Wearable accelerometers, accelerometry data and automatic steps segmentation in `R`:

`strideter` and `convo` `R` packages

Marta Karaś
July 3rd, 2018
Wrocław

Outline

- What are wearable devices?
- Overview of accelerometry data
- Individual strides segmentation from raw accelerometry data with **`strideter`** **`R`** package: adaptive movelet approach
- Speed matters - **`convo`** **`R`** package for fast "running" statistics computation via Fast Fourier Transform

Accelerometers measure acceleration of a body

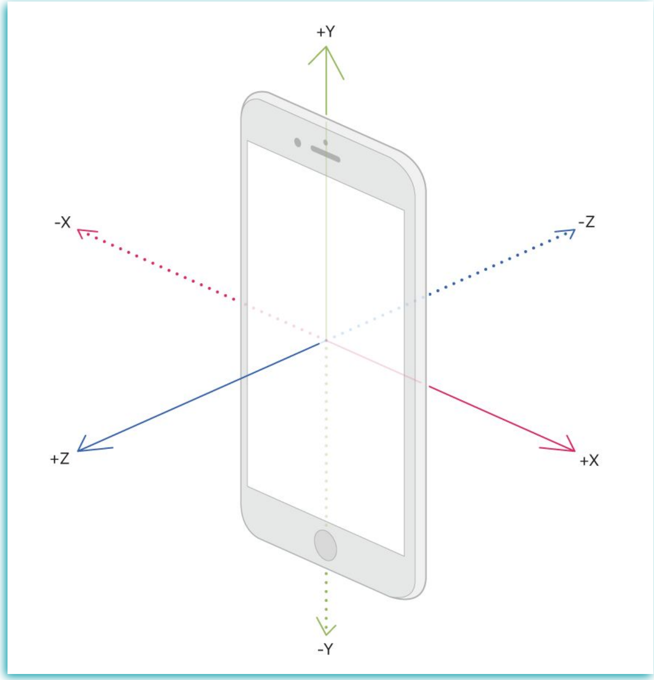
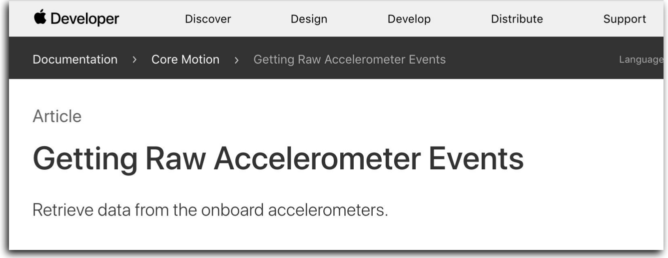
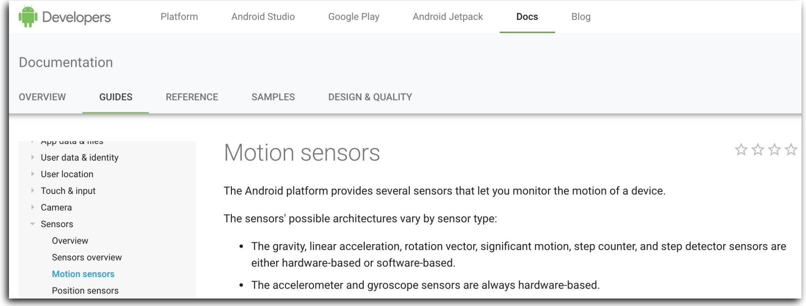


Fig. 1: A three-axis accelerometer measures acceleration along the x, y, and z axes.

[Fig. 1 source: https://developer.apple.com/documentation/coremotion/getting_raw_accelerometer_events]

- Raw accelerometry data accessible from many modern mobile phone devices (e.g. iOS, Android mobile operating systems)



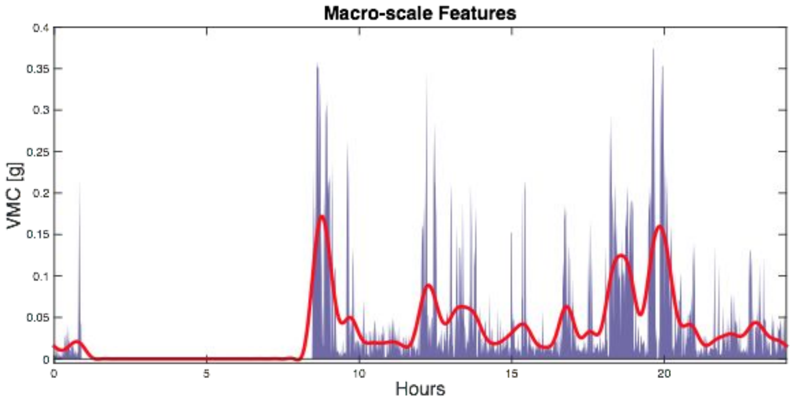


Actigraphy and actigraph units

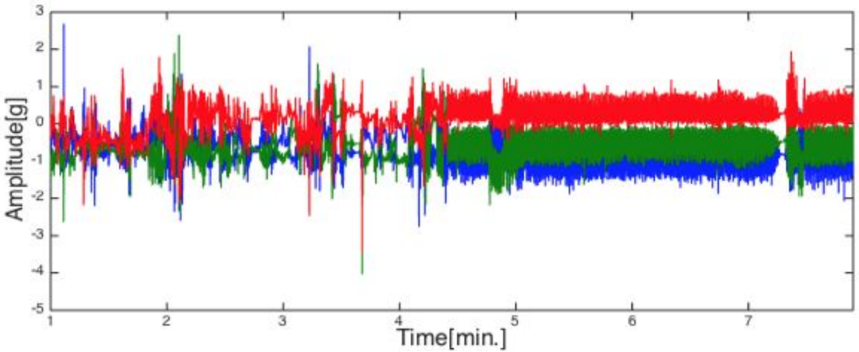
- Actigraphy: non-invasive monitoring of human activity
- Actigraph unit (wearable accelerometer)
  - has accelerometer, memory to store the collected measurements, interface (usually USB) to program the timer and download the data

Measurements: Macro- and Micro-scale

- **Macro-scale** – summarized data (e.g. 1 minute intervals)
  - Activity counts
  - Number of steps
  - Number of calories burned



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



Wearable accelerometers in health research

- **Objective measurement of physical activity**
- Compare physical activity in subgroups of a study participants
    - Time spent in different activity intensity levels, physical activity type classification,
- Epidemiology of aging
- Response to treatment, recovery after surgery
- Sleep quality

nhanes
National Health and Nutrition Examination Survey

DECOS
Developmental Epidemiologic Cohort Study

WOMEN'S HEALTH INITIATIVE

BLSA BALTIMORE LONGITUDINAL STUDY OF AGING

Examples of large epidemiological studies that include wearable accelerometer devices in their study protocol.

Monitoring of walking - Why important?

- Features of human gait are reportedly related to health outcomes
  - stroke, dementia, survival
  - gait speed, gait variability
- Accelerometry data allows to extract subject-specific gait pattern
  - athletic track running monitoring, post-stroke recovery monitoring
  - raw accelerometry data
    => $\boxed{\text{steps segmentation}}$
    => extraction features of human gait / gait pattern

Gait Speed and Survival in Older Adults

Dr. Stephanie Studenski, MD, MPH, Dr. Subashan Perera, PhD, Dr. Kushang Patel, PhD, Dr. Caterina Rosano, MD, PhD, Dr. Kimberly Faulkner, PhD, Dr. Marco Inzitari, MD, PhD, Dr. Jennifer Brach, PhD, Dr. Julie Chandler, PhD, Dr. Peggy Cawthon, PhD, Dr. Elizabeth Barrett Connor, MD, Dr. Michael Nevitt, PhD, Dr. Marjolein Visser, PhD, Dr. Stephen Kritchevsky, PhD, Dr. Stefania Badinelli, MD, Dr. Tamara Harris, MD, Dr. Anne B. Newman, MD, Dr. Jane Cauley, PhD, Dr. Luigi Ferrucci, MD, PhD, and Dr. Jack Guralnik, MD, PhD

Author information ► Copyright and License information ► Disclaimer

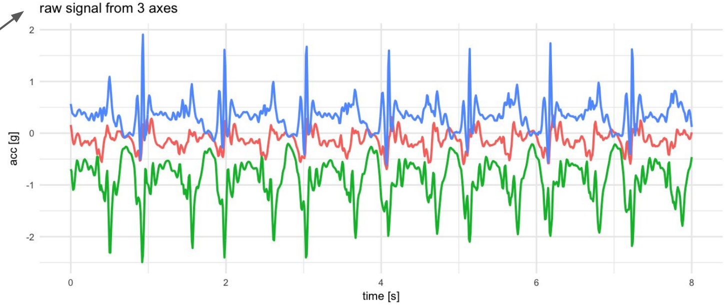
The publisher's final edited version of this article is available at JAMA
See other articles in PMC that cite the published article.

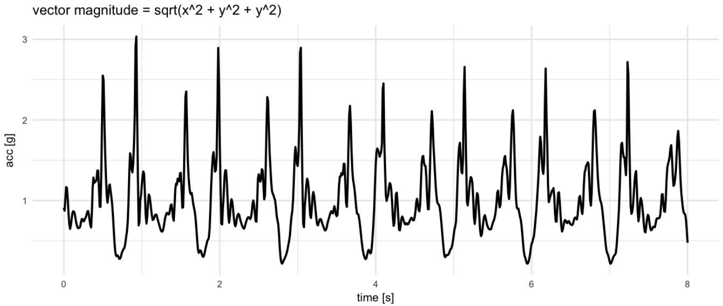
Abstract — Go to:

Context

Survival estimates help individualize goals of care for geriatric patients, but life tables fail to account for the great variability in survival. Physical performance measures, such as gait speed, might help account for variability, allowing clinicians to make more individualized estimates.

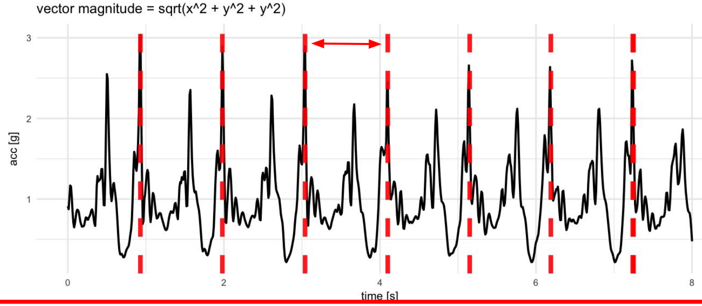
Person performing walking activity, wearing accelerometer device on their ankle





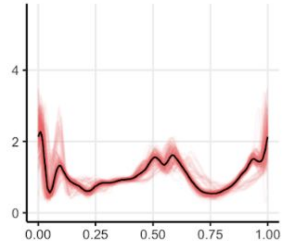
strideter R package: **automatic segmentation of individual strides** from accelerometry data of walking

stride = 2 consecutive steps



Extract:

- gait speed
- stride-to-stride variability statistics
- subject-specific gait pattern:

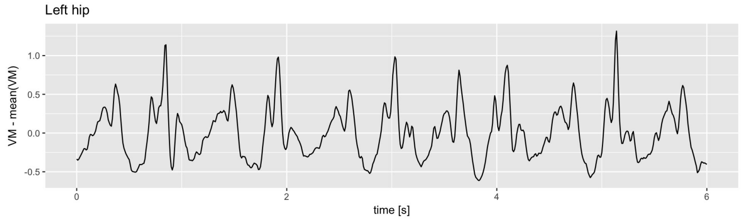


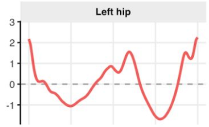
Adaptive movelets method for strides segmentation: big picture

Assume new raw accelerometry data which contains walking. We then consider **vector magnitude (vm)**

$$vm_t = \mathrm{sqrt}(x_t^{\,2} + y_t^{\,2} + z_t^{\,2})$$

for x, y, z - acceleration values collected from 3 device axes at each time point t.

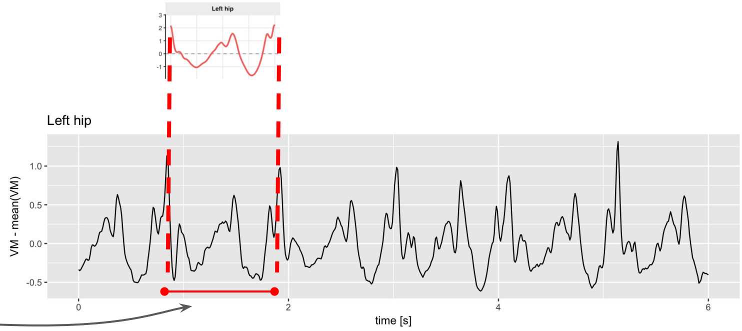




stride = 2 consecutive steps

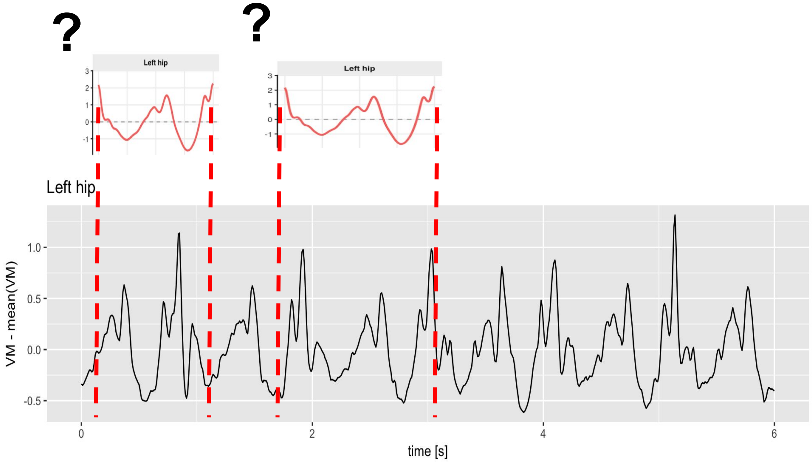
- Consider **stride movelet**: a predefined pattern of stride (based on some segmented data available)
- Use stride movelet to iteratively find windows of vm signal with which movelet has highest similarity (e.g. correlation / covariance)

- Call these windows a newly identified strides

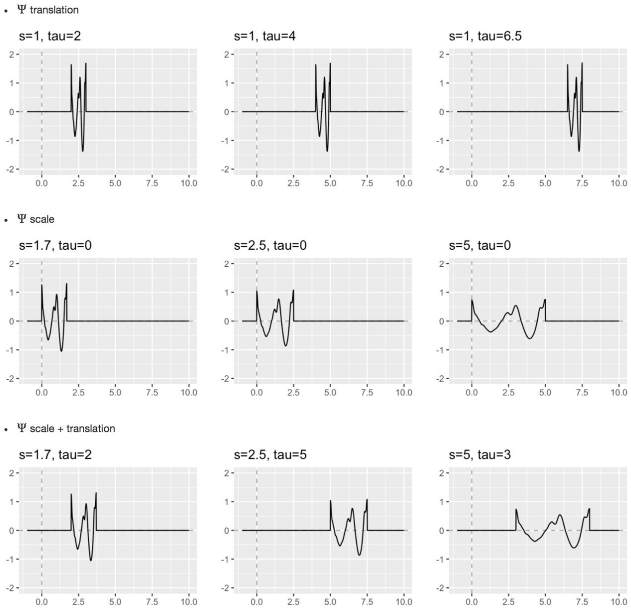


Method key feature: scaling and shifting the movelet

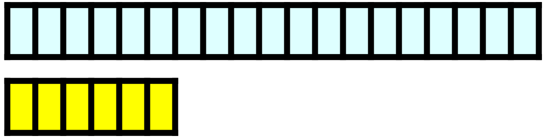
- To identify VM signal windows of the highest similarity with a movelet, we consider **different <u>scales</u> and <u>locations</u>** of a movelet (with respect to VM signal)



- Scaling and translating the movelet: example



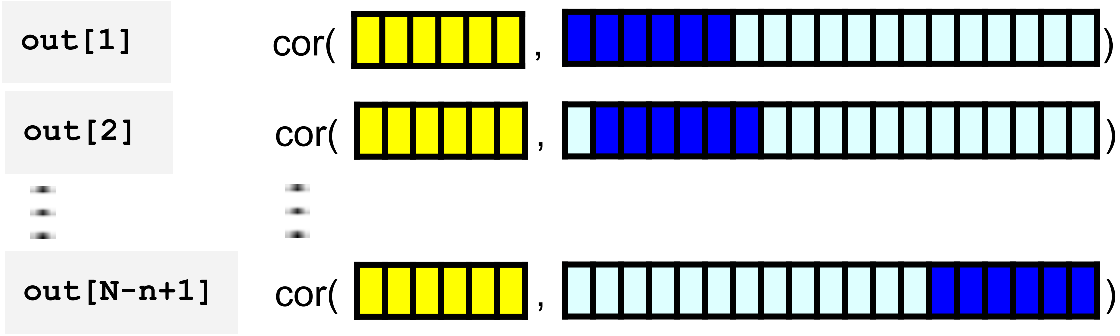
convo R package: fast computation of "running" (aka "rolling") statistics



Input 1: vector **x**: `len(x) = N`

Input 2: (shorter) vector **y**: `len(y) = n`

```
# compute running correlation between x and y
library(convo)
out <- RunningCor(x, y)
```



Author Marta Karas [aut, cre],
Jacek Urbanek [aut]

Maintainer Marta Karas <mkaras2@jhu.edu>

R topics documented:

Available on GitHub:
https://github.com/martakarass/convo

Example: Running Correlation with `convo`

```
library(convo)

N <- 1000000   ## million obs
n <- 100
x <- sin(seq(0, 1, length.out = N) * 2 * pi * 10)
y <- x[1:n]

RunningCor.CONVENTIONAL <- function(x, y){
  N <- length(x)
  n <- length(y)
  sapply(1:(N - n + 1), function(i){
    cor(x[i:(i + n - 1)], y)
  })
}

system.time(out  <- RunningCor(x, y))
# user  system elapsed
# 1.592    0.098    1.691
system.time(out2 <- RunningCor.CONVENTIONAL(x, y))
# user  system elapsed
# 34.132    0.489   34.620
```

`convo` package implementations of rolling statistics:

- formulate **statistics formulas via convolutions**
- convolutions are computed via Fast Fourier Transform
  - speed gain: complexity of the convolution can be **reduced from O(n^2) to O(n log n)** with FFT compared to conventional convolution computation

Example: strides segmentation with **strideter**

```
library(strideter)

## consider raw acc signal from left hip
## vm = sqrt(x^2 + y^2 + z^2)
head(vm)
# [1] 1.112560 1.141720 1.143985 1.130872 1.130655 1.104831

## create list of sub-population specific stride templates
## (template_lh is attached to package)
movelets <- list(template_lh$clustersN_2[1,],
                 template_lh$clustersN_2[2,])

# run algorithm
out <- SegmentStrides(vm,
                      movelets,
                      fs = 100,
                      similarity.measure = "cor",
                      similarity.measure.thresh = 0.7,
                      cadmin = 1,
                      cadmax = 3.5,
                      run.parallel = TRUE)
```

`vm`: accelerometry signal vector magnitude (numerical vector)

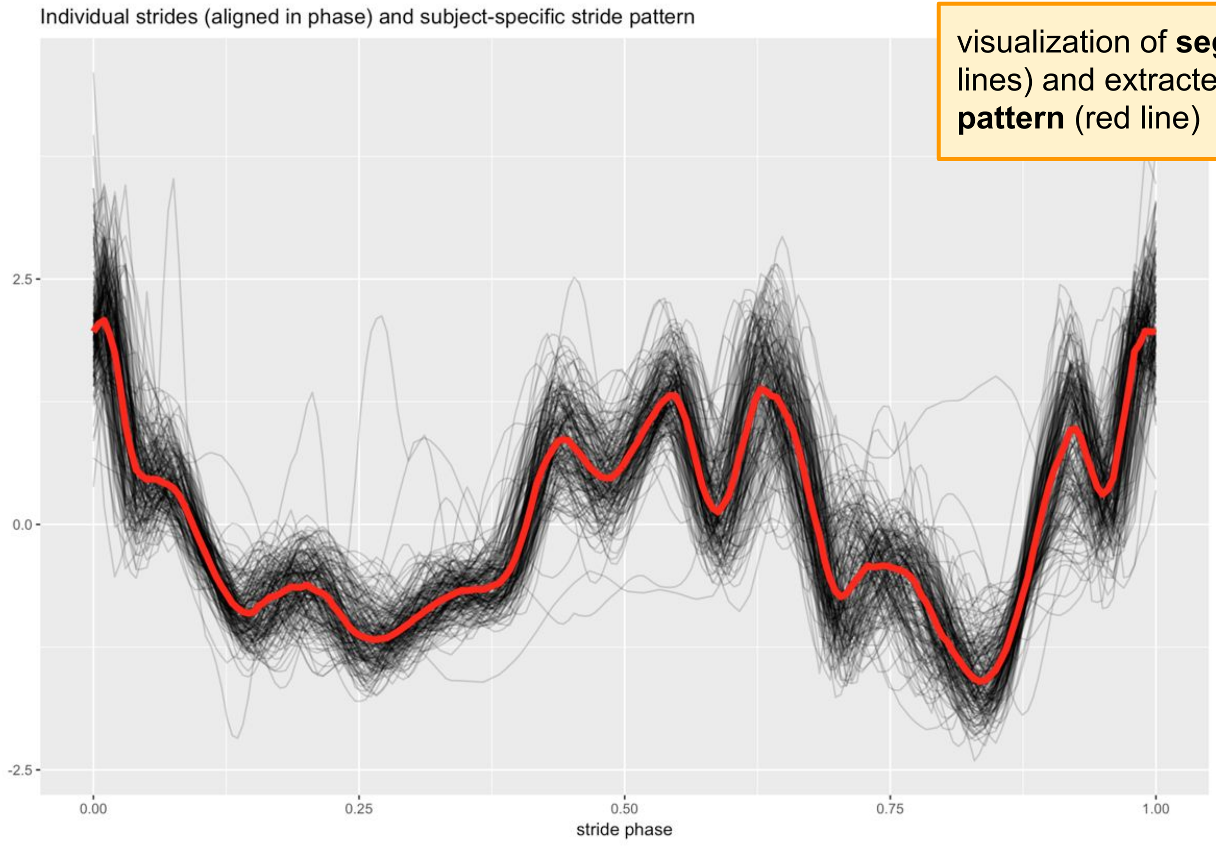
`movelets`: list of stride patterns (might be multiple) of data from left hip; attached to package

similarity measure type and its value lower threshold to identify a stride

`cadmin/cadmax`: definition of range of parameters defining stride lengths we expect in data result: identified strides summary table

```
# result
head(data.frame(tau_i = out$tau_i,     ## index of stride start
                T_i   = out$T_i,        ## length of stride
                sim_i = out$sim_i))     ## signal to movelet similarity

#    tau_i T_i     sim_i
# 1     21  84 0.8900953
# 2    104 100 0.9273885
# 3    203 101 0.8627565
# 4    303 100 0.8948592
# 5    402 100 0.9133245
# 6    501  98 0.8631847
```



visualization of **segmented strides** (black lines) and extracted subject-specific **stride pattern** (red line)

Take-home messages

- Wearable accelerometers provide for objective measurement of physical activity
- Walking accelerometry monitoring allows for extracting gait features of scientific importance
- `strideter` **R** package implements adaptive movelet approach for automatic stride segmentation from subsecond accelerometry data of continuous walking
- `convo` **R** package implements methods for fast computation of “Running” statistics via FFT

Collaborators

Johns Hopkins University

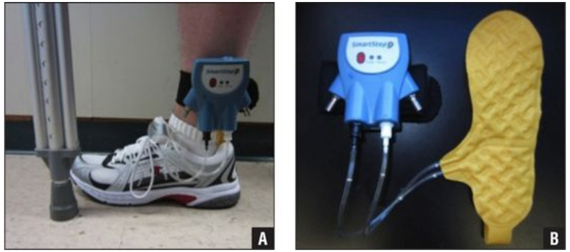
Wearable and Implantable Technology (WIT) working group:

- Jacek K. Urbanek (School of Medicine, Department of Medicine)
- Ciprian Crainiceanu (School of Public Health, Department of Biostatistics)

Thank you for your attention!

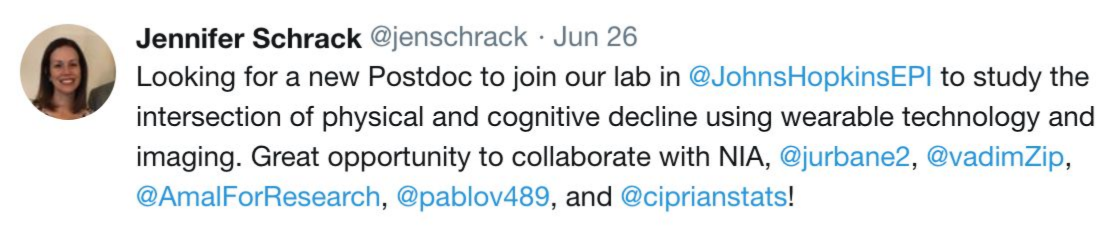
Comments? Questions? mkaras2@jhu.edu

@martakarass



Notes

- **Postdoc position open to work with wearable technology** at Johns Hopkins Bloomberg School of Public Health



- July 2018 - a good time to start working at applications for USA PhD programs (most Biostatistic program deadlines: December 2018)