My First Exposure to Accelerometer Data was for 100000 People from UK Biobank

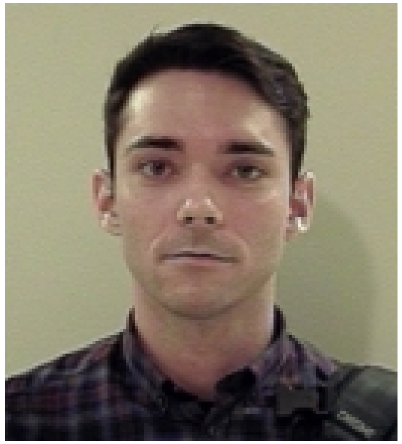
John Muschelli, @StrictlyStat, Johns Hopkins University

presentation: http://johnmuschelli.com/CMStat_2018
source: http://github.com/muschellij2/CMStat_2018

Processing math: 58%

Disclaimer (Acknowledgements?)

Work done with
Andrew Leroux
PhD Student at Johns
Hopkins



Processing math: 58%

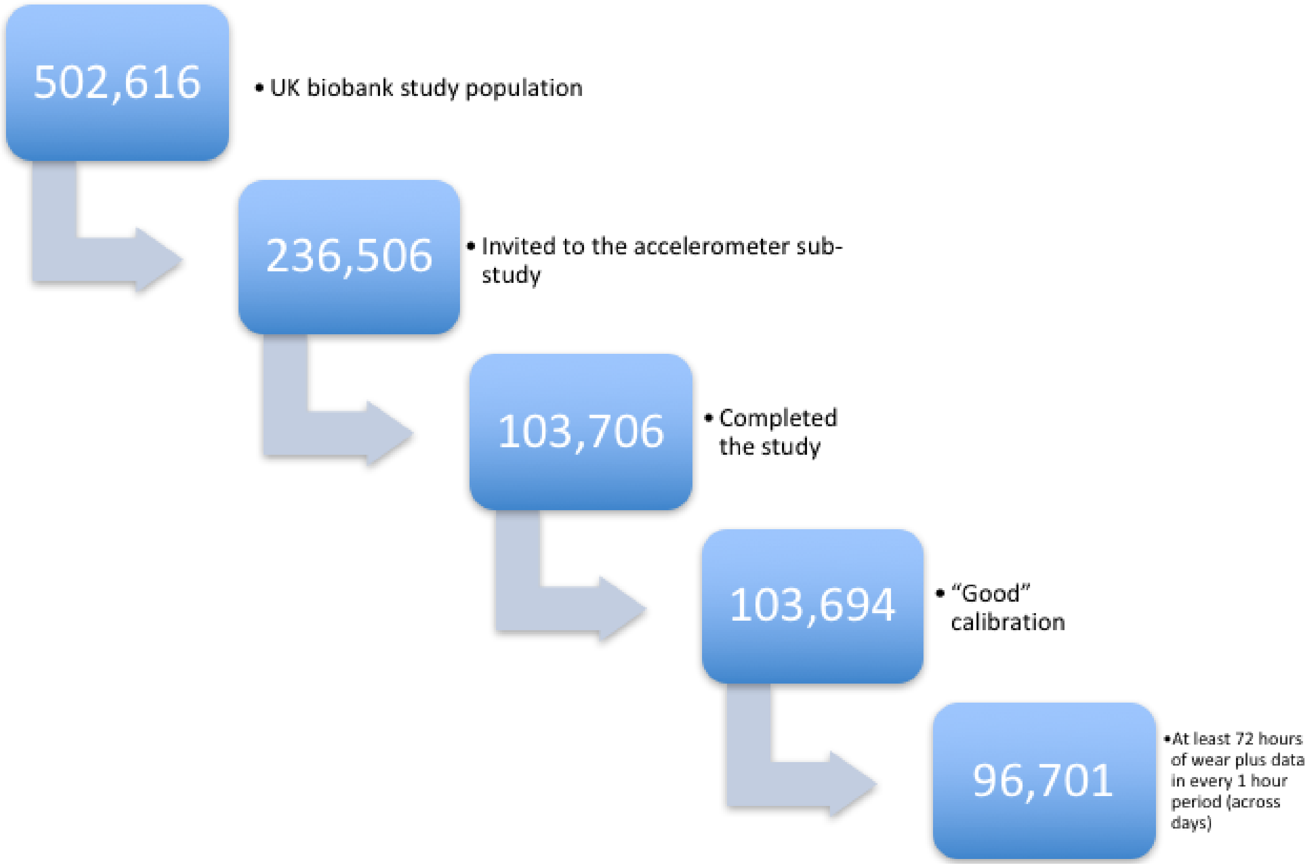
UK Biobank Data

- Overall 500,000 participants (UK), ≈ 100,000 included in the sub-study

In this presentation:

- Explore the data
- Assess “bias” in different devices (see if “autocalibration” is working)
- Also discuss inclusion criteria “bias”
- Get similar findings (Doherty et al. 2017) analysis

Where do 100K come from?

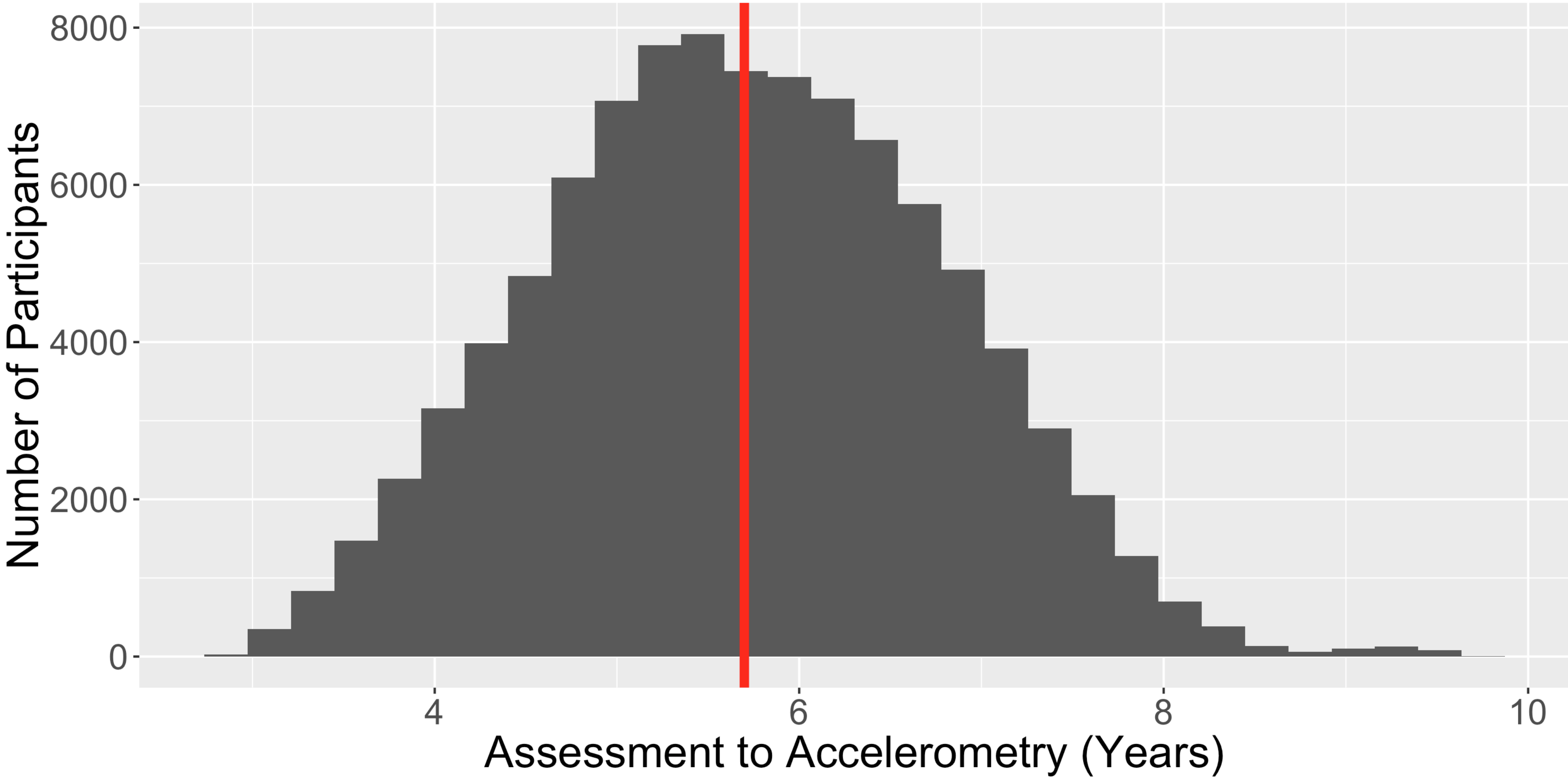


Demographics: Lots of Non-Response

	Completed: good data	Completed: bad data	No response	Not asked
n	96701	7005	**132800**	**266110**
Age at Initial Visit (mean (sd))	56.6 (7.8)	55.2 (7.9)	**56.4 (8.0)**	**57.5 (8.1)**
Male (% Male)	42255 (43.7)	3156 (45.1)	**62601 (47.1)**	**121151 (45.5)**
Ethnicity (% Non-White)	2983 (3.1)	335 (4.8)	**7617 (5.8)**	**16102 (6.1)**

Many people DIED before being able to be asked

Assessment to Accelerometry can be a WHILE



Processing math: 58%

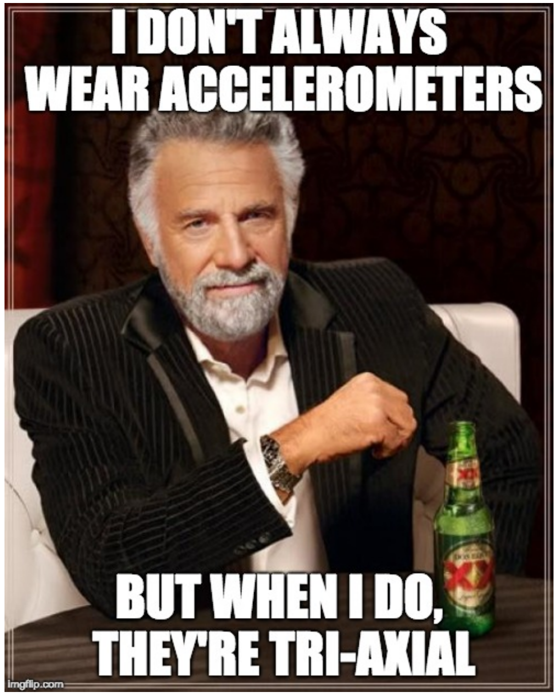
Responders are Healthier (Self-Reported)

	Completed: good data	Completed: bad data	No response	Not asked
Overall health (%)				
Excellent	**20987 (21.8)**	**1464 (21.0)**	**21583 (16.3)**	**37849 (14.4)**
Good	57849 (60.0)	4057 (58.1)	78968 (59.7)	148196 (56.2)
Fair	15149 (15.7)	1261 (18.0)	26669 (20.2)	62313 (23.6)
Poor	2482 (2.6)	205 (2.9)	4969 (3.8)	15124 (5.7)

Lesson #1: The devil is in the inclusion criteria (or can be)

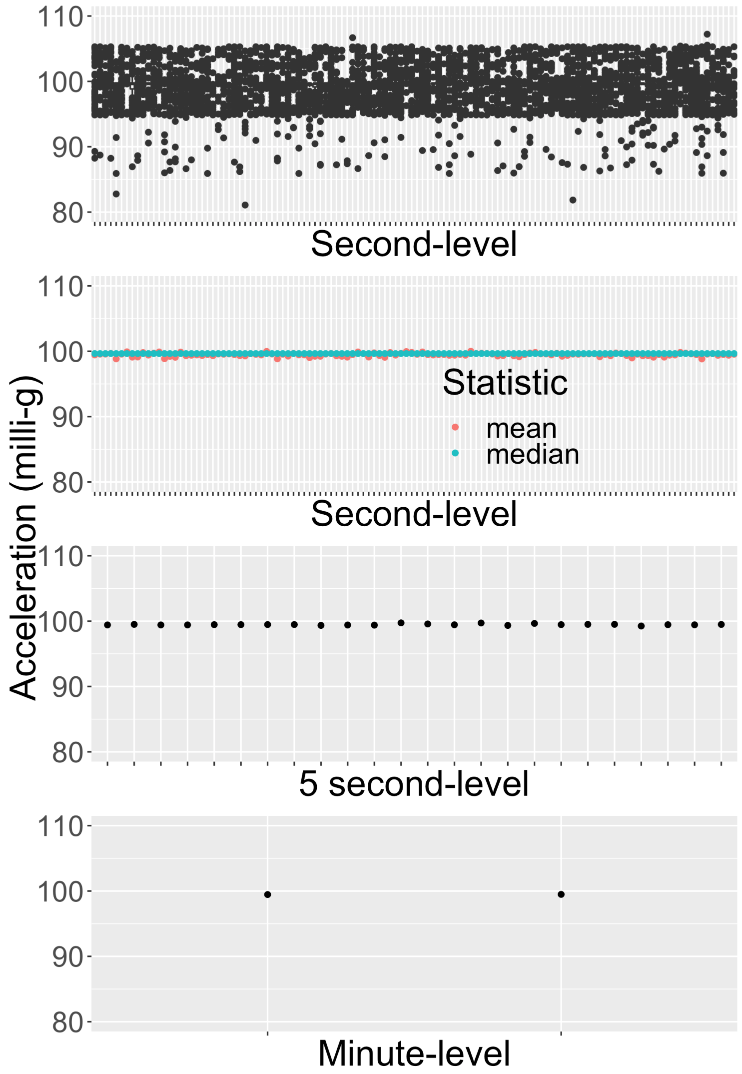
Data Gathered

- Tri-axial Axtivity 100Hz over 7 days
- Started at 10AM and ended at 10PM (spoiler: will be important)
- Data measured in milli-g (1g = 9.80665 ms^2)
    - not counts or steps as other devices



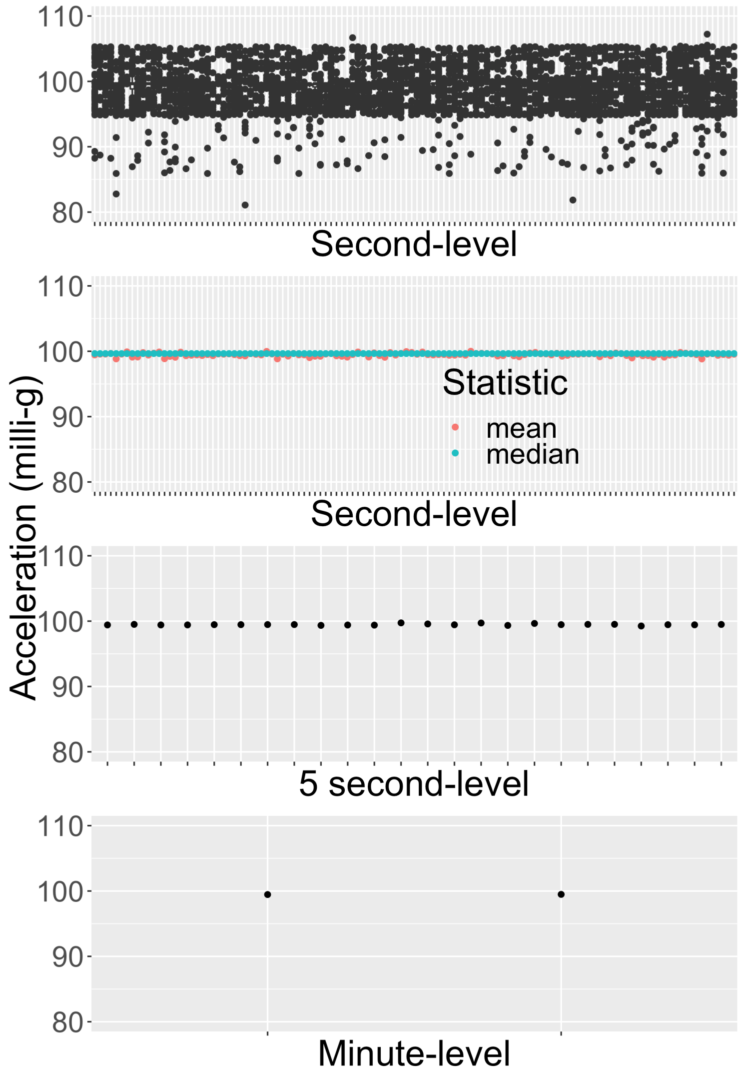
Accelerometry Data Available

- Data at varyling levels
  - Axtivity CWA format (Highest resolution, 100Hz) (200Mb per user)
    - very large for 100K subjects (20Tb)
- 5 second level data
  - UKBB imputation/processing done
  - averaged into 1440 minute-level data
- Overall statistics (mean/median): overall, daily, hourly, day of week
  - removed “non-wear” periods



Accelerometry Data Available

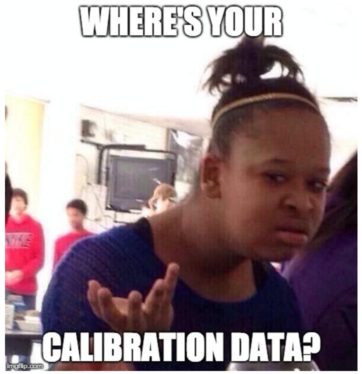
- Data at varyling levels
  - Axtivity CWA format (Highest resolution, 100Hz) (200Mb per user)
    - very large for 100K subjects (20Tb)
- **5 second level data**
  - **UKBB imputation/processing done**
  - **averaged into 1440 minute-level data**
- Overall statistics (mean/median): overall, daily, hourly, day of week
  - removed “non-wear” periods

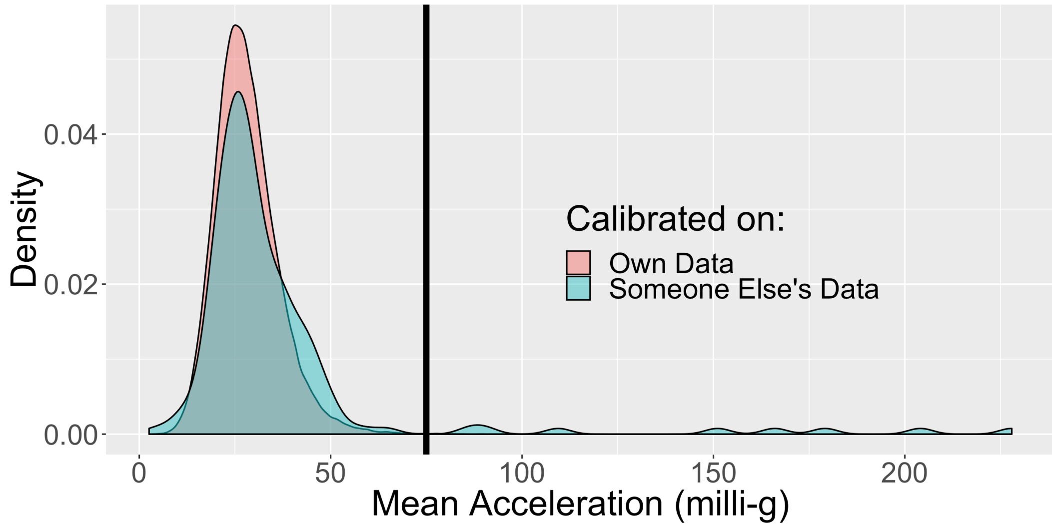


UKBB Processing: Auto-calibration

(Hees et al. 2014)

1. Use 10s window all axes SD < 13.0 mg.
2. Fit a unit gravity sphere using OLS.
3. If 3 axes had values outside a ±300mg range - use calibration coefficient
4. If not, use **next person's** calibration coefficient from the same device

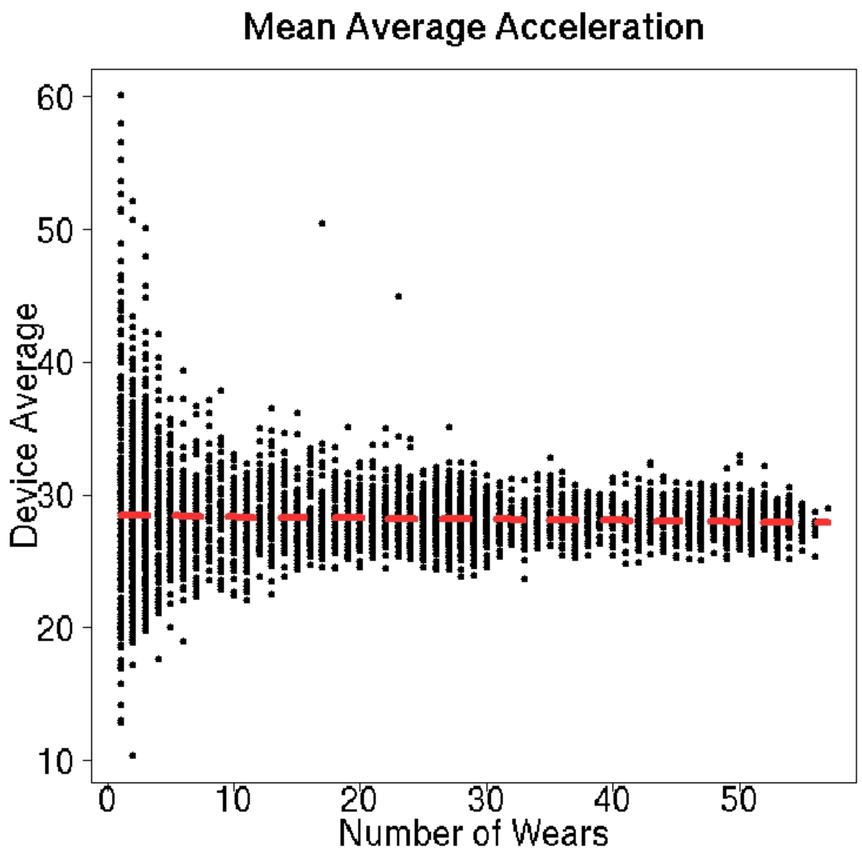




Processing math: 58%

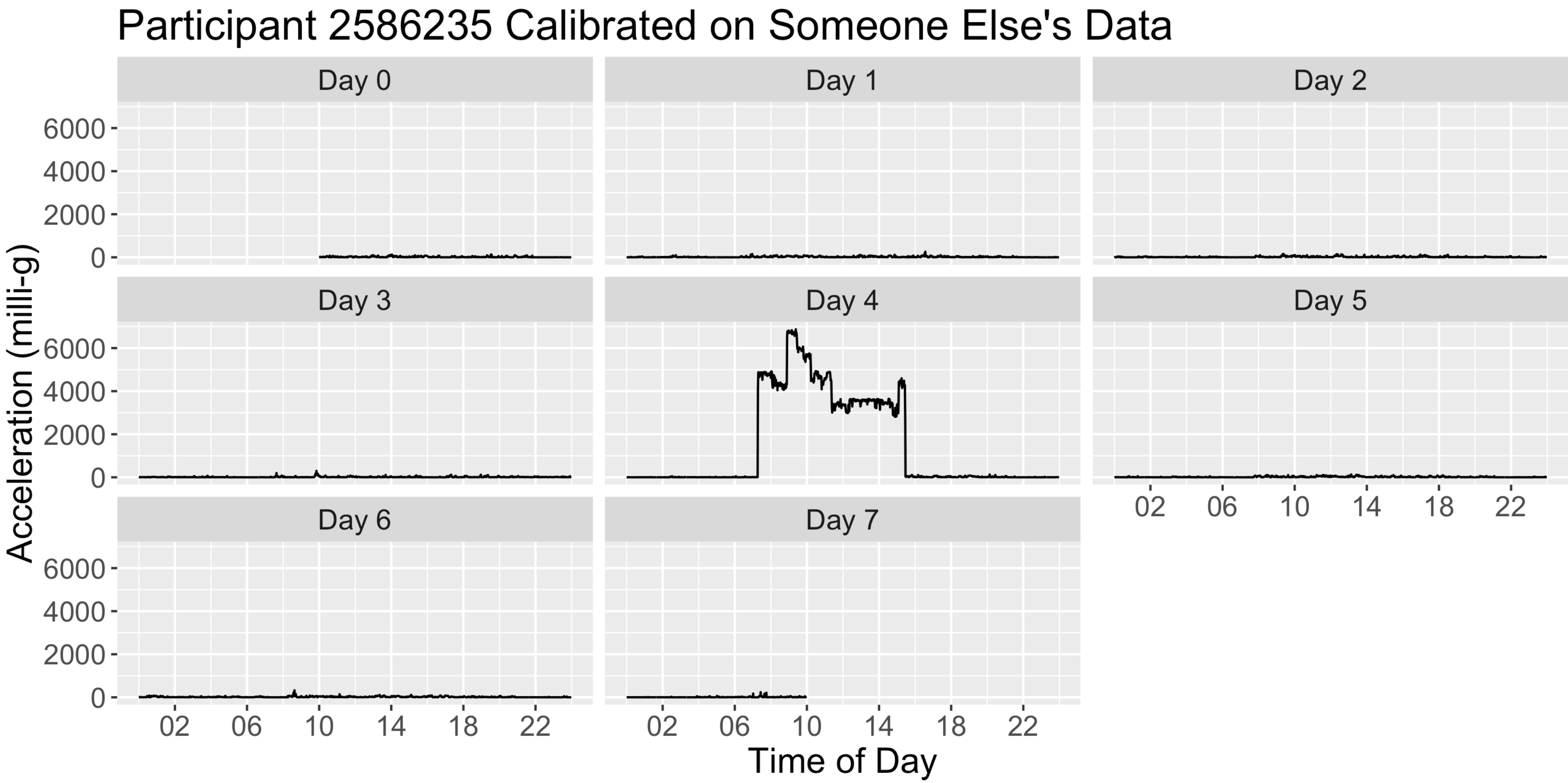
Auto-calibration seems to “work”

- Within-person average, within-device average (one point per device)
- Plotted against # of wears per device(σ should decrease with \sqrt{n})

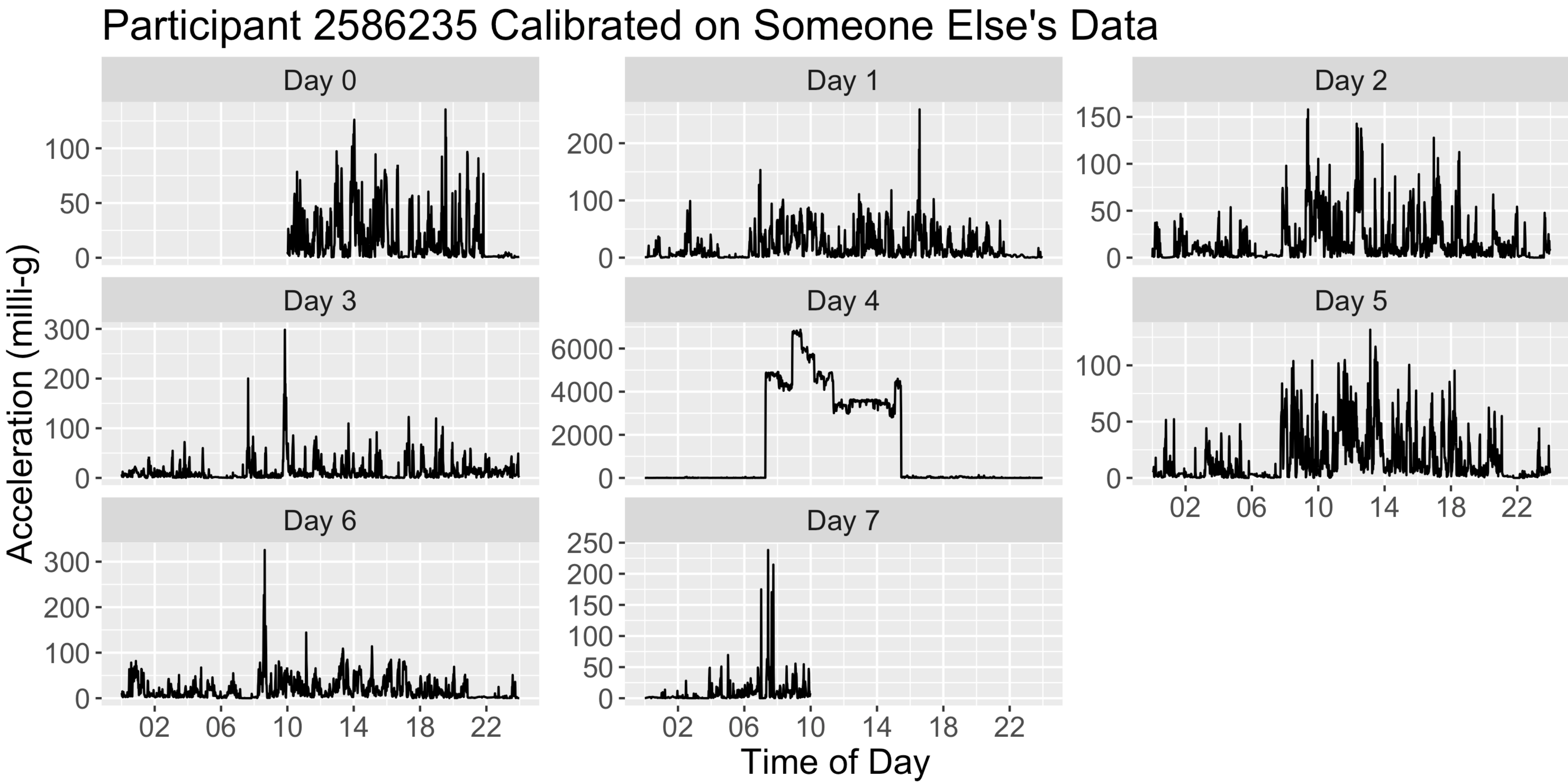


Processing math: 58%

Doesn’t work for all cases

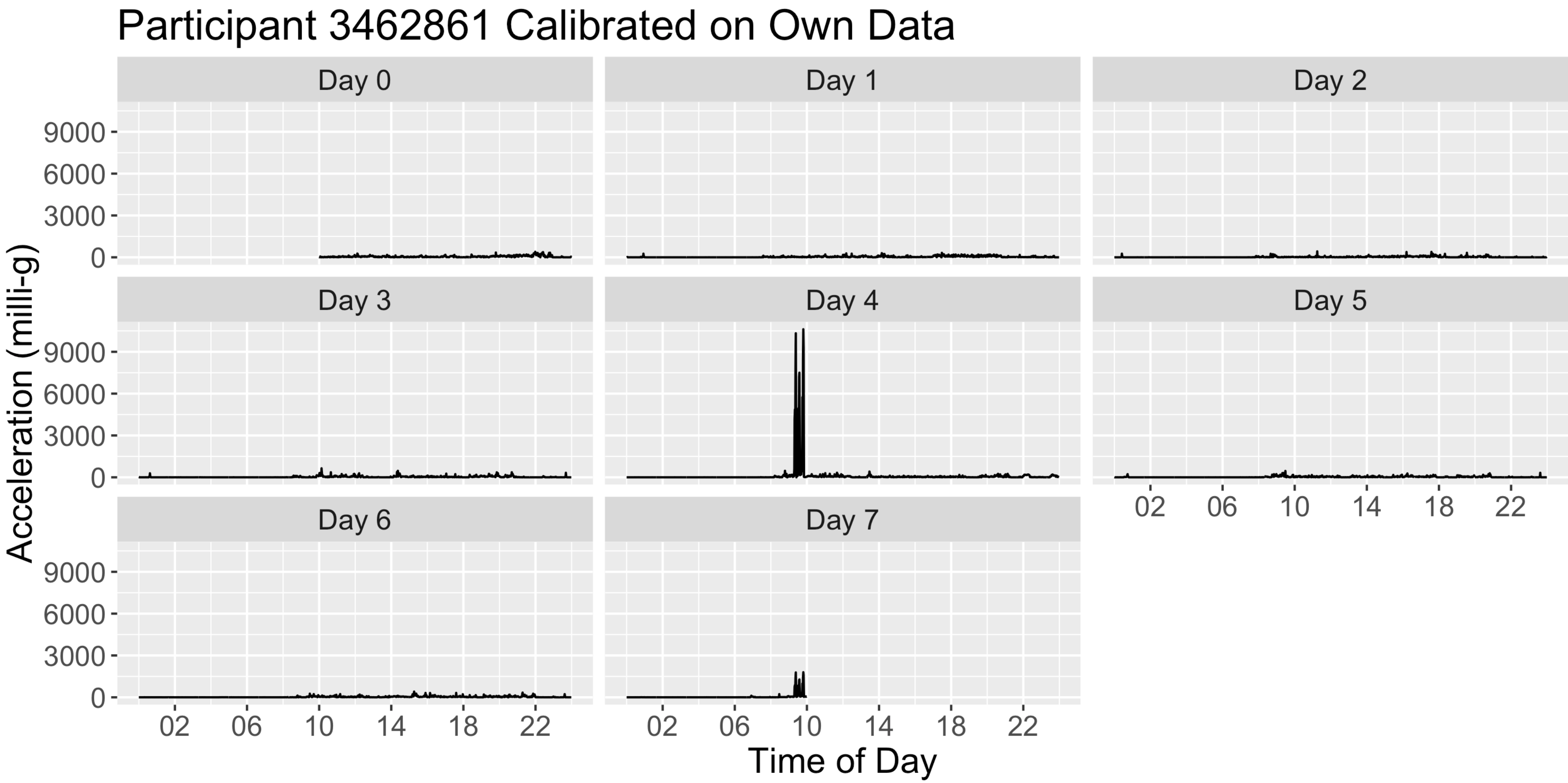


Doesn’t work for all cases



Processing math: 58%

Doesn’t work for all cases



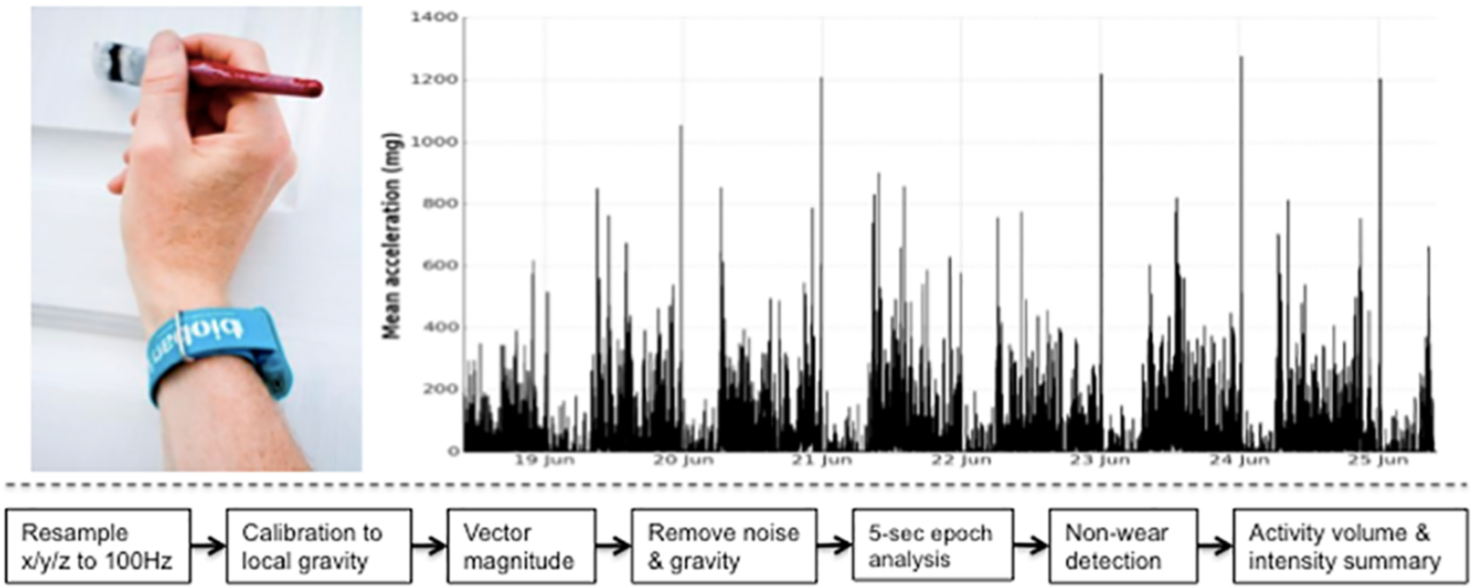
Processing math: 58%

Lesson #2: If magnitude is important, need calibration (“batch effect” correction), but may not be perfect

Processing math: 58%

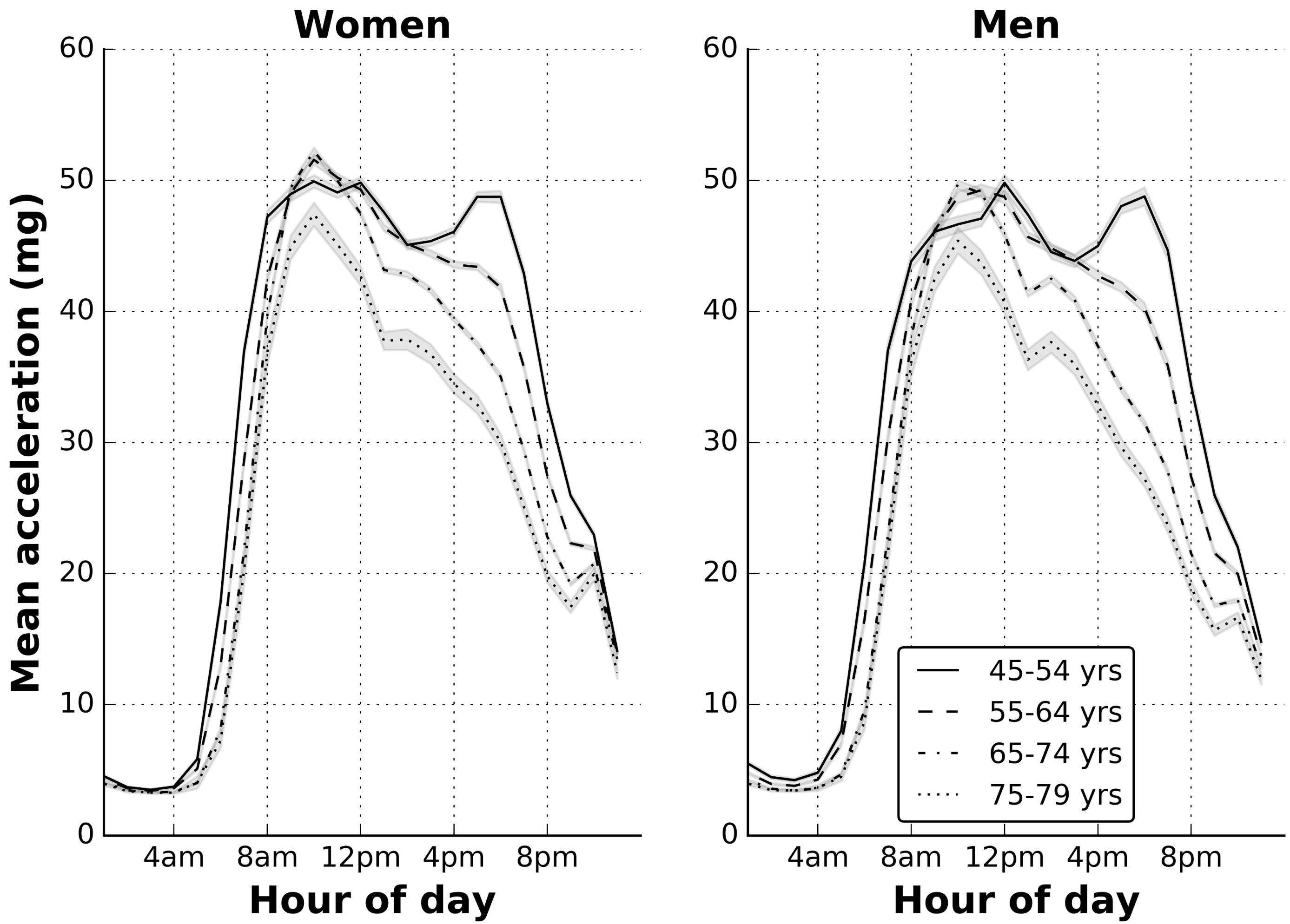
UKBB Processing: Doherty et al. (2017)

- Recording errors and ‘interrupts’ flagged (plug in accelerometer to computer)
- \pm8g flagged
- Resampled to 100 Hz (interrupts > 5 seconds set to missing)
- Euclidean norm, fourth order Butterworth low pass filter (f = 20Hz).
- Subtract 1g, negative values set to 0



They have software (Python) on GitHub

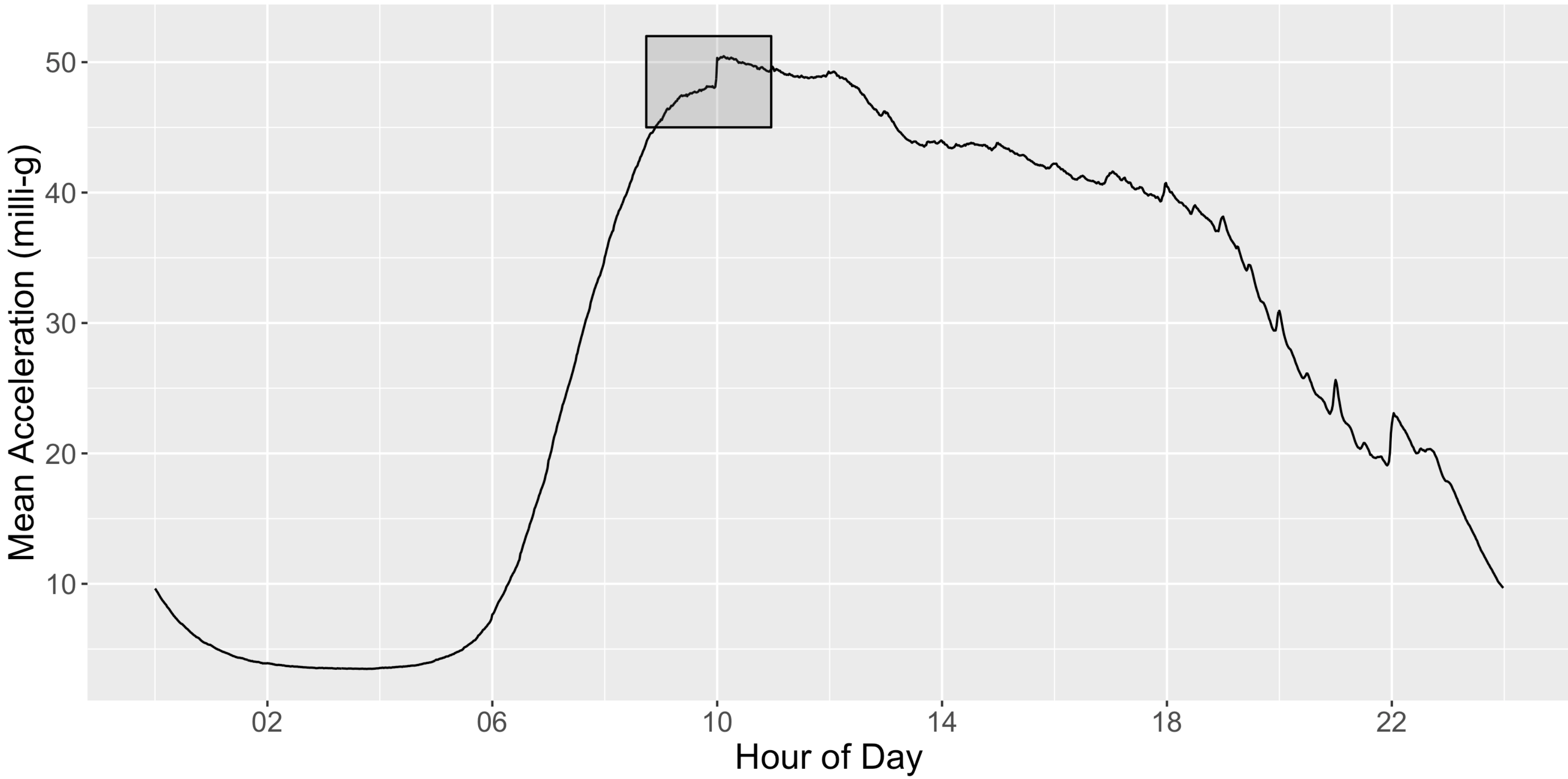
One result from (Doherty et al. 2017) analysis



Processing math: 58%

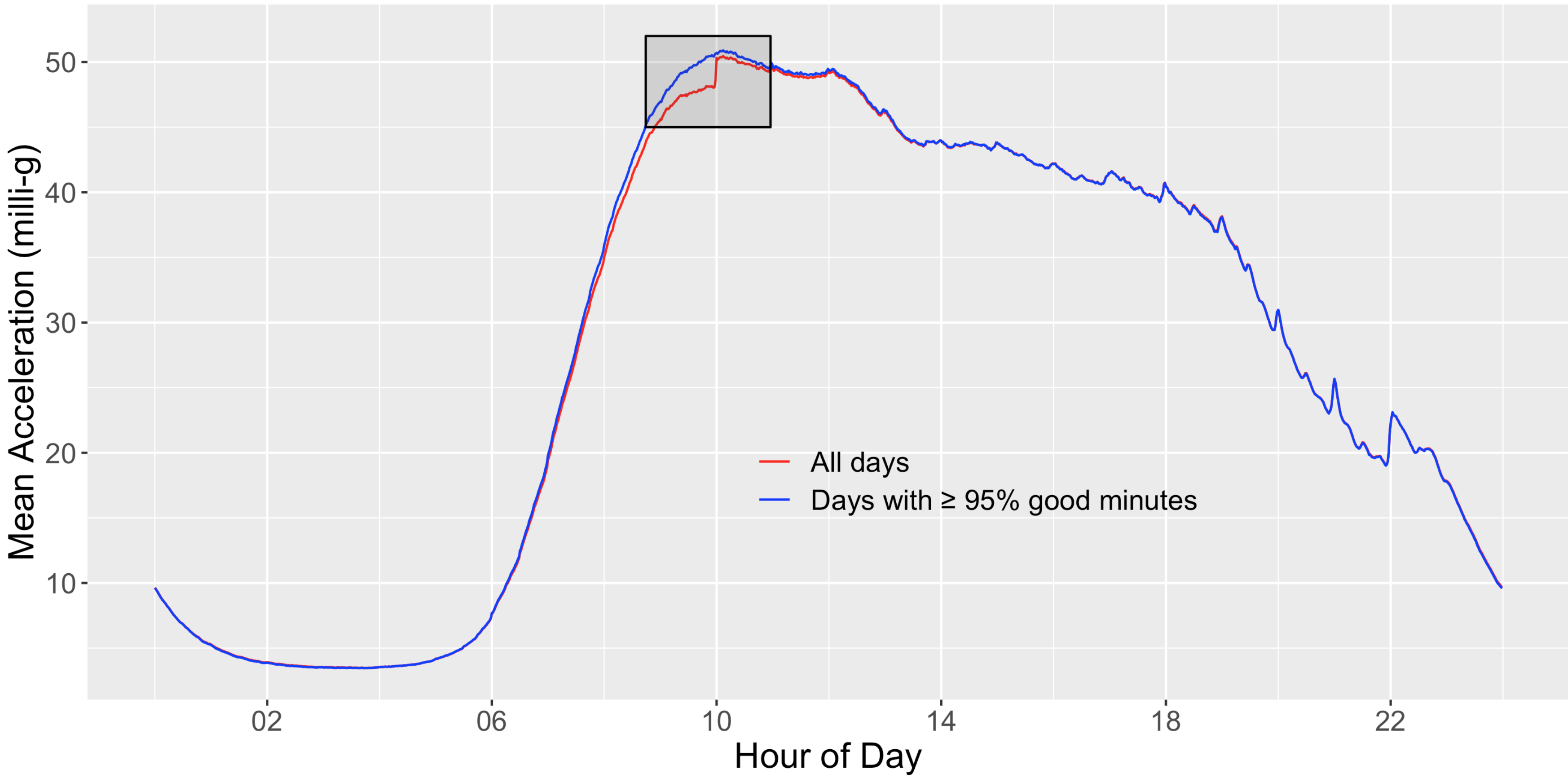
How do people typically move?

- Average over minute - regardless of mulitple visits per person



How do people typically move?

- Average over minute with days > 95% non-missing data

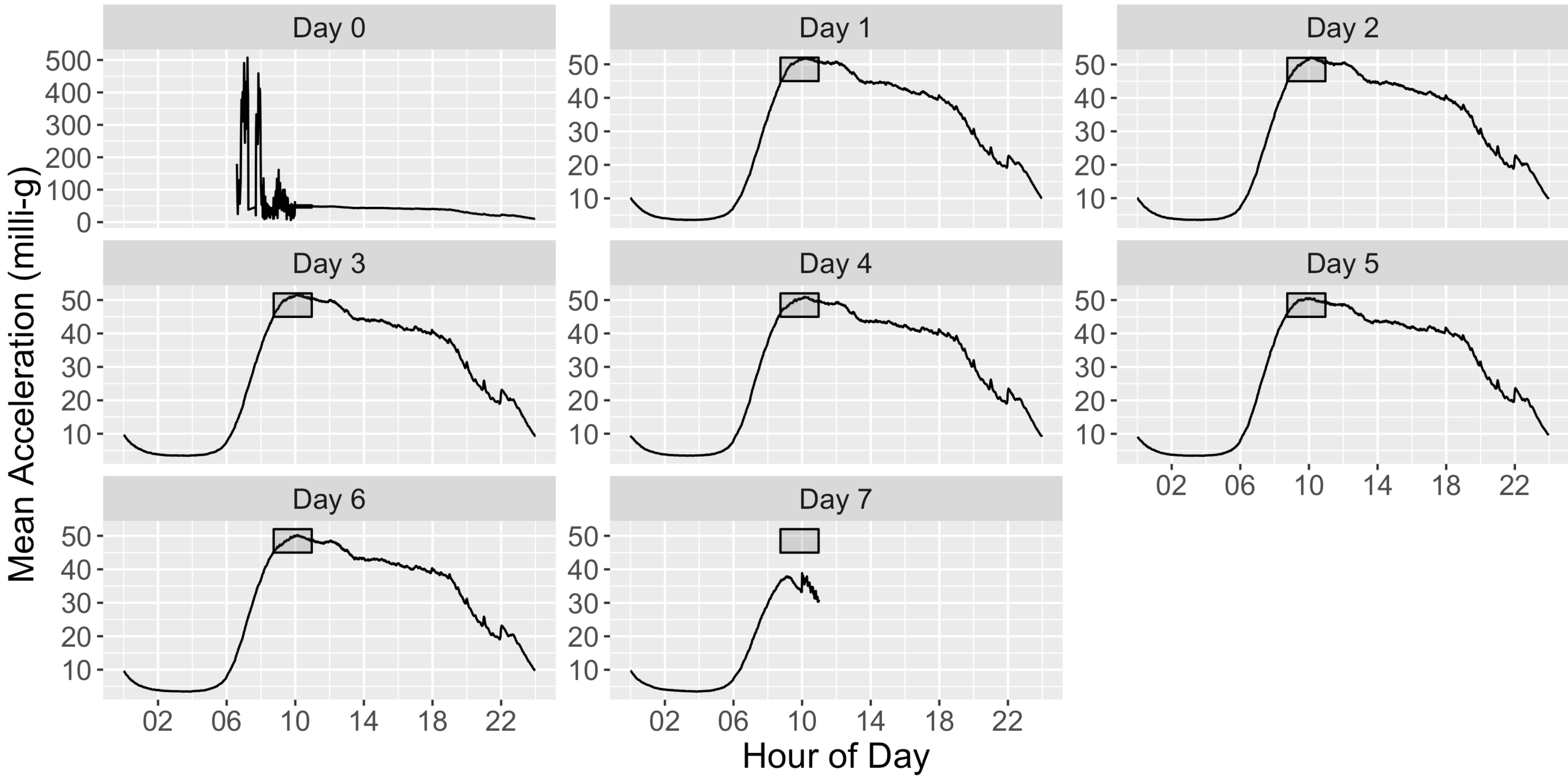


(Maybe) Lesson #3: Keep only “full” days

Processing math: 58%

Not so fast, let's look day by day

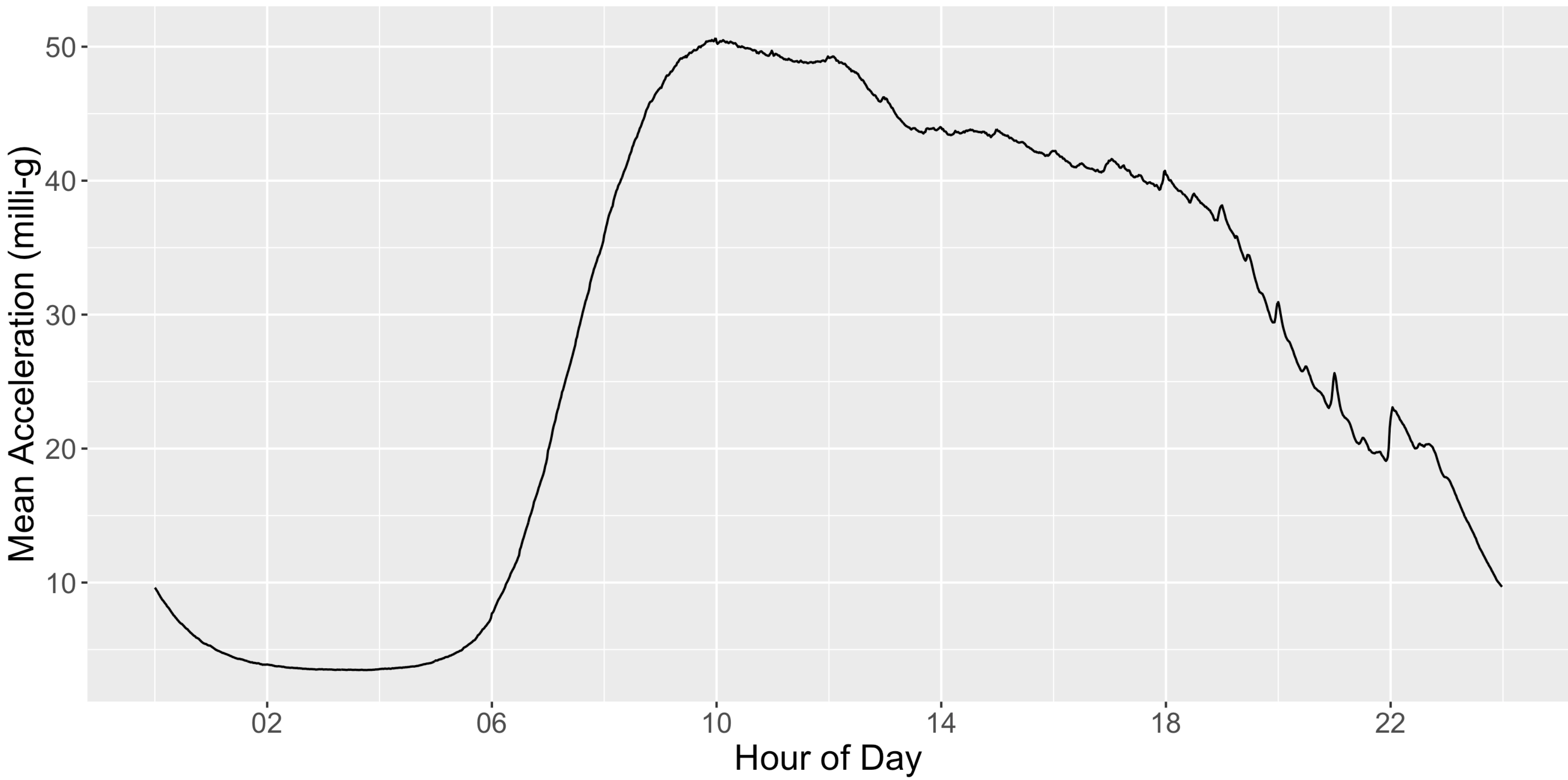
- Average over minute for days separately, one row per subject



Processing math: 58%

Removing Day 7

- Average over minute for days 0 - 6

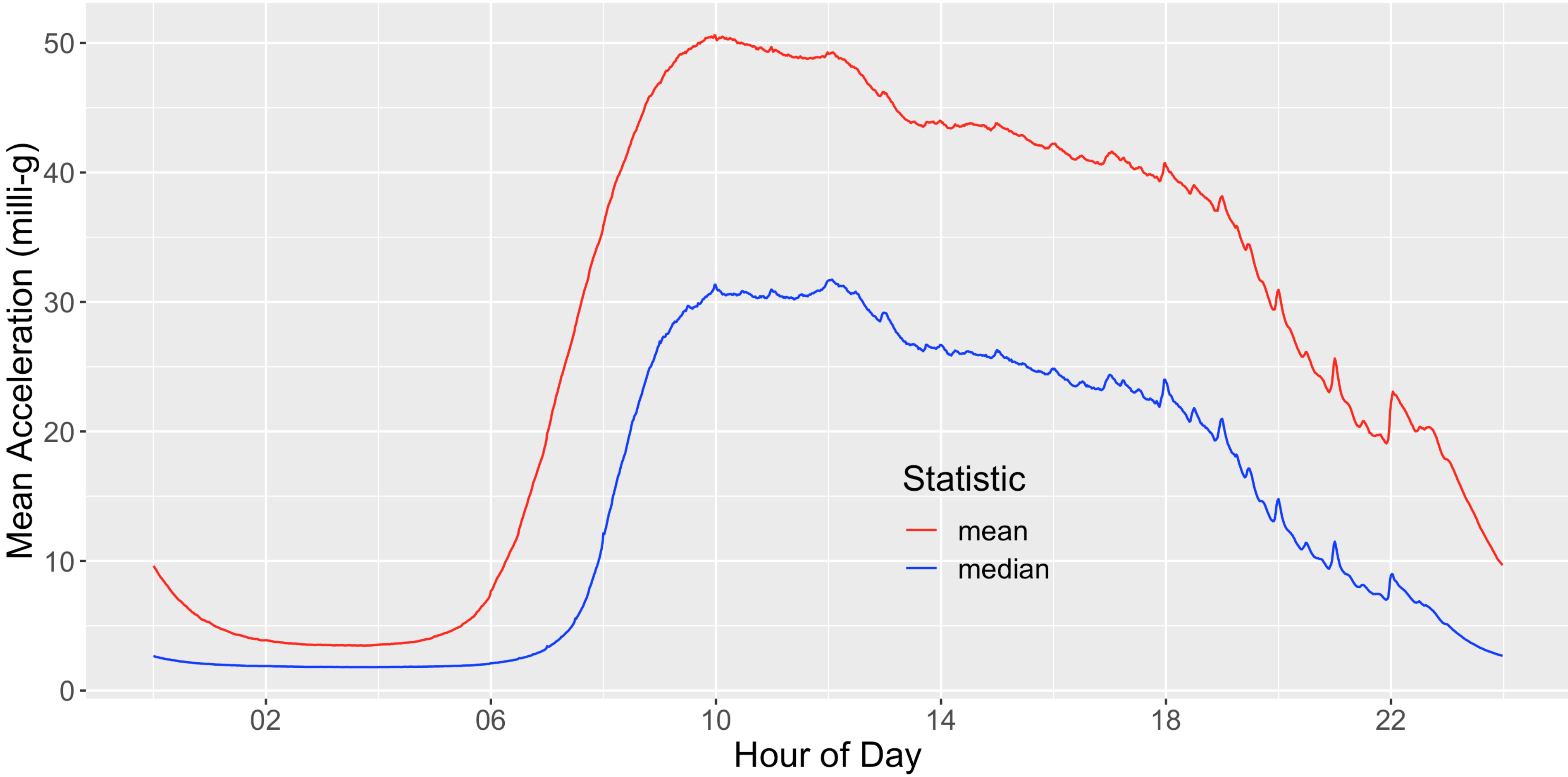


Processing math: 58%

Lesson #3: Explore days before averaging across in individual?

Processing math: 58%

Well use the MEDIAN then!

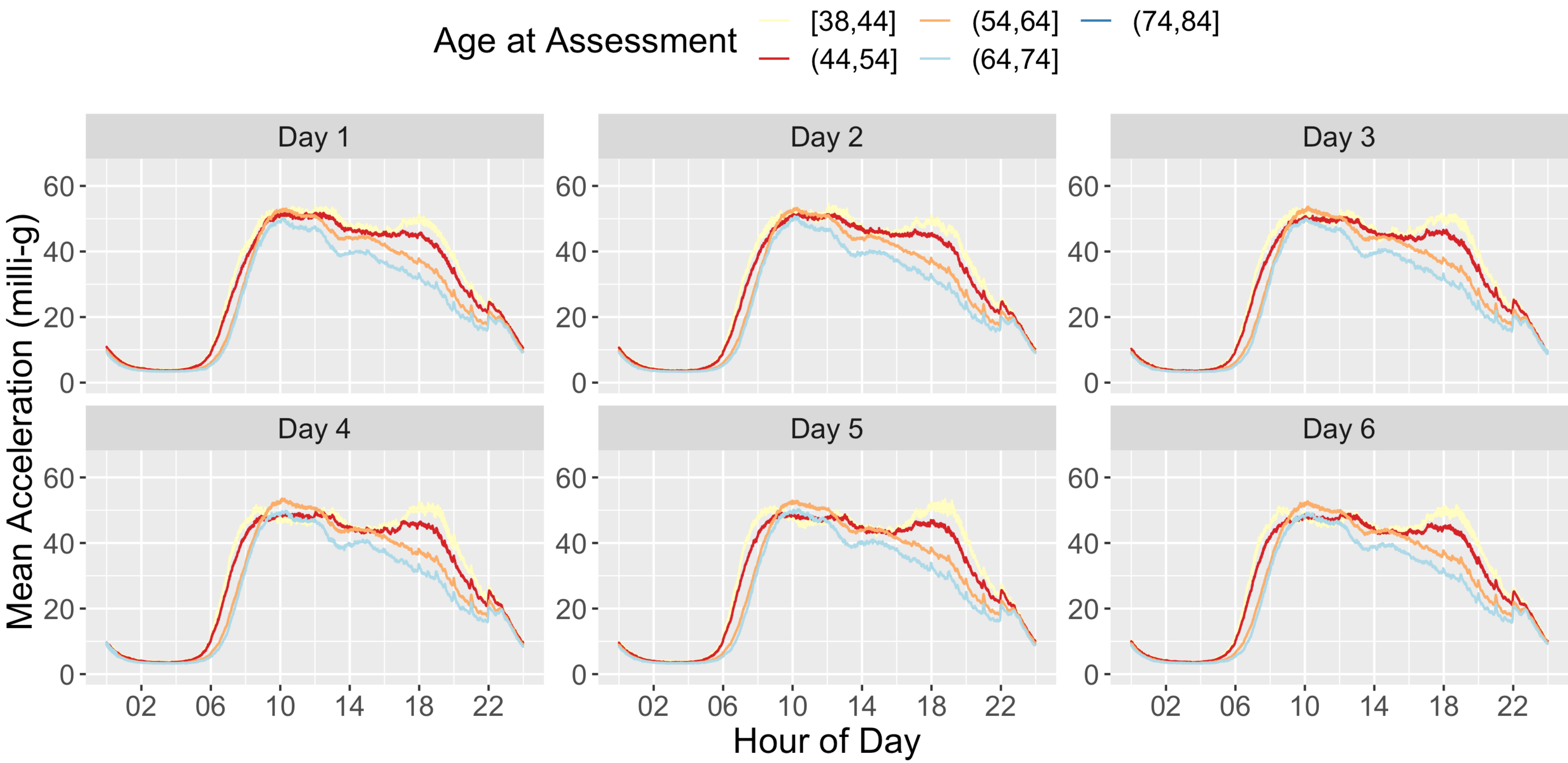


Processing math: 58%

Maybe it's a "few bad apples"

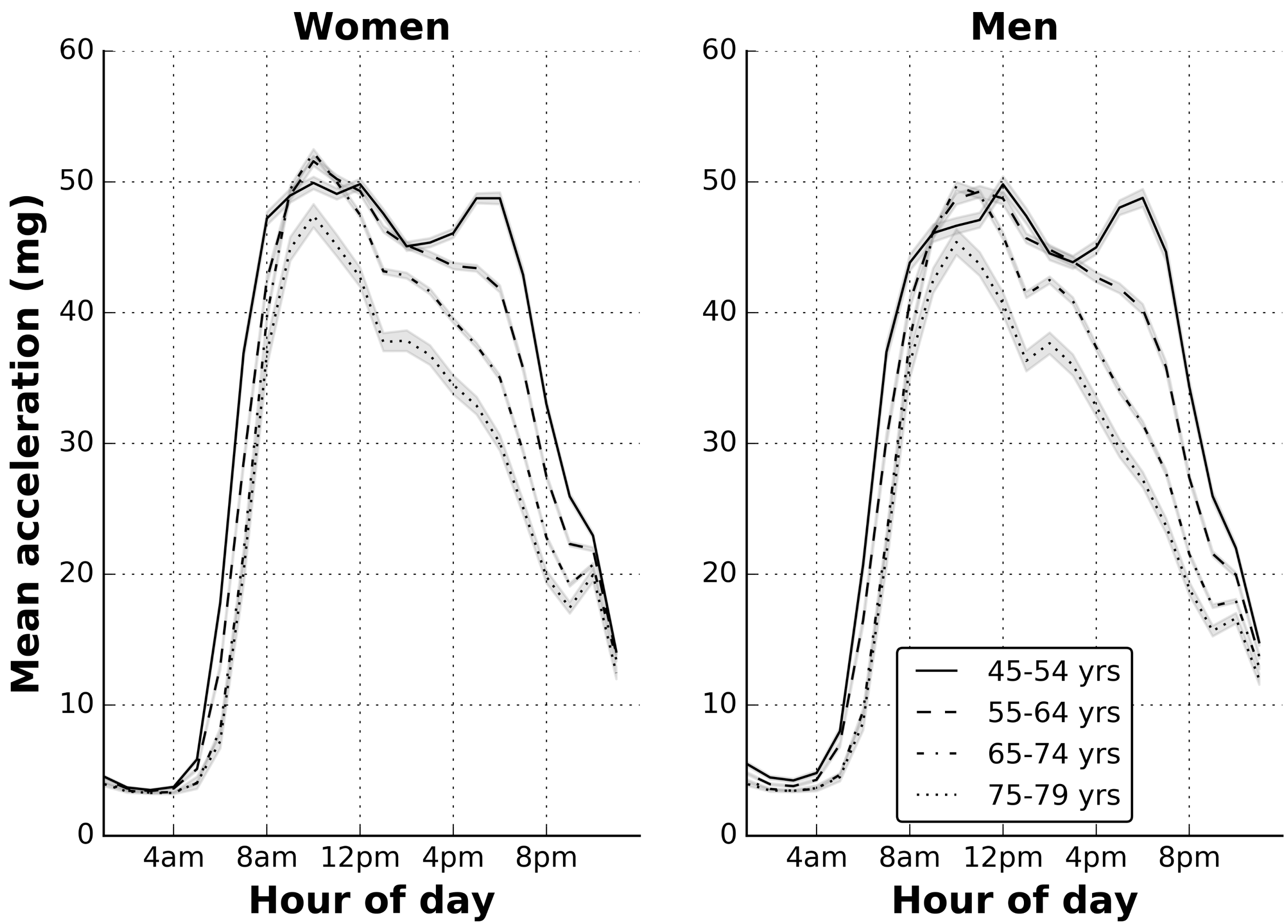
Heatmap of 2000 randomly sampled people

Age at Assessment Plot



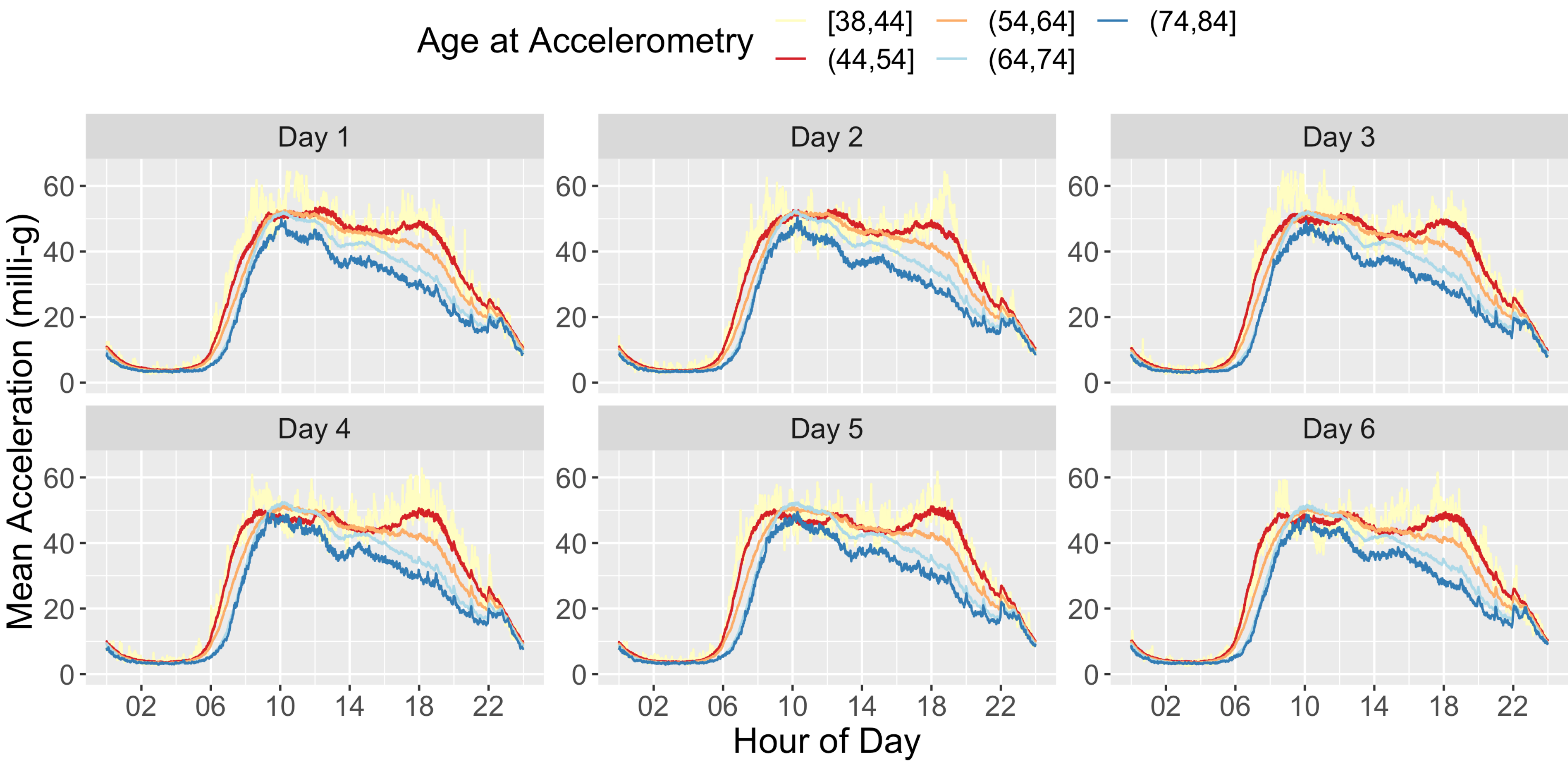
One result from (Doherty et al. 2017) analysis

- Look at age range

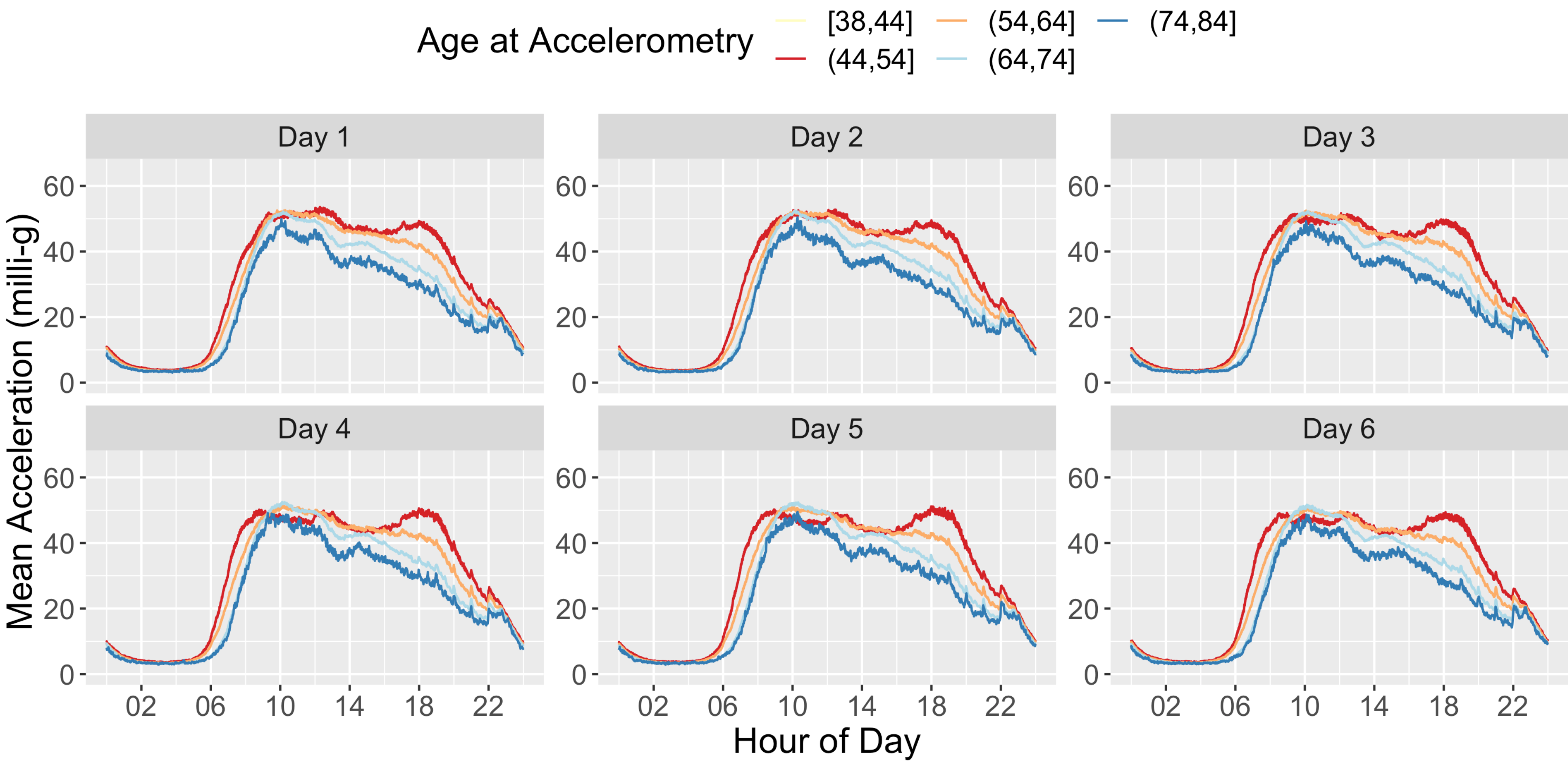


Processing math: 58%

Age at Acceleration Plot

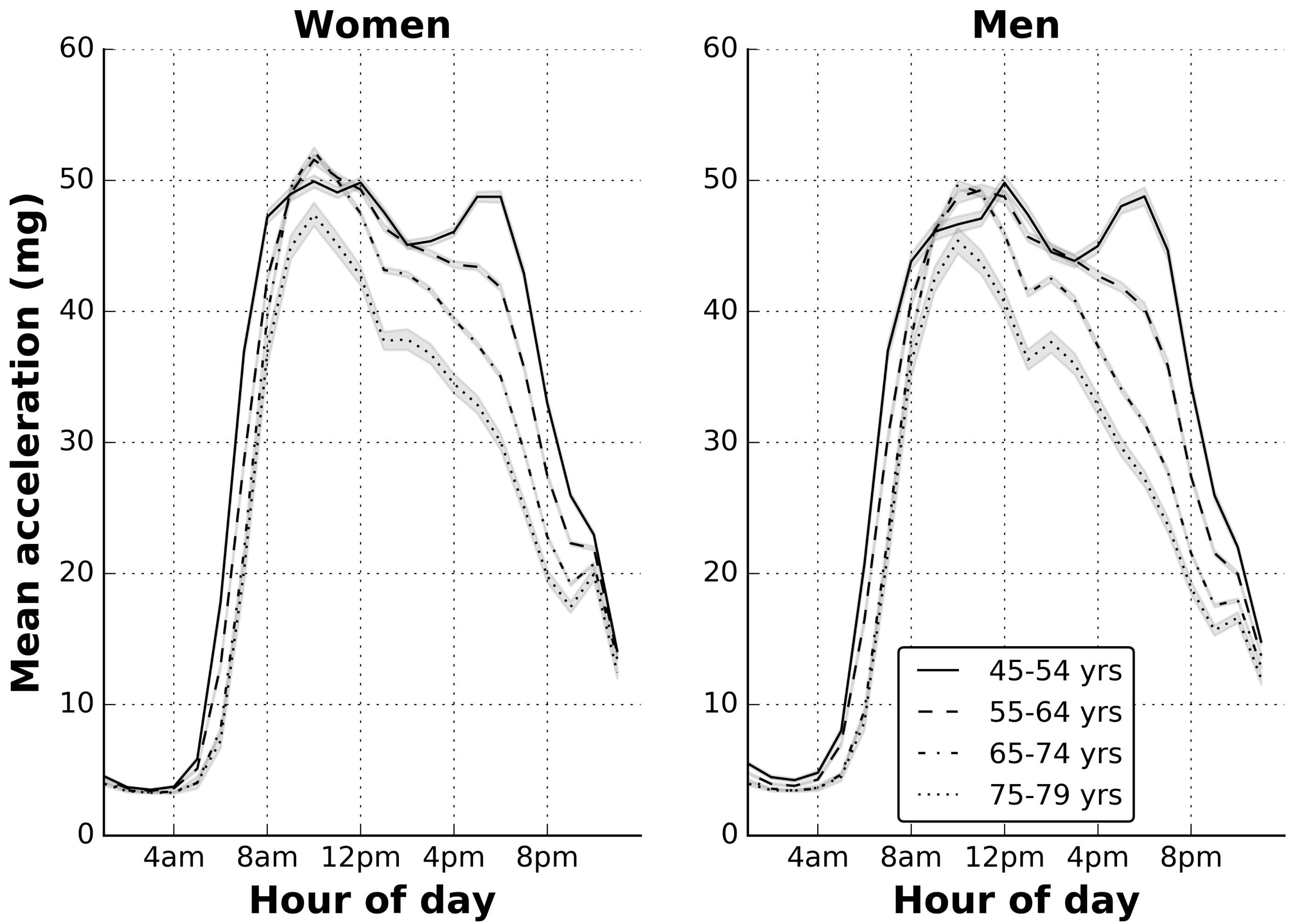


Age at Acceleration Plot - Remove 38-44 y/0



Processing math: 58%

One result from (Doherty et al. 2017) analysis



Processing math: 58%

Takehome Messages

1. Start off smaller than 100K people
2. Inspect the raw(ish) data
3. Processing highly affects results

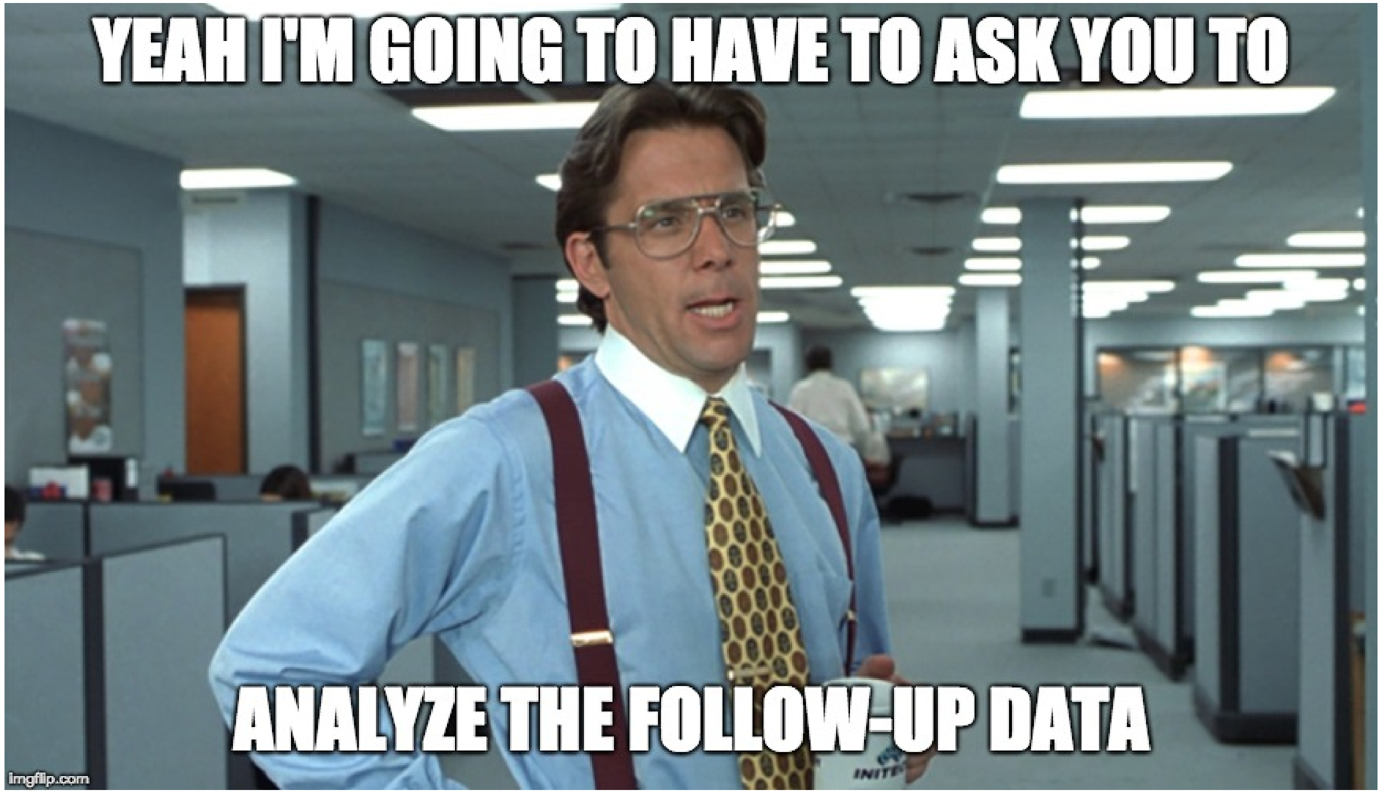
- Autocalibration seems to work well on gross features (with 100K people)
  - Artifacts still seem present in the data

1. Inclusion criteria matters (esp. for inference)
2. **Back to the 100Hz data** we go!

Processing math: 58%

Next data installment

“We invited some participants to wear an activity monitor for a week, four times a year. … finished in early 2019.”



Processing math: 58%

References (and Thanks)

Doherty, Aiden, Dan Jackson, Nils Hammerla, Thomas Plötz, Patrick Olivier, Malcolm H Granat, Tom White, et al. 2017. “Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The Uk Biobank Study.” 12 (2): e0169649.

Hees, Vincent T van, Zhou Fang, Joss Langford, Felix Assah, Anwar Mohammad, Inacio CM da Silva, Michael I Trenell, Tom White, Nicholas J Wareham, and Søren Brage. 2014. “Autocalibration of Accelerometer Data for Free-Living Physical Activity Assessment Using Local Gravity and Temperature: An Evaluation on Four Continents.” 117 (7): 738–44.