LONDON,^{a,c,*} J. D. JENTSCH,^{b,*}
D. STOTT PARKER,^d R. A. POLDRACK,^{b,*} C. EVANS^{a,*}
AND N. B. FREIMER^{a,*}

^aDepartment of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine at University of California, Los Angeles

promise of personalized medicine and the rational diagnosis and treatment of neuropsychiatric syndromes. © 2009 IBRO. Published by Elsevier Ltd. All rights reserved.

Key words: phenotype, genetics, genomics, informatics, cognition, psychiatry.

The Human Phenome Project

Nelson Freimer^{1,2,3} & Chiara Sabatti^{3,4}

A principal goal of genetic research is to identify specific genotypes that are associated with human phenotypes. It will soon be possible to conduct genome-wide genotyping on a massive scale. Our current approaches for defining and assaying phenotypes may be inadequate for making optimal use of such genotypic data. We propose an international effort to create phenomic databases, that is, comprehensive assemblages of systematically collected phenotypic information, and to develop new approaches for analyzing such phenotypic data. We term this effort the Human Phenome Project and suggest a scientific and organizational scope for the project.

Deep Phenotyping for Precision Medicine

Peter N. Robinson*

Institut für Medizinische Genetik und Humangenetik, Charité - Universitätsmedizin Berlin, Berlin, Germany

For the Deep Phenotyping Special Issue

Received 27 February 2012; accepted revised manuscript 2 March 2012.

Published online in Wiley Online Library (www.wiley.com/humanmutation).DOI: 10.1002/humu.22080

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

Jukka-Pekka Onnela^{*1} and Scott L. Rauch^{2,3,4}

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

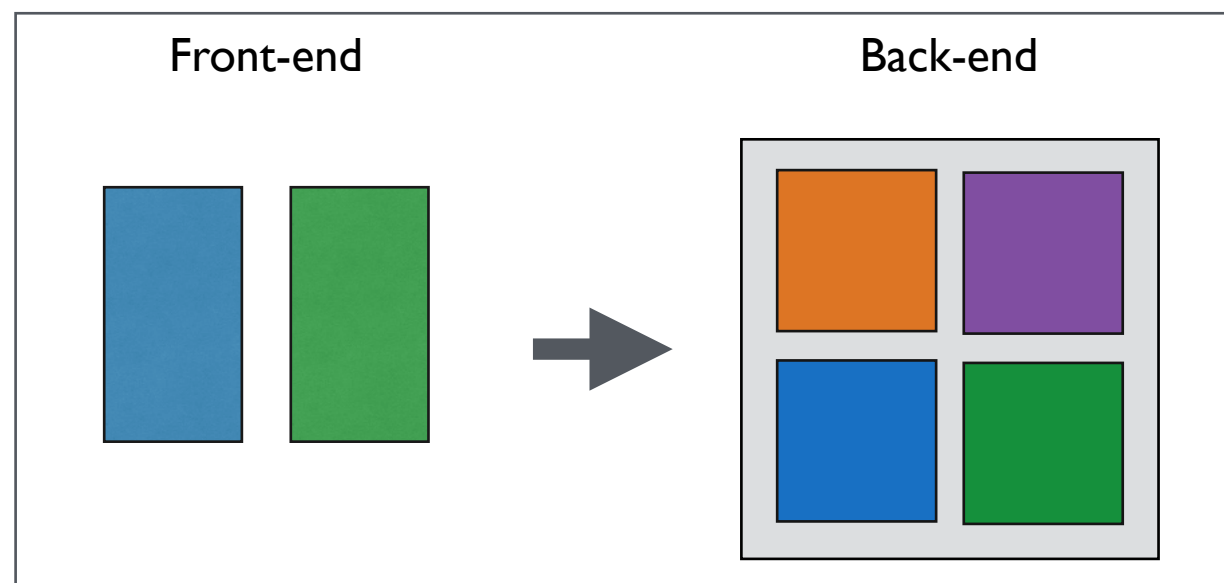
J Torous, MV Kiang, J Lorme, & JP Onnela. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR mental health, 3(2), 2016.

DIGITAL PHENOTYPING

- **Digital Phenotyping Project** (2013 NIH Director's New Innovator Award DP2MH103909):
 - **Infrastructure:** Develop customizable, scalable, open source research platform for high-throughput smartphone-based digital phenotyping (Beiwe)
 - **Methods:** Big and noisy data, temporally dense, missingness, high-dimensional, etc.
 - **Studies:** Used in ~30 studies mainly at Harvard Medical School teaching hospitals (McLean, MGH, BWH, BIDMC) in various fields (psychiatry, neuroscience, neurology, neurosurgery)
- Overall goal is to systematize data collection and analysis in smartphone-based digital phenotyping

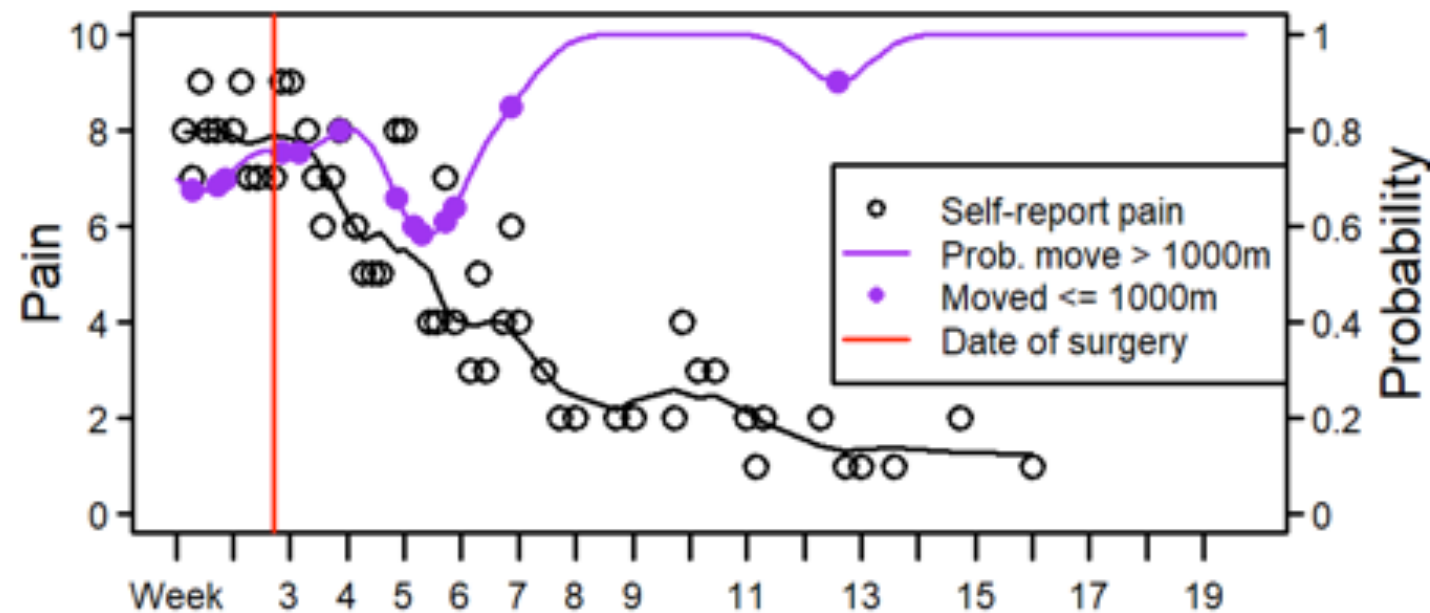
BEIWE RESEARCH PLATFORM

- **Front-end:** Android and iOS applications for collecting active and passive data
- **Back-end:** Amazon Web Services (AWS) cloud computing infrastructure
 - Scalability is key: about 1 GB of **raw data** per subject-month
 - **Web interface:** Study portal for managing studies and accessing data
 - **Data collection:** EC2 instances and Elastic Beanstalk
 - **Data storage:** S3 for medium-term and Amazon Glacier for long-term
 - **Data analysis pipeline:** Data processing and modeling (automatic & manual)
- **Open source** under the 3-clause BSD license (<https://github.com/onnella-lab>)



SPINE SURGERY

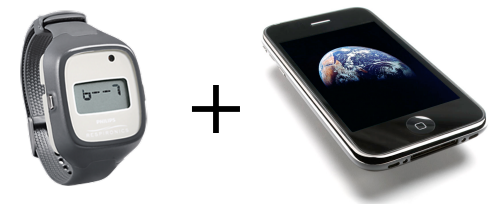
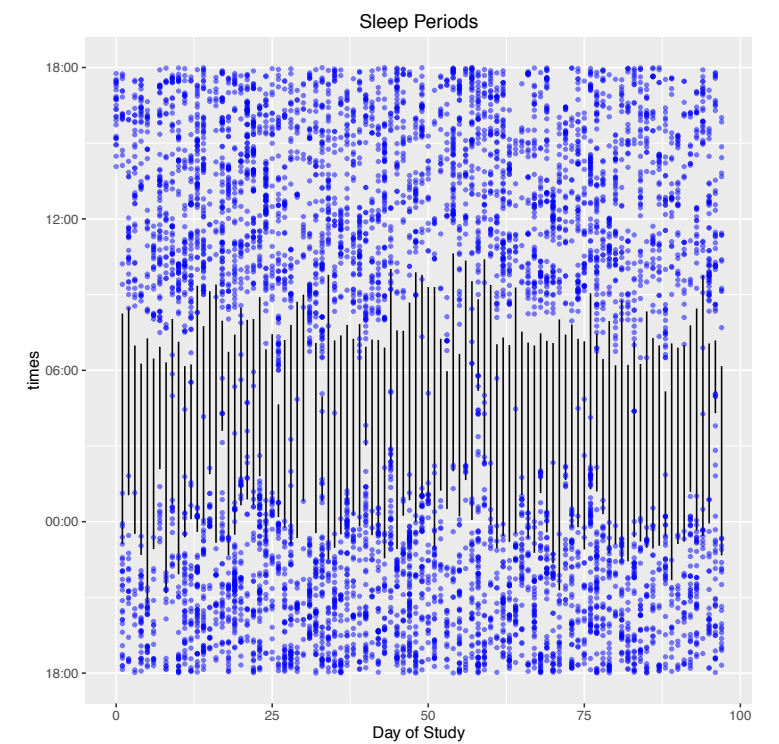
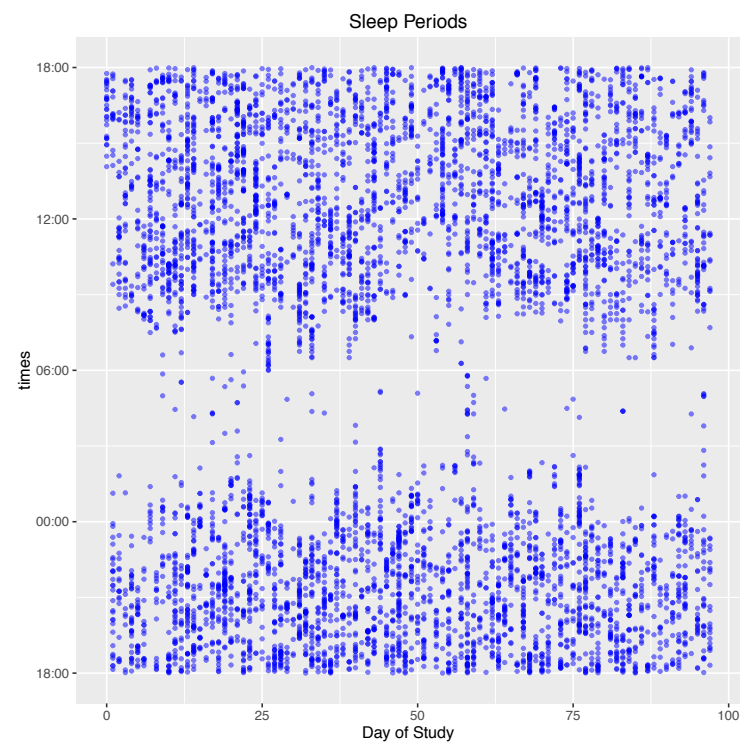
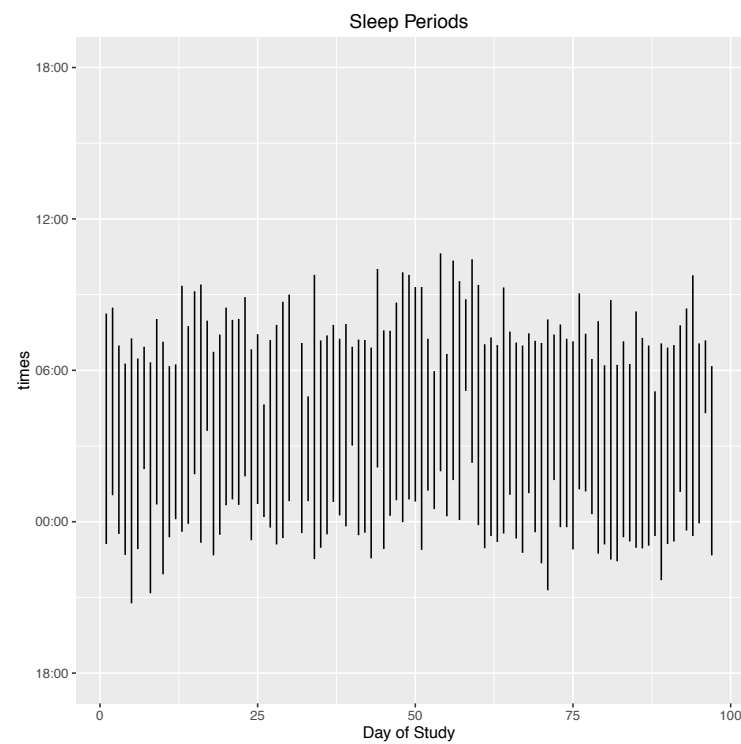
- Studied 105 patients with spine disease at BWH
- 55 patients underwent a surgical intervention during follow up
- By the end of follow up, the mean change in pain for all patients was -1.3 points (from 4.96 to 3.66)
- Increased pain was significantly associated with reduced patient mobility as measured by three daily GPS summary statistics



Join work with Ian Barnett, David J Cote, Timothy R Smith

QUANTIFYING SLEEP

- Social and academic transitions of Harvard undergraduate students
- Wearable, Beiwe on smartphone, 15 MRI scans



Join work with Randy Buckner, Maheen Shermohammed, Justin Baker, Ian Barnett, Patrick Staples

OPPORTUNITIES AND CHALLENGES

Opportunities:

- Scale (N) and length (T) of measurements
- Essentially identical measurements everywhere using smartphones (not so for wearables)
- Building partnerships across industry, academia, government

Challenges:

- Data are very high dimensional and very noisy
- Many analytical and statistical challenges in these early days
- Data standards and reproducibility
- Data security and patient privacy
- Regulatory considerations
- Patient engagement and feedback (possibly useful, can be harmful, constitutes an intervention)
- Integration into clinical care and EHR (information overload)