Wearable devices - statistical learning to the rescue

Jaroslaw Harezlak, Ph.D.

Professor and Interim Co-Chair
Department of Epidemiology and Biostatistics
School of Public Health
Indiana University, Bloomington, IN, USA

May 22, 2020
Zoom, Universe

Outline

Wearable and implantable devices

- Overview of the accelerometry data
- Micro-scale analysis
  - Detection and analysis of walking
  - Gait features
- Macro-scale analysis

Accelerometry data in health research: challenges and opportunities

Review and examples

Statistics in Biosciences, 1-28, 2019

Authors:

Marta Karas

Jiawei Bai

Marcin Strączkiewicz

Jaroslaw Harezlak

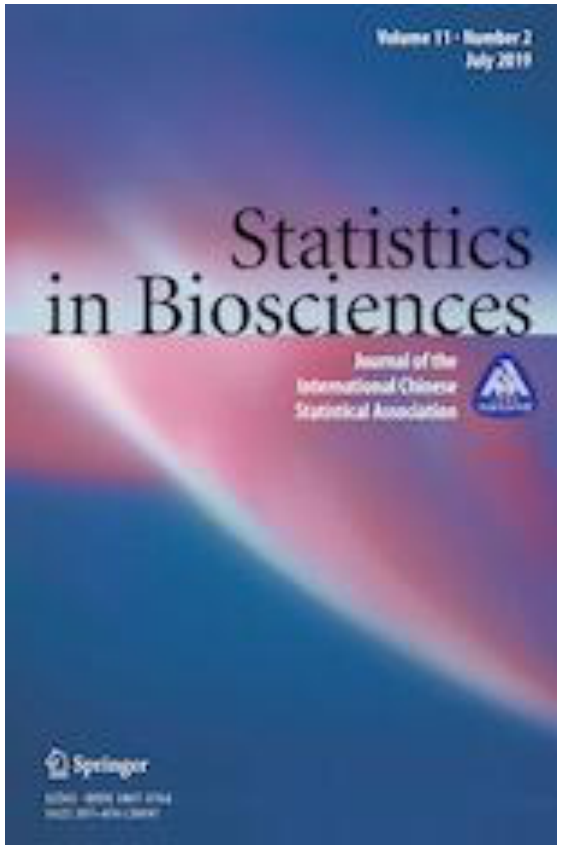
Nancy Glynn

Tamara Harris

Vadim Zipunnikov

Ciprian Crainiceanu

Jacek Urbanek



Wearable and Implantable Technology

- Wearable and implantable devices are smart electronic devices (electronic device with micro-controllers) that can be worn on the body as implants or accessories
- Emerging technology
- Growing popularity in health research

Actigraphy

- Actigraphy: non-invasive monitoring of activity
- Accelerometer records human movement





Other wearable/implantable devices

Major data challenges

- Data size
- Complexity and heterogeneity
- Lack of standardized data collection protocols
  - Device type
  - Device location
  - Wear/non-wear
  - No. of days of observations

Collaboration

Epidemiologists

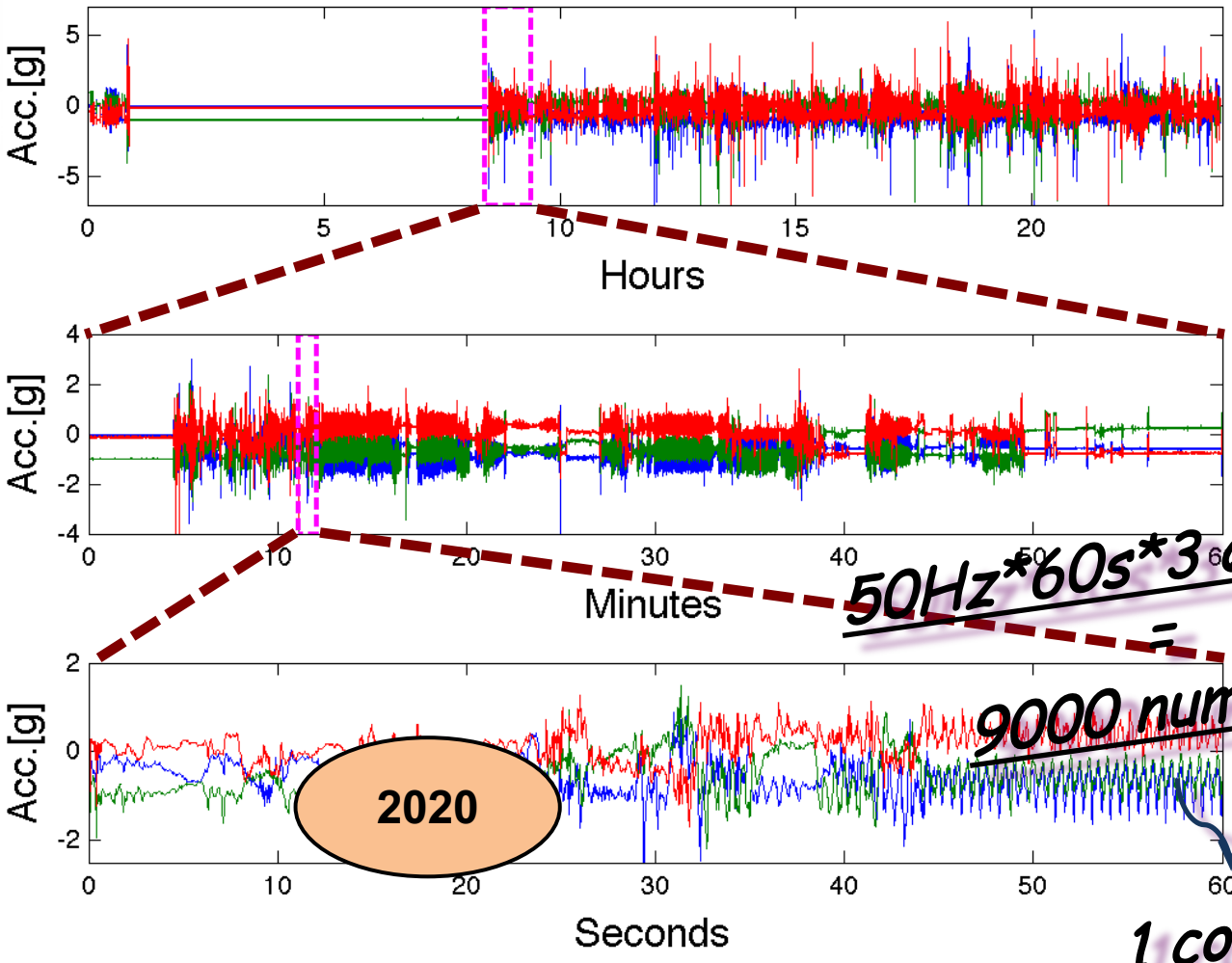
Biostatisticians and data scientists

Health Research

Clinicians

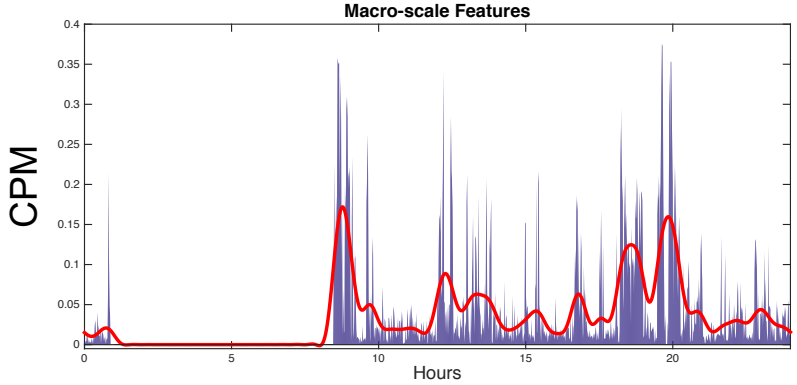
Software engineers

Raw accelerometry data

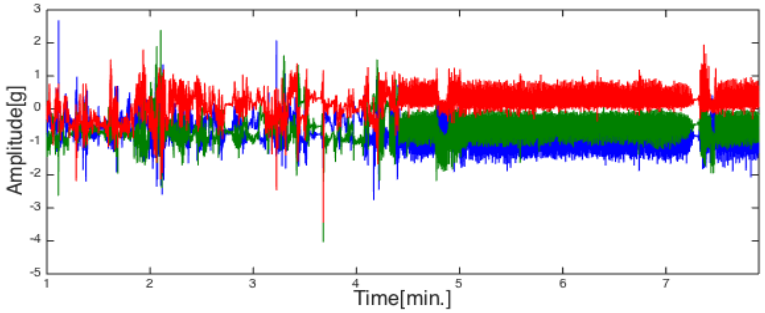


Macro- and Micro-scale

- **Macro-scale** – summarized data (1-minute intervals)

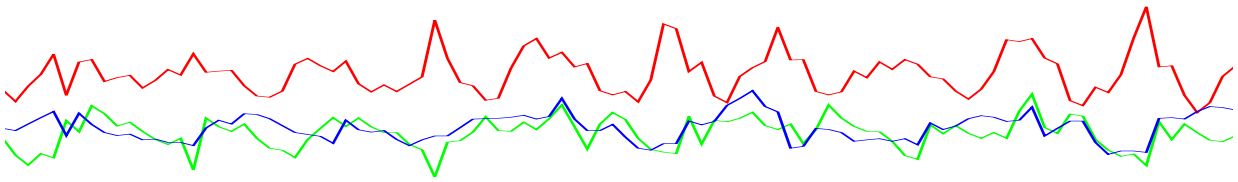


- **Micro-scale** – raw accelerometry data (10Hz+)



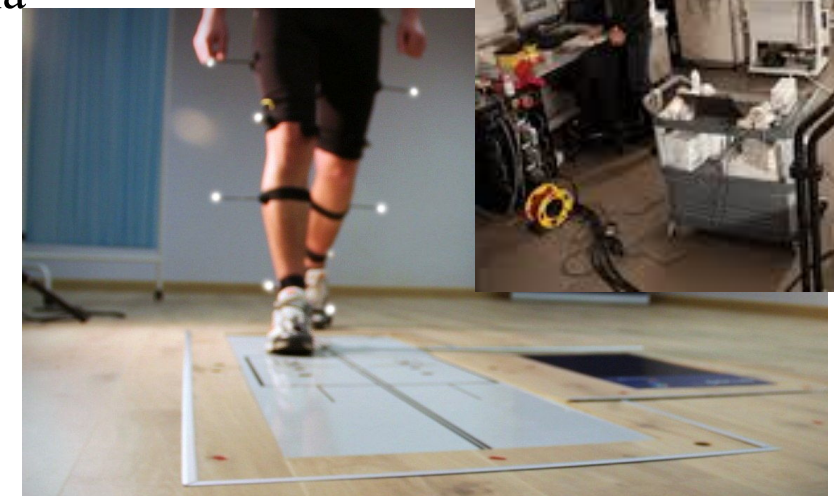
Micro-scale analysis

- Activity type recognition
  - **Detection of walking**, driving
  - Climbing stairs, resting, sedentary vs. upright



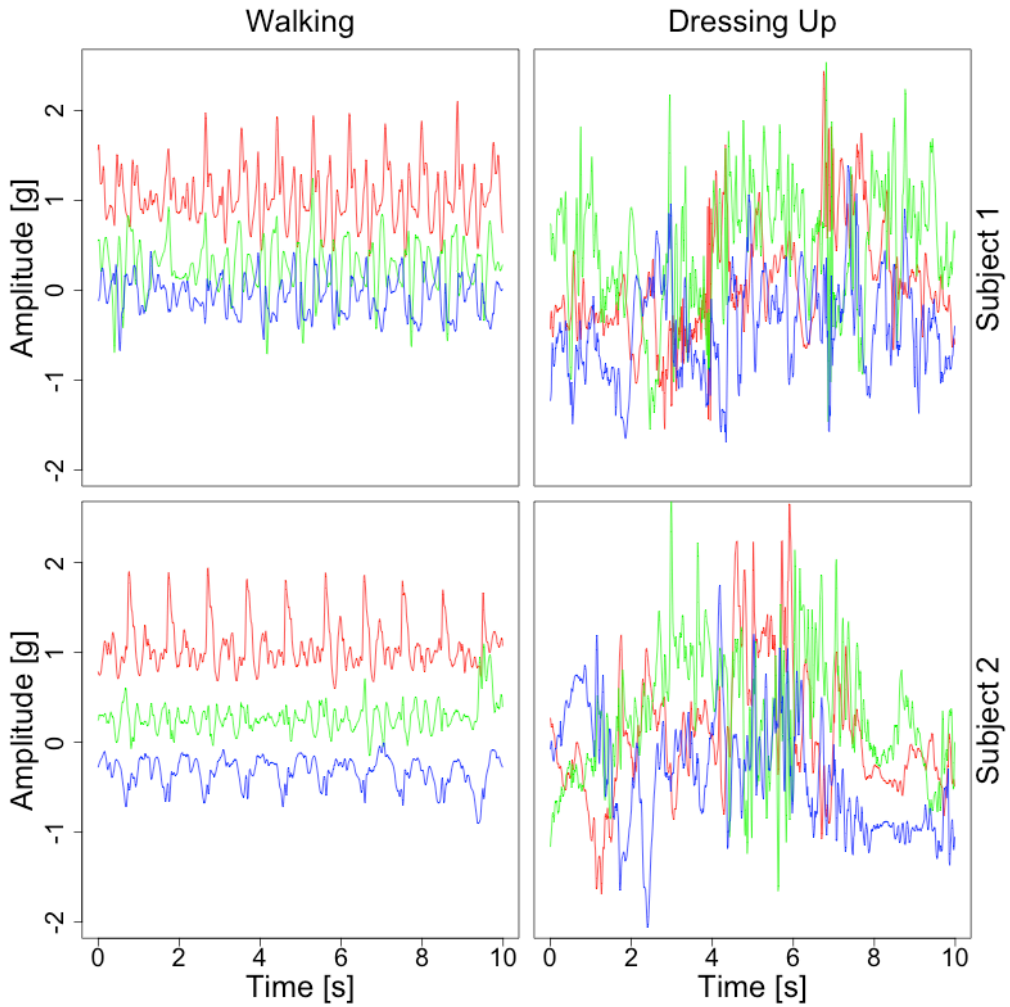
Why walking?

- Common form of physical activity
- Features of walking related to:
  - Survival
  - Mild Cognitive Impairment
  - Dementia
  - Stroke

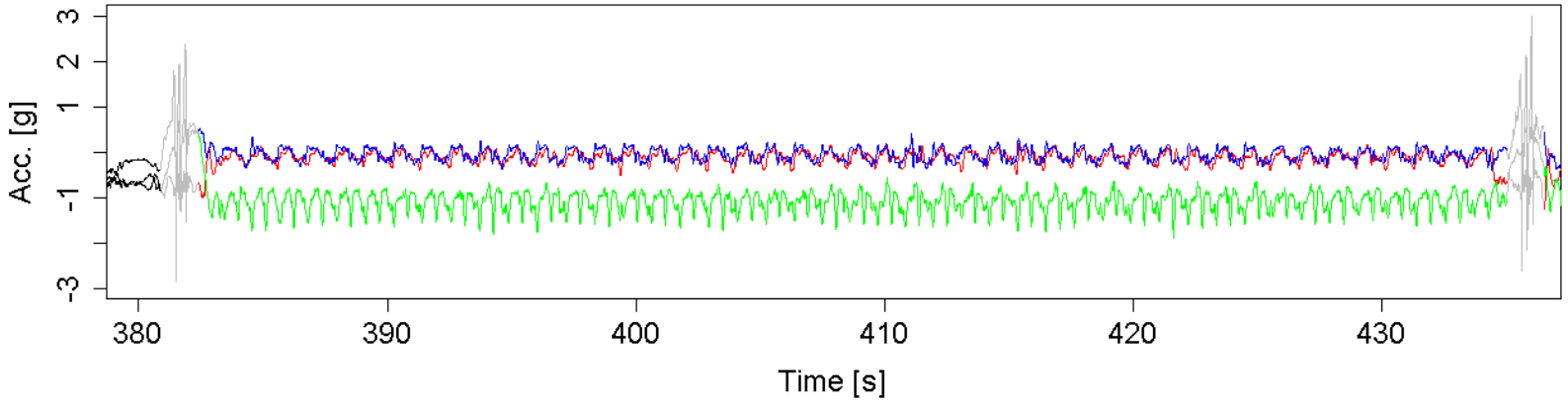


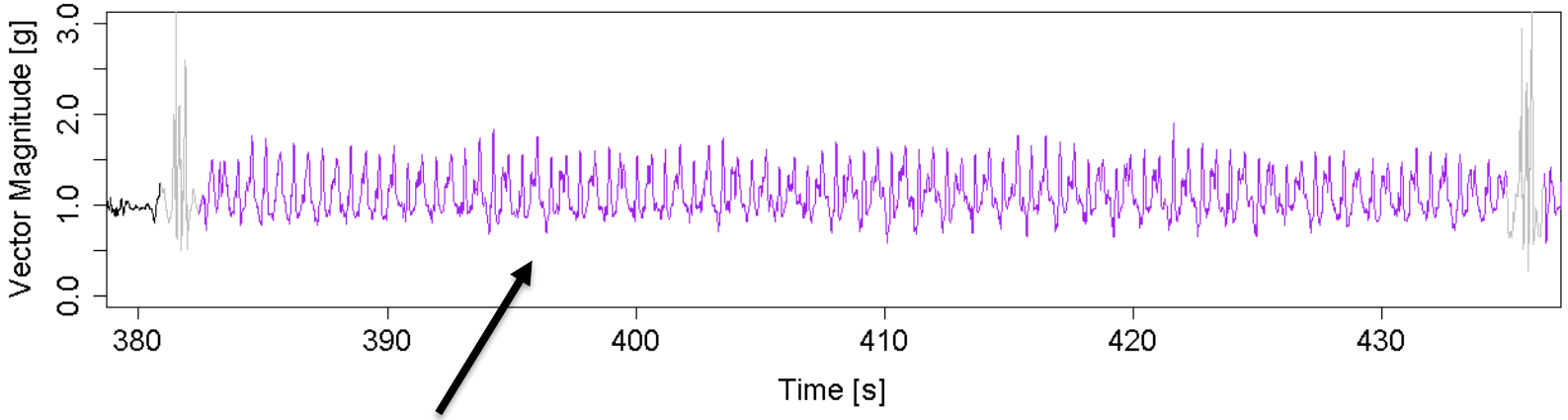
Urbanek, J, Zipunnikov, V, Harris, T, Crainiceanu, C, **Harezlak, J**, Glynn, NW, Validation of gait characteristics extracted from raw accelerometry during walking against measures of physical function, mobility, fatigability, and fitness, The Journals of Gerontology: Medical Sciences, 2018 Apr 17;73(5):676-681

Data heterogeneity



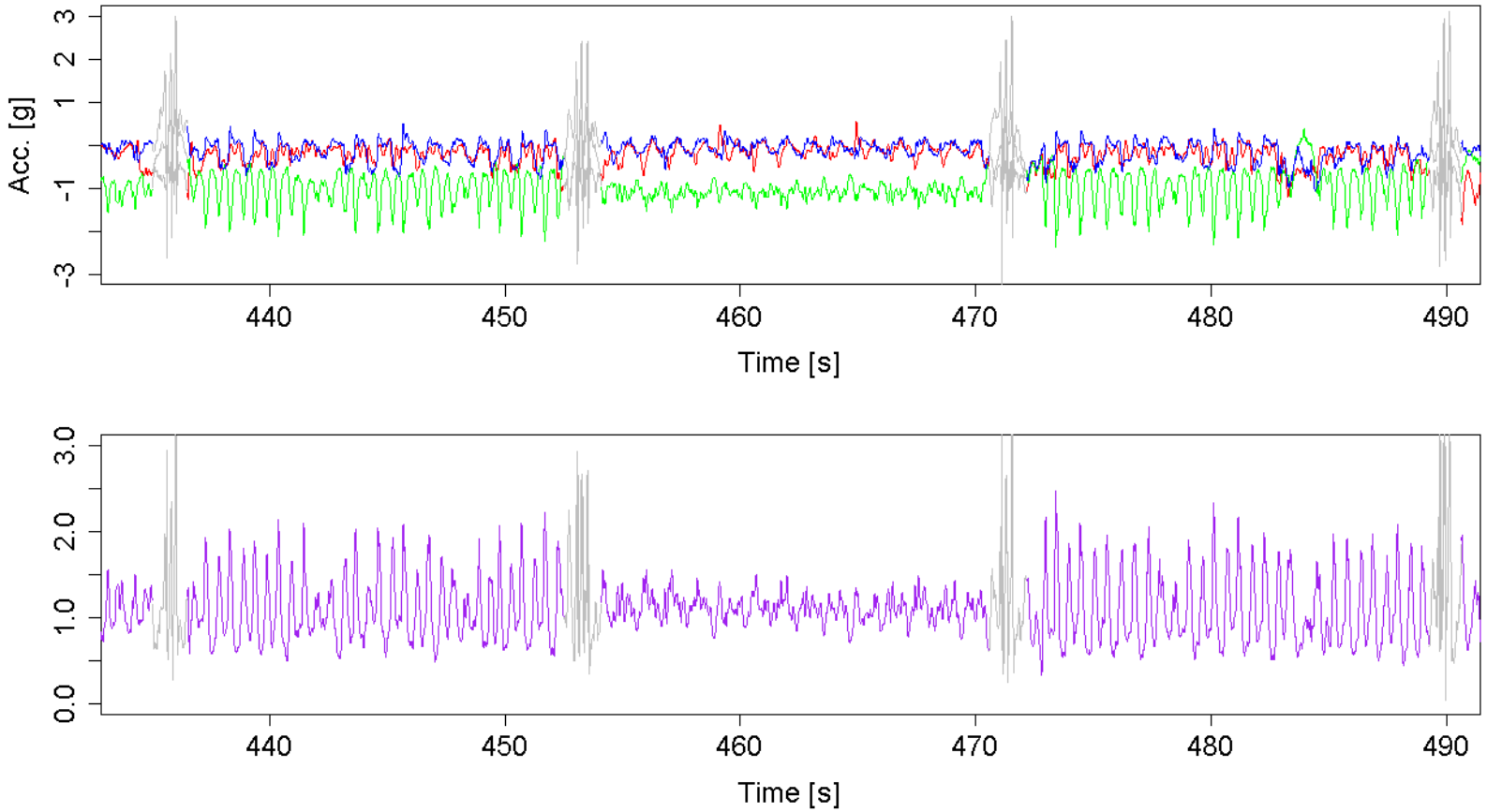
Walking on level ground





$$vm(t) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2}$$

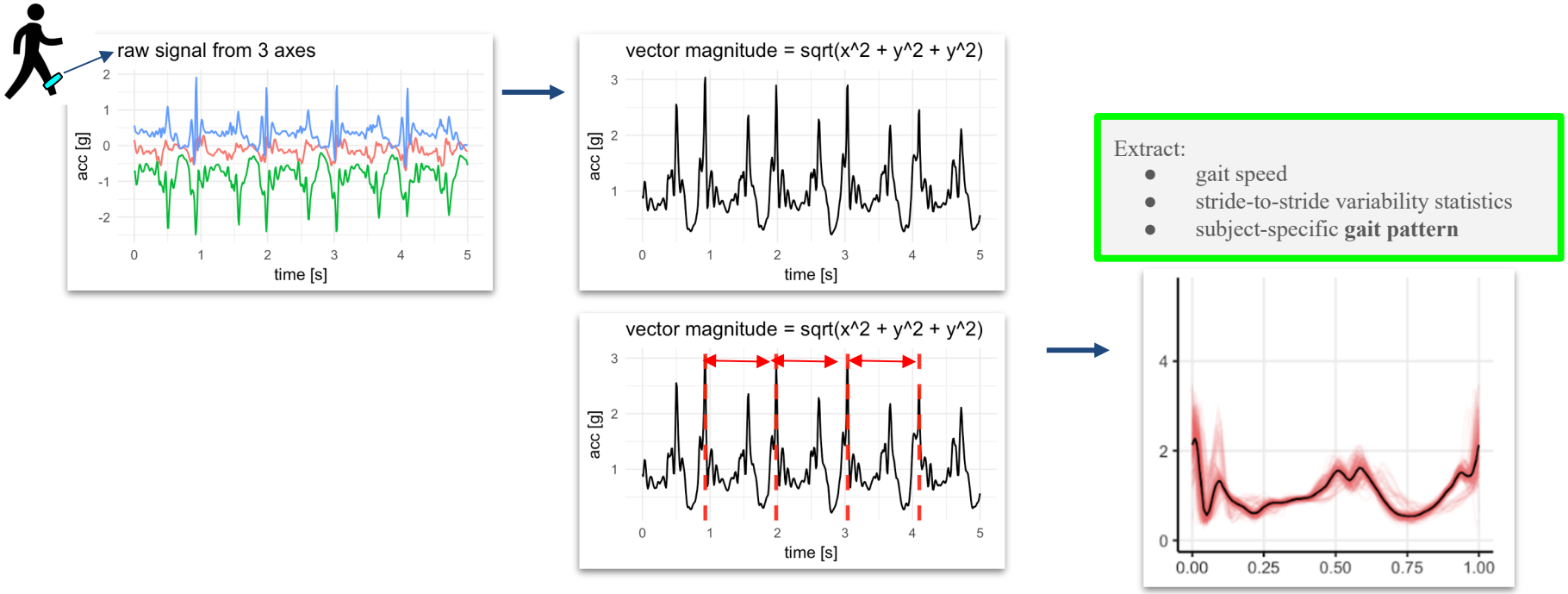
Down/up/down stairs



Courtesy of Marta Karas

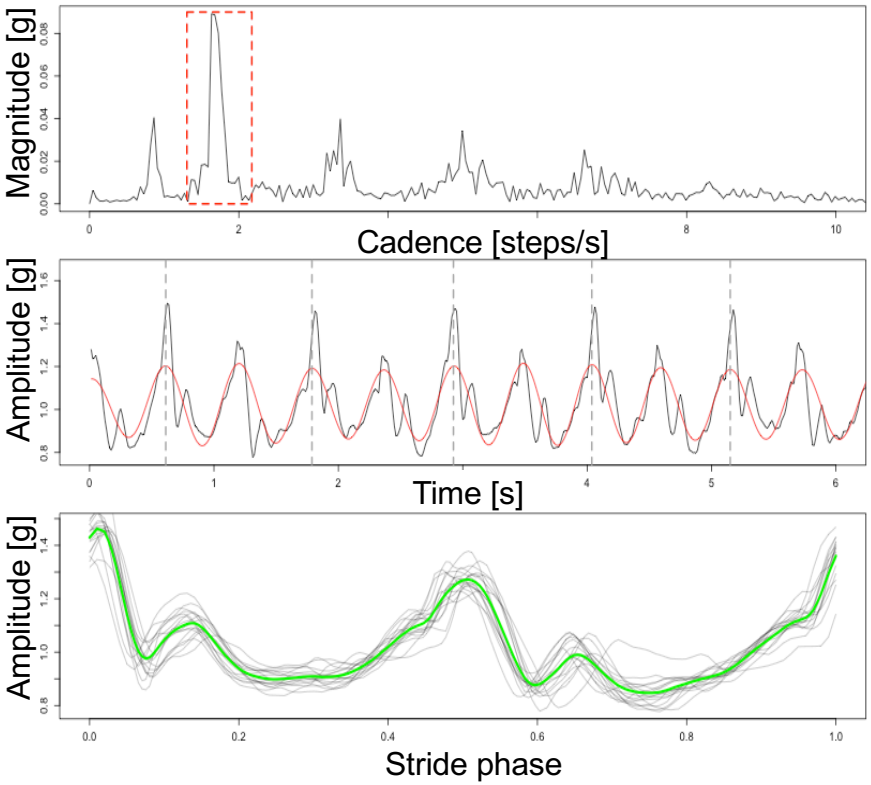
Goal: segment strides from accelerometry walking data

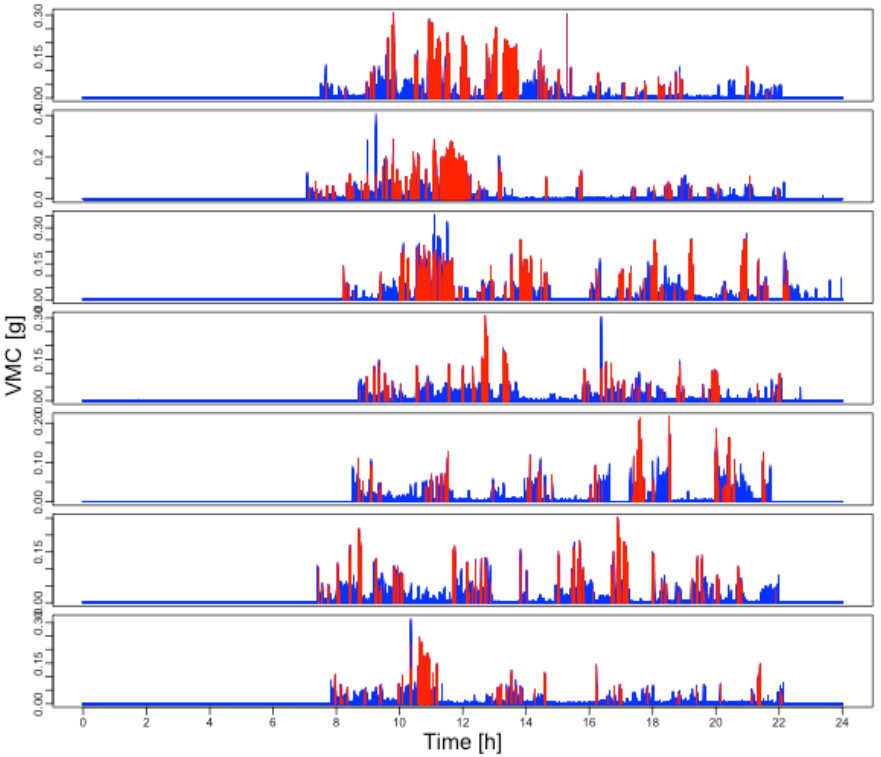
Raw subsecond-level accelerometry walking data → Segmentation of Strides (2 subsequent steps) → Feature extraction of human gait and gait pattern



Karas, M., Straczkiewicz, M., Fadel, W., **Harezlak, J.,** Crainiceanu, M., Urbanek, J., Adaptive empirical pattern transformation (ADEPT) with application to walking stride segmentation, Biostatistics (accepted)

Characteristics of walking





Urbanek, J, **Harezlak, J,** Glynn, NW, Harris, T, Crainiceanu, C, Zipunnikov, V, Stride variability measures derived from wrist-and hip-worn accelerometers, Gait & Posture 52 (2017) 217–223

Urbanek, J, Zipunnikov, V, Harris, T, Fadel, W, Glynn, N, Crainiceanu, C, **Harezlak J,** Prediction of sustained harmonic walking in the free-living environment using raw accelerometry data. Physiological measurement 39 (2), 02NT02

DECOS Study @ University of Pittsburgh

Population:

N = 51 (26 women) enrolled in the Development Epidemiologic Cohort Study (DECOS)

Age: between 70 and 90 (median = 78, SD = 5.68),

BMI: between 20.5 and 37.9 (median 25.9, SD = 3.91)

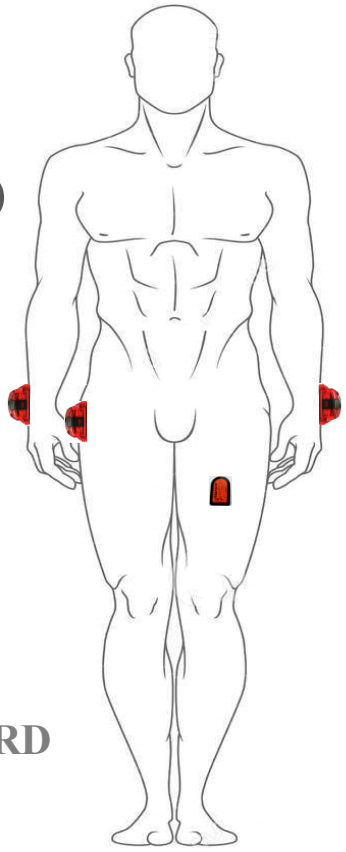
Data:

Free-living data collected for 7 days

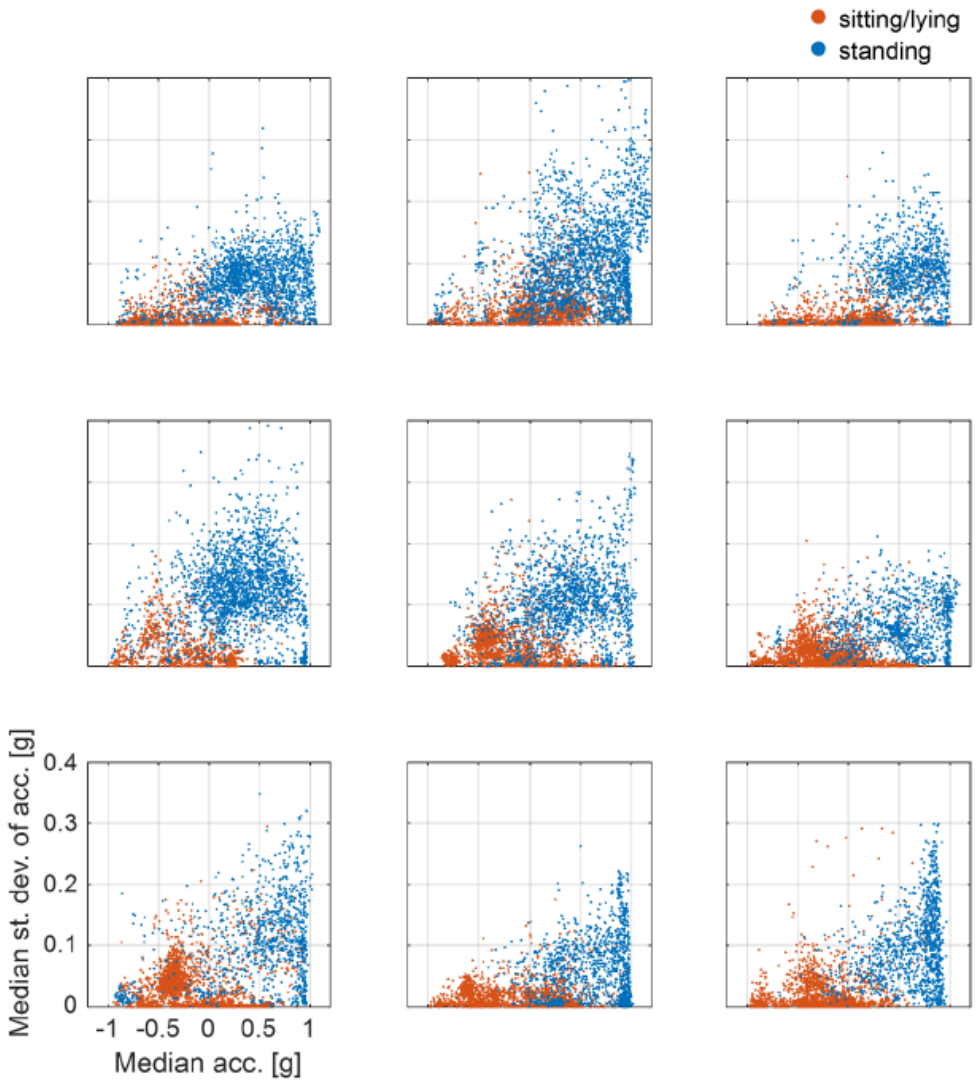
Left and right wrists: ActiGraph GT3X+ (80Hz)

Hip: ActiGraph GT3X+ (80Hz)

Thigh: activPAL 3 (20Hz) treated as silver standard



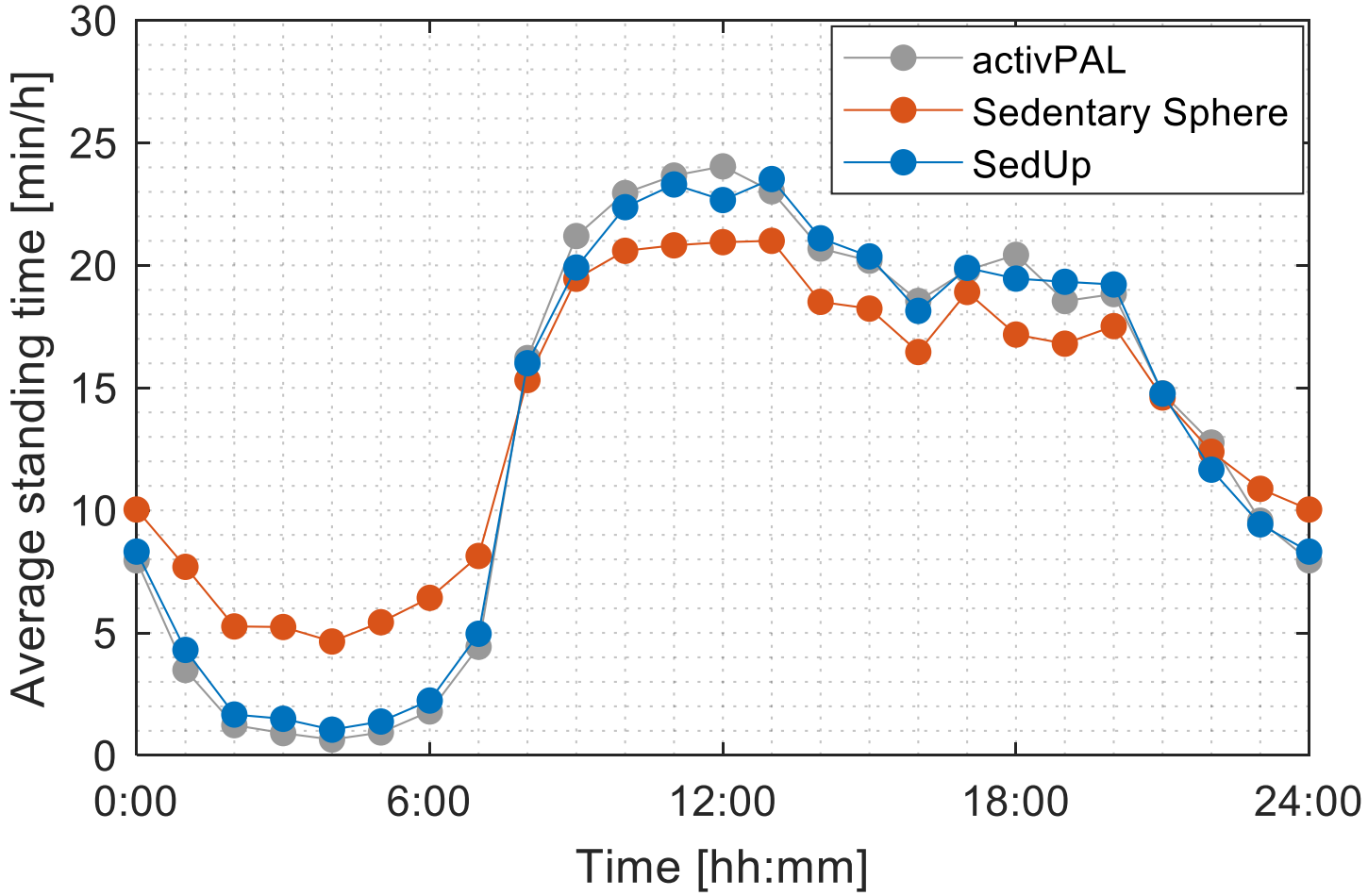
Sedentary vs. upright



Sitting/Standing Results

		Left wrist							Right wrist						
Method		**SedUp**						**SS**	**SedUp**						**SS**
Window [s]		**15**	**30**	**45**	**60**	**75**	**90**	**-**	**15**	**30**	**45**	**60**	**75**	**90**	**-**
TPR	**Median**	0.79	0.81	0.83	0.83	0.84	**0.83**	**0.66**	0.82	0.83	0.84	0.85	0.86	**0.86**	**0.65**
TNR	**Median**	0.90	0.90	0.91	0.91	0.91	**0.91**	**0.85**	0.91	0.92	0.92	0.93	0.93	**0.93**	**0.88**
MAPE [%]		13.3	13.0	12.7	12.6	12.6	**12.5**	**18.2**	15.7	15.3	15.2	15.0	15.1	**15.1**	**19.5**
MPE [%]		4.1	4.5	3.7	3.4	3.5	**2.9**	**4.1**	5.3	4.6	5.6	4.3	4.5	**4.5**	**6.7**

Results (left wrist)



Micro-scale analysis Summary

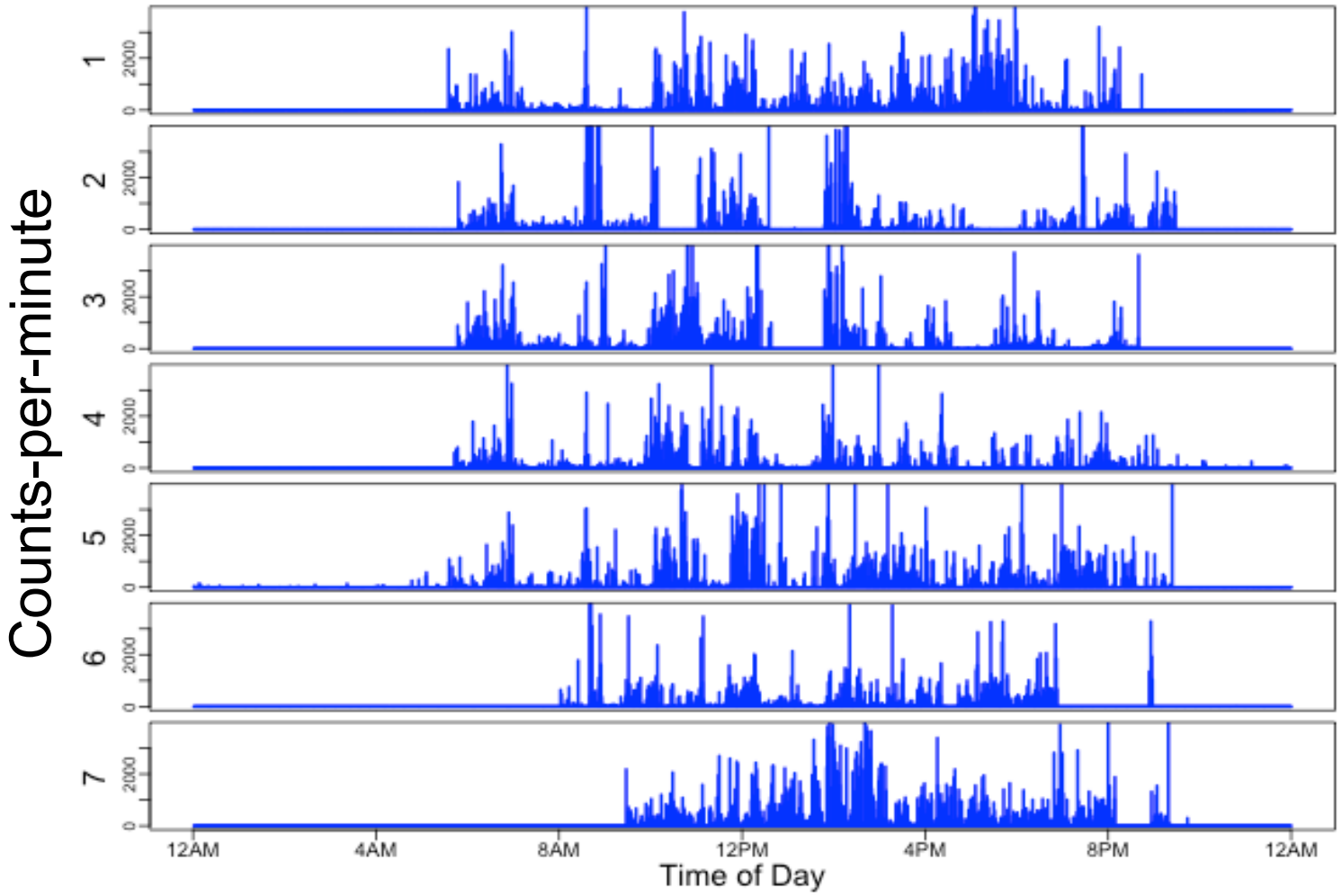
- Objective quantification of physical activity
- Objective measures of physical performance, fatigability, mobility and fitness (free-living)
- Algorithm development
- Statistical learning methods

Macro-scale analysis

- Aggregated 1-minute intervals
- General questions

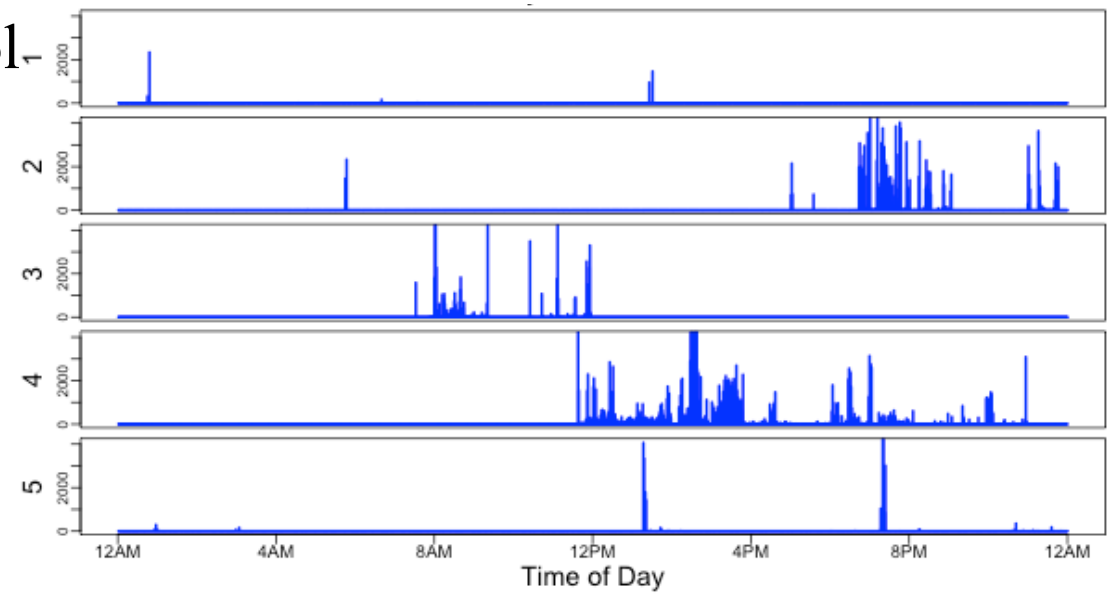
- US nationally representative sample
- Physical activity data for ~12 000 participants
- 7 consecutive days of observation
- ~10,000 observations per subject
- Variety of clinical and demographic outcomes
- Linked mortality data
- Publicly available dataset
- Big data challenges

PA data

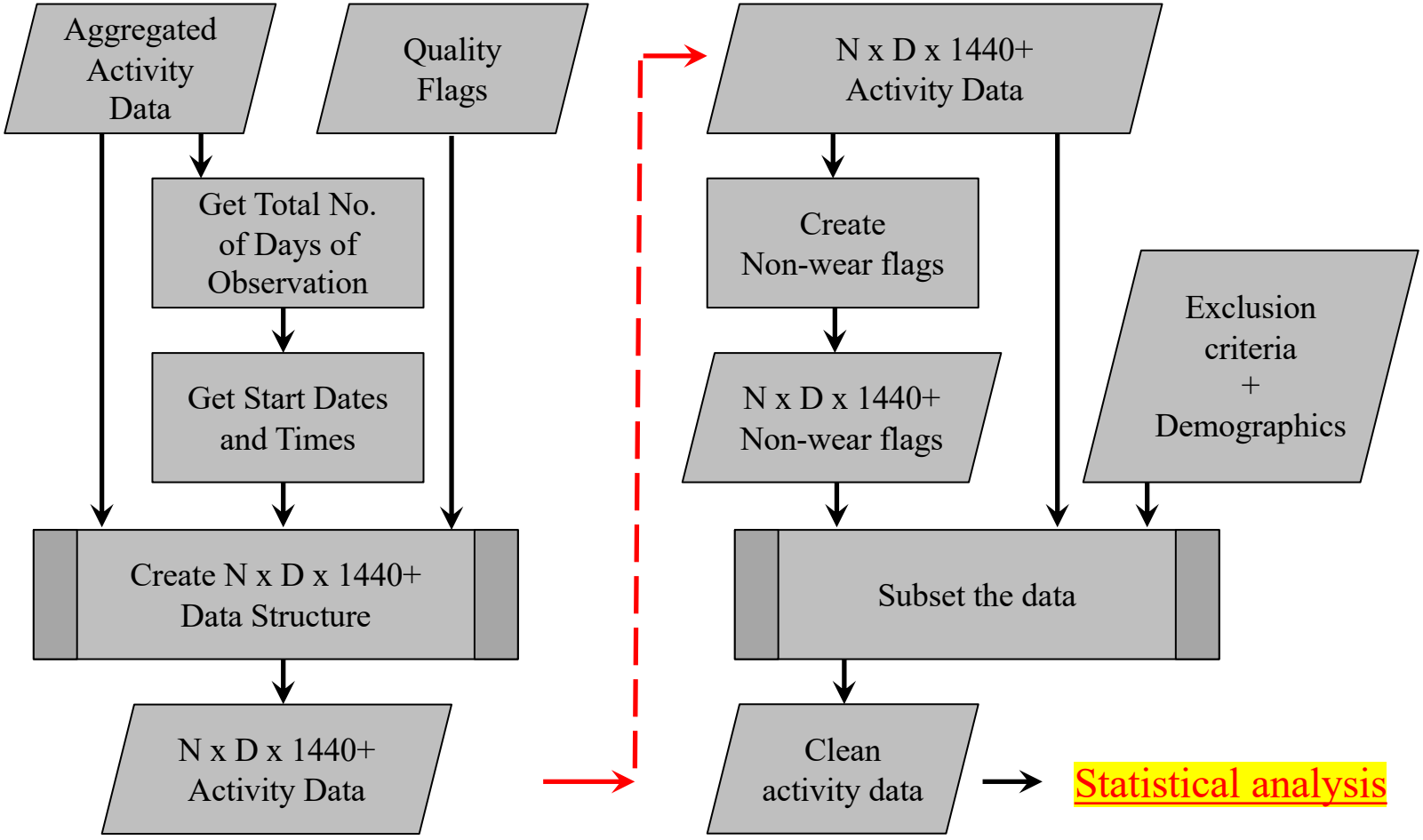

Counts-per-minute
1
2
3
4
5
6
7
0
2000
12AM
4AM
8AM
12PM
4PM
8PM
12AM
Time of Day

Macro-scale challenges

- Data format, storage and structure
- Start times of monitoring periods
- Different number of days across subjects
- Missing data due to non-wear
- Quality control



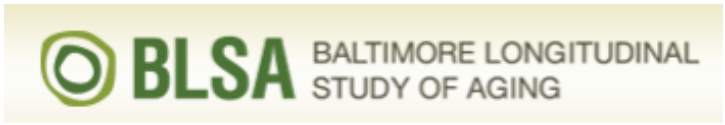
Data processing pipeline



Macro-scale: summary

- Curating data is extremely time intensive

- Proposed pipeline:
  - efficient analysis of PA data
  - fast complex analysis
  - applied to data from NHANES and BLSA



Acknowledgements

University of Wroclaw

- Michal Kos (Statistics)

Johns Hopkins School of Medicine

- Jacek Urbanek (Biostatistics)
- Marta Karas (PhD student)
- Ciprian Crainiceanu (Biostatistics)
- Vadim Zipunnikov (Biostatistics)

Indiana University

- William Fadel (Biostatistics)

Harvard University

- Marcin Straczkiewicz (Biostatistics)

University of Pittsburgh

- Nancy Glynn (Epidemiology)

Thank you