Some Practical Tools and Lessons Learned from Working with Wearable Sensor Data

Jaejoon Song

University of Texas
School of Public Health & MD Anderson Cancer Center
Department of Biostatistics
jjsong2@mdanderson.org

March 16, 2017

Objectives

1. To introduce opportunities from wearable sensor data
2. To introduce some caveats in wearable sensor monitoring
3. To introduce some tools for working with wearable sensor data

Opportunities from Wearable Sensor Data

Opportunities from wearable sensor data

- Recently, wearable devices have received increased attention
- Wearable devices available today can track
  - Physical activity and sleep patterns
  - Blood pressure, heart rate
  - Blood sugar, insulin
  - Electroencephalography (EEG)

Why use wearable sensors?

- Can measure in real-living conditions
- Can measure for extended periods of time
- Can be less burdensome to participants
- Can be a fun factor and encourage adherence (e.g. Fitbit)

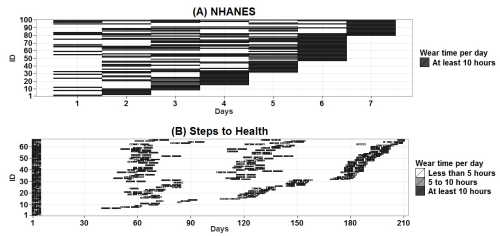
How widely used are wearable sensors?

- NHANES cycles 2003-04, 2005-06, 2011-12, and 2013-14
- UK Biobank (7-day activity data for 100,000+ individuals)
- Consumers (13.4 million sold in US, 2015)

What does the data look like?

- The raw data often exists in a form of high-frequency time-series
- Often, a particular type of processed data summary can be of interest to behavioral scientists or clinicians (e.g., minutes of 'moderate to vigorous' physical activity per day (MET ≥ 3))

Some Caveats in Wearable Sensor Monitoring



Some caveats in wearable sensor monitoring

1. Requires extensive data pre-processing
2. Requires special attention to missing data
3. Analysis of longitudinal physical activity data collected using these devices may necessitate specialized statistical methods

Data processing procedures for Activity Counts

1. Define non-human obs. and replace with missing indicators
2. Differentiate wear and non-wear time
3. Identify bouts of physical activity among wear time
4. Summarize the wear time and min. of physical activity per day

However, most of the processing procedures are ad-hoc, and following different guidelines will give you different summaries (see Troiano et al., 2014, British J. of Sports Med; Troiano et al., 2006, J. of App. Phys.).

Issues with wearing the wearables

- We rely on participants to wear the devices
- Missing data occur naturally:
  - Factors un-associated with physical activity
    1. Forgetting to wear
    2. Device failures or administrative errors
  - Factors associated with physical activity
    1. Physically active participants may be more adherent in wearing the monitoring devices
    2. Participants with certain body composition may be more likely to be physically active, and at the same time, be more adherent in wearing the monitoring devices

Caveats in analyzing accelerometer data

- Informative observation times: Association between device wear times and measurement outcome
- Censored observations: Study participants may stop wearing the device from a certain measurement day and onwards
- Informative censoring: The early termination of the wearable sensor monitoring may be related to the measurement outcome

Some Tools for Working with Wearable Sensor Data

Software

- Despite the increased attention to wearable sensor monitoring, availability of software has been limited
- We developed a free software for processing, visualizing, and analyzing accelerometer data
- This software is available for download at a public repository `Download`
- We also released a web-based application for exploring accelerometer data `Link`

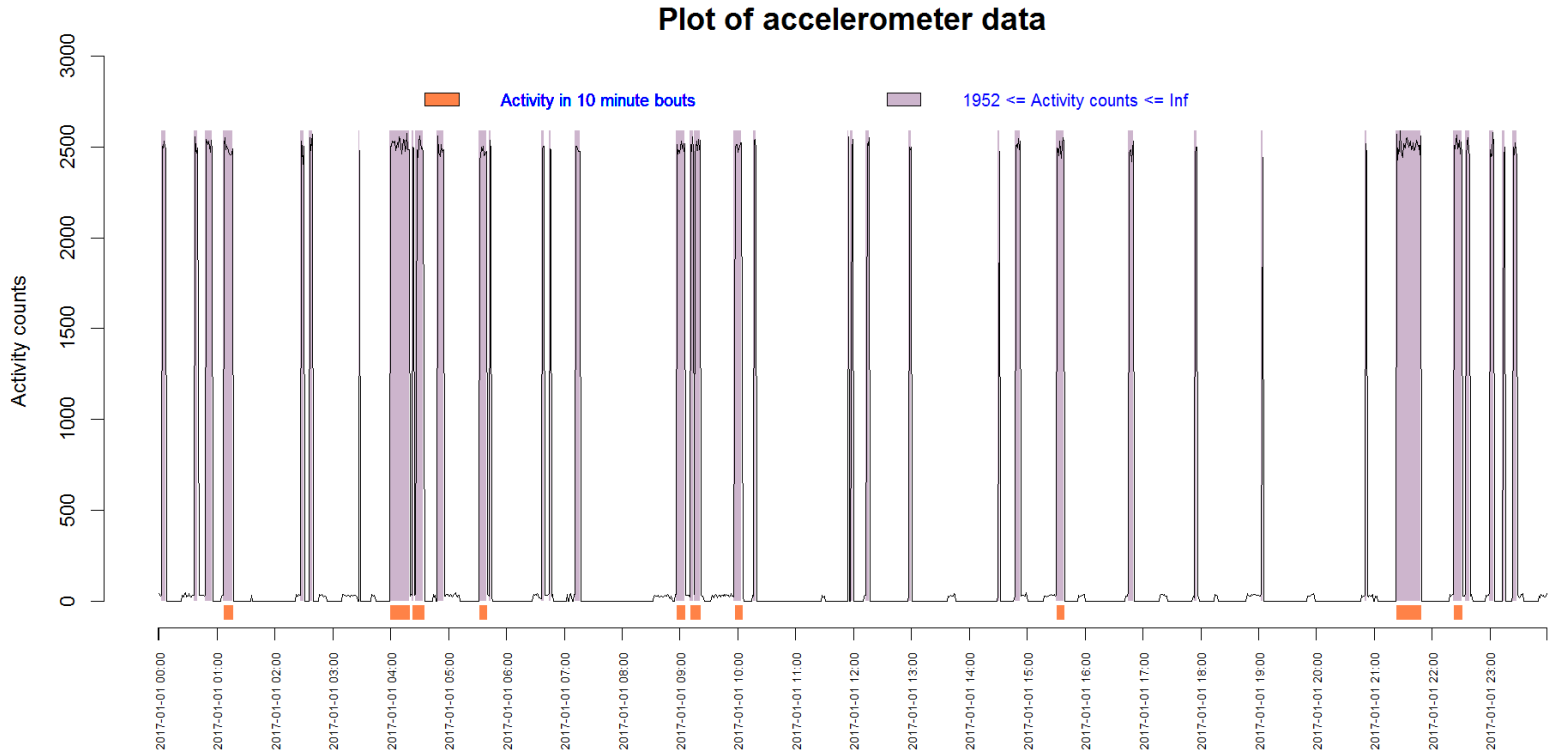
Software capabilities

1. Accelerometer data processing
   - Activity counts data
   - Raw accelerometer data
2. Accelerometer data visualization
3. Accelerometer data analysis
4. Accelerometer data simulation

Accelerometer data processing

The R package `acc` provides functions to process *activity counts* data based on the established rules in the literature.

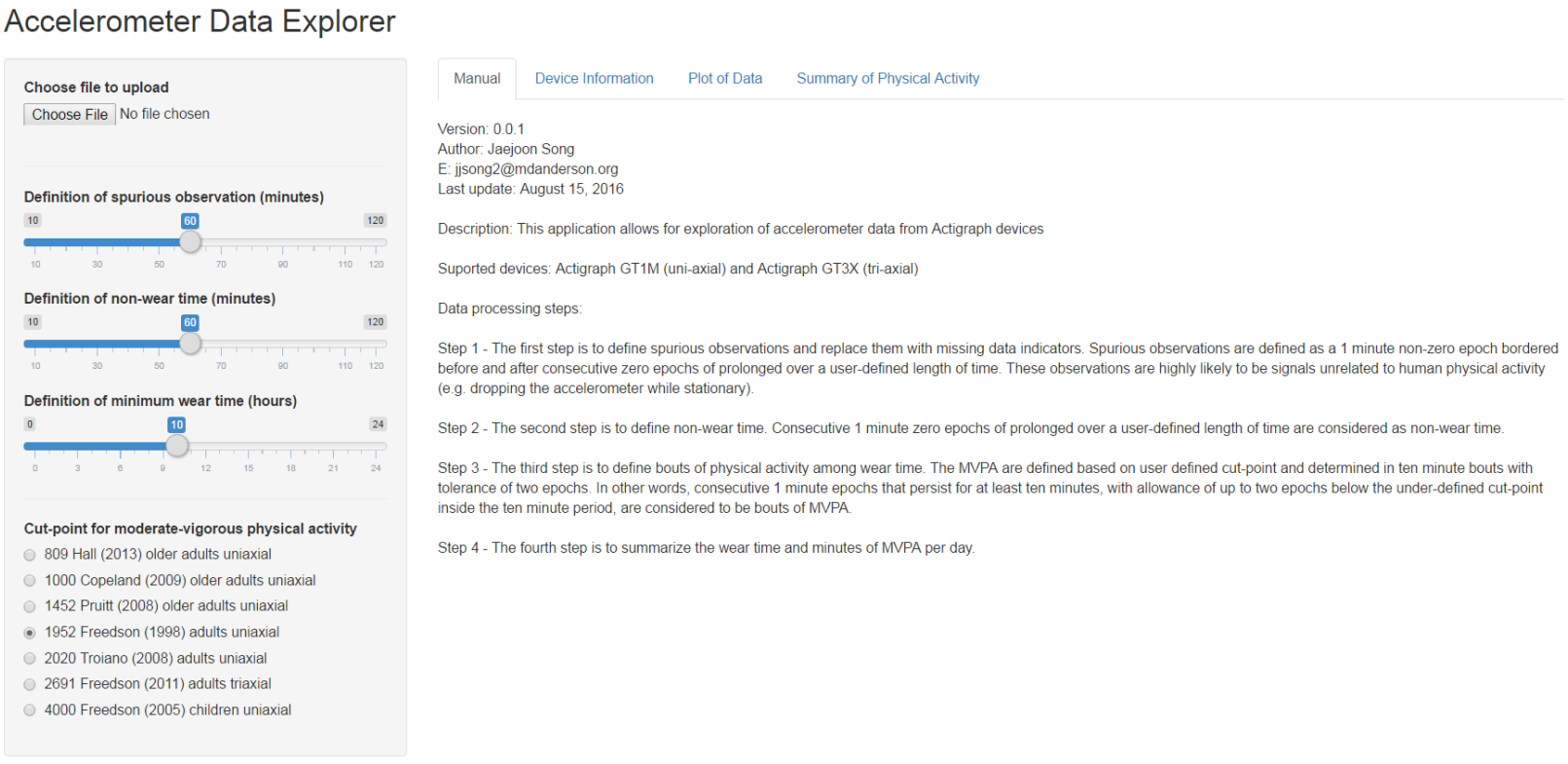
```
acc(data, tri, axis, spuriousDef, nonwearDef,
minWear, patype, pacut, boutsize, tolerance)
```



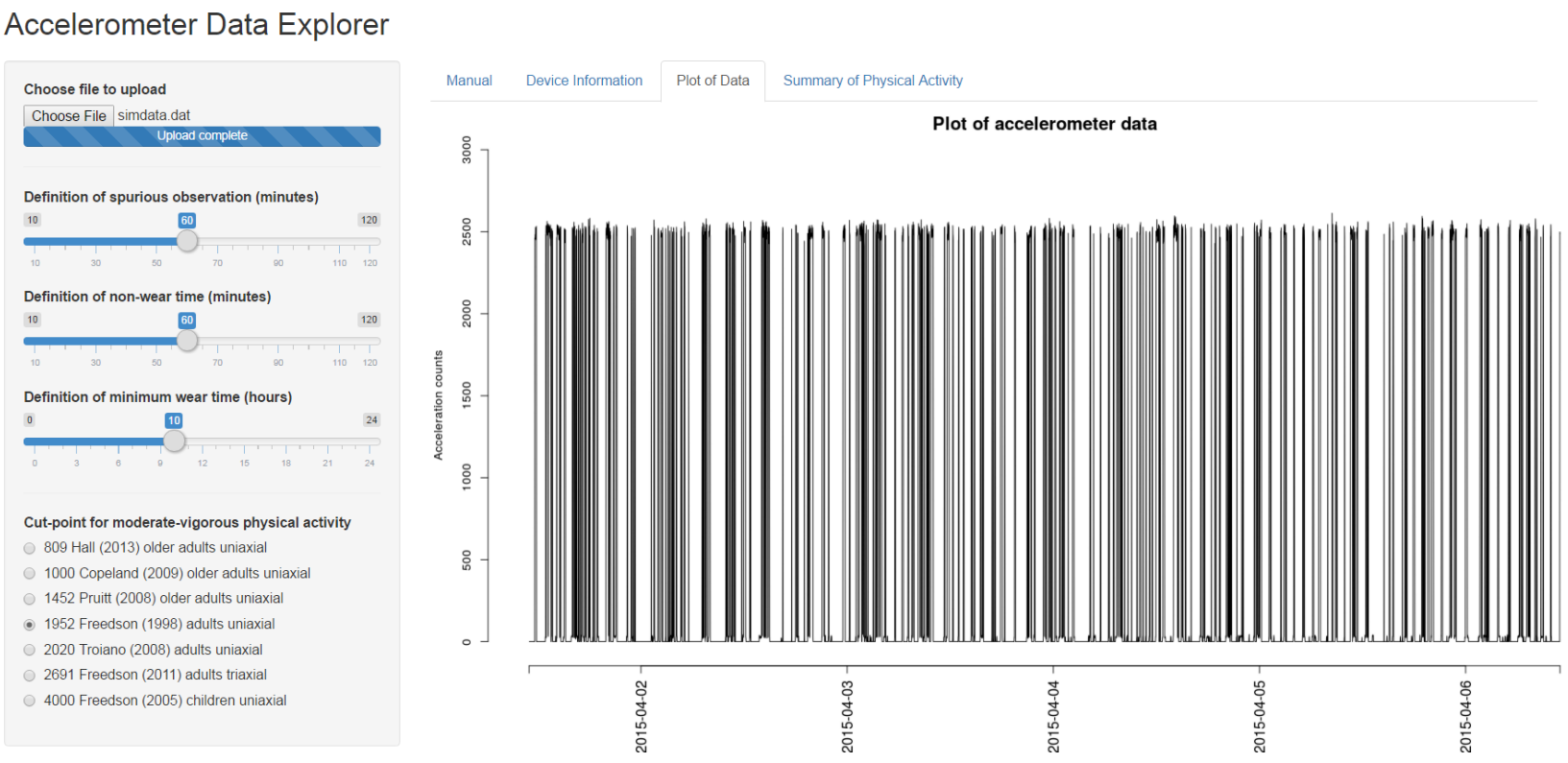
Accelerometer data processing

- Data processing can be also done in batch mode
- You can also use multiple cores
- A web-based application is designed for exploring the impact of different data processing specifications for accelerometer data Link

Accelerometer data processing: web application



Accelerometer data processing: web application



Accelerometer data processing: web application

Accelerometer Data Explorer

Choose file to upload

Choose File | simdata.dat

Upload complete

Definition of spurious observation (minutes)

10 60 120

10 30 50 70 90 110 120

Definition of non-wear time (minutes)

10 60 120

10 30 50 70 90 110 120

Definition of minimum wear time (hours)

0 10 24

0 3 6 9 12 15 18 21 24

Cut-point for moderate-vigorous physical activity

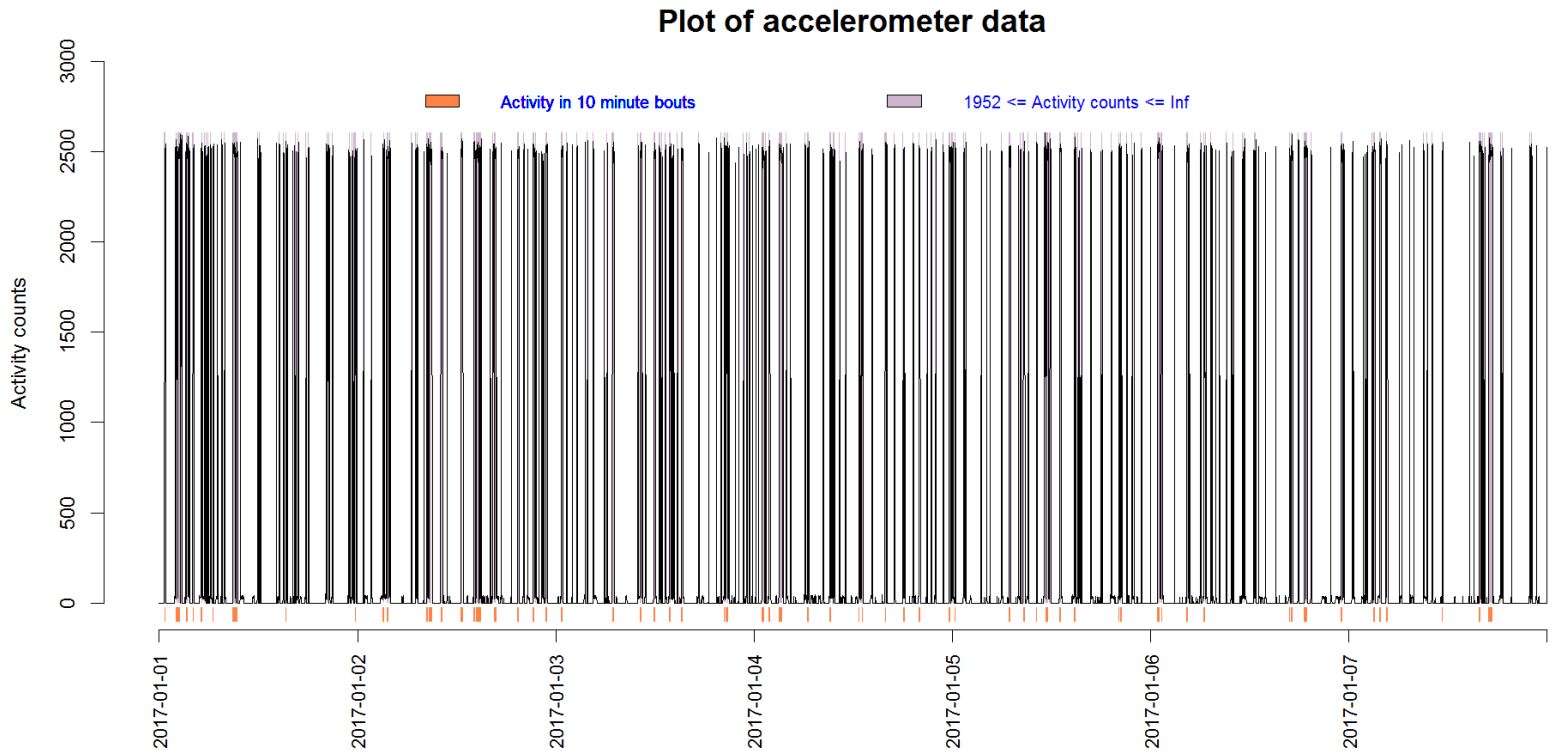
- 809 Hall (2013) older adults uniaxial
- 1000 Copeland (2009) older adults uniaxial
- 1452 Pruitt (2008) older adults uniaxial
- (•) 1952 Freedson (1998) adults uniaxial
- 2020 Troiano (2008) adults uniaxial
- 2691 Freedson (2011) adults triaxial
- 4000 Freedson (2005) children uniaxial

Manual | Device Information | Plot of Data | Summary of Physical Activity

	Date	Wear Time	Sedentary Minutes	Sedentary Bouts	MVPA Minutes	MVPA Bouts
1	2015-04-01	613.00	334	14.00	115	6.00
2	2015-04-02	1015.00	655	25.00	186	12.00
3	2015-04-03	1225.00	785	28.00	199	11.00
4	2015-04-04	1344.00	999	35.00	145	8.00
5	2015-04-05	1366.00	1057	26.00	132	10.00
6	2015-04-06	591.00	433	11.00	53	4.00

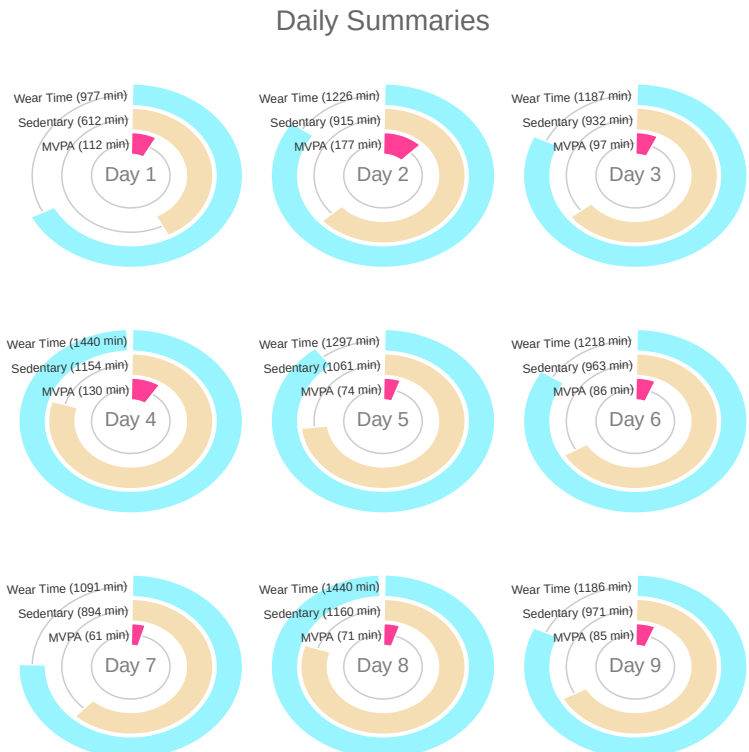
Accelerometer data visualization

Our software offers a variety of options for visualizing data



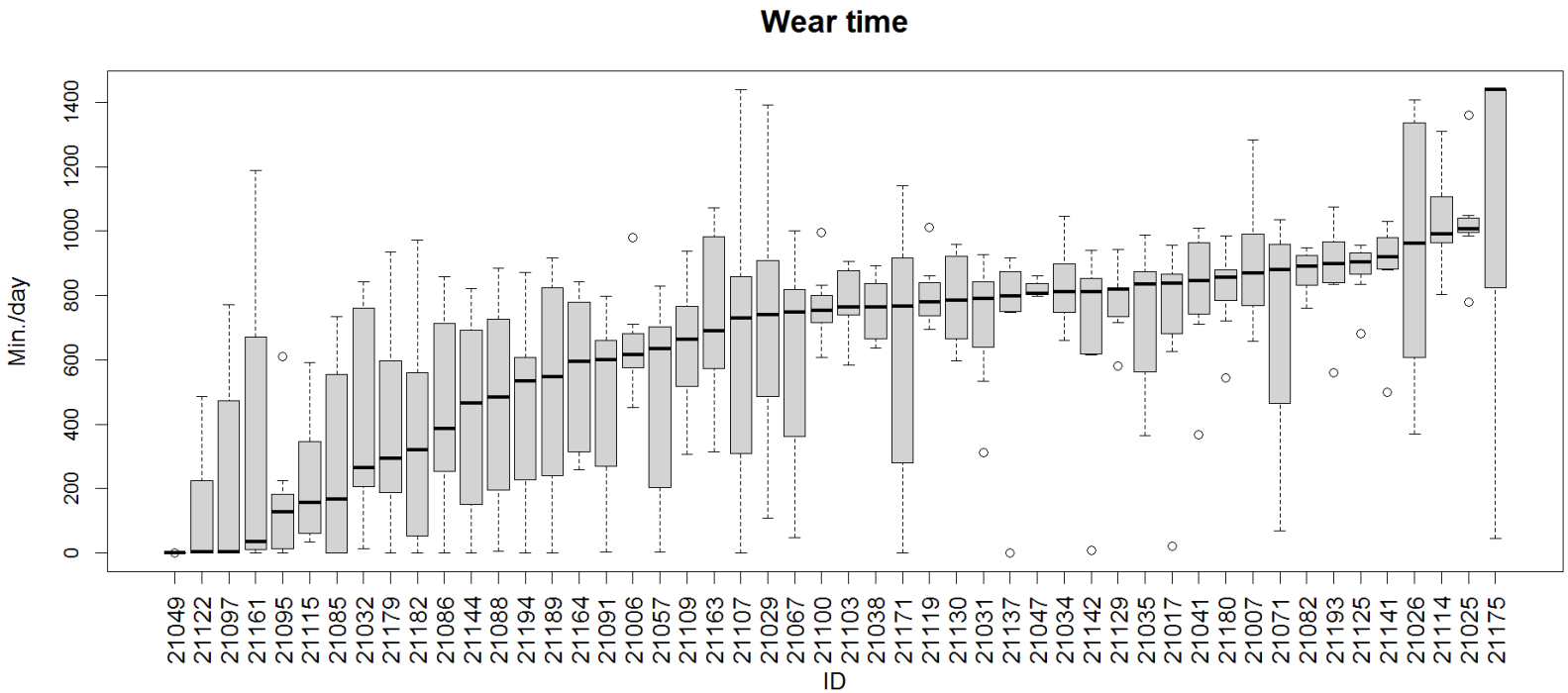
Accelerometer data visualization

Daily activity summaries can be represented compactly using race track plots



Accelerometer data visualization

Boxplots of wear times for multiple individuals



Accelerometer data visualization

Heatmaps are useful for exploring group level patterns

Analysis of longitudinal wearable sensor data

The function `aeexfit` can be used for analyzing wearable sensor physical activity data with missing observations.

```
aeexfit(formula, data, weight, se, control, boot)
```

Near future in wearable technology

- Wearable devices may become (nearly) invisible
- Will be more efficient, and powered without external charge
- Will be connected with Internet of Things
- May be integrated into Electronic Health Records
- May require more and more sophisticated methodology

Discussion

1. Introduced opportunities from wearable sensor data
2. Introduced some caveats in wearable sensor monitoring
3. Introduced some tools for working with wearable sensor data

Acknowledgements

The study was supported by the National Institutes of Health/National Cancer Institute through R01 CA 109919, R25T CA057730, R25E CA056452, P30 CA016672 (M. D. Anderson's Cancer Center Support Grant and PROSPR Shared Resource), and the Center for Energy Balance in Cancer Prevention and Survivorship, Duncan Family Institute for Cancer Prevention and Risk Assessment.

Thank you

- Questions?
- jjsong2@mdanderson.org