Wearable and Implantable Technology (WIT)

Ciprian Crainiceanu
Professor of Biostatistics
www.ciprianstats.org

JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

Ranking predictors of five-year all-cause mortality in the US

Rank	Variable	AUC	Rank	Variable	AUC
1	**TAC**	**0.770**	**16**	**sPC6**	**0.657**
2	Age	0.757	**17**	**$TLAC_{6\text{-}8am}$**	**0.633**
3	**$TLAC_{8\text{-}10pm}$**	**0.753**	**18**	Education	0.611
4	**MVPA**	**0.748**	**19**	Drinking	0.593
5	**$TLAC_{4\text{-}6pm}$**	**0.740**	**20**	Smoking	0.574
6	**$TLAC_{12\text{-}2pm}$**	**0.735**	**21**	CHF	0.569
7	**ASTP**	**0.734**	**22**	BMI	0.550
8	**$TLAC_{10am\text{-}12pm}$**	**0.734**	**23**	Cancer	0.559
9	**$TLAC_{2\text{-}4pm}$**	**0.730**	**24**	Diabetes	0.556
10	**ST**	**0.728**	**25**	Gender	0.554
11	**TLAC**	**0.722**	**26**	Stroke	0.548
12	**$TLAC_{8\text{-}10am}$**	**0.684**	**27**	CHD	0.548
13	Mobil. Prob.	0.672	**28**	Race	0.514
14	**$TLAC_{8\text{-}10pm}$**	**0.671**	**29**	**$TLAC_{12am\text{-}2am}$**	**0.519**
15	**SATP**	**0.660**	**30**	**Wear time**	**0.459**

NHANES 2003-2006, age: 50-84, total: 2969, cases: 294, controls: 2675

Getting the organized NHANES accelerometry data

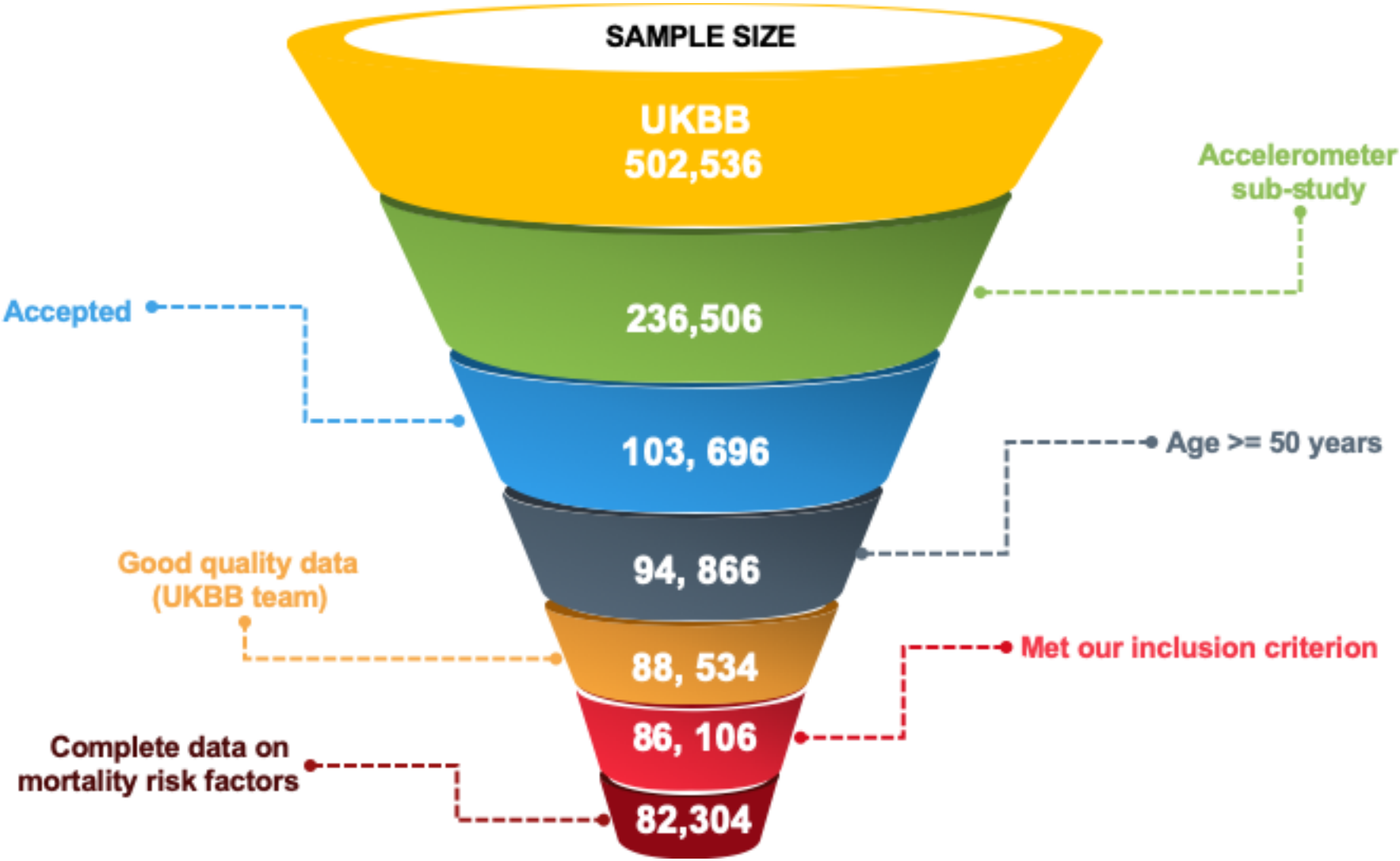
- NHANES data package (*rnhanesdata*):

  https://github.com/andrew-leroux/rnhanesdata

- Installing the rnhanesdata

  devtools::install_github(andrew-leroux/rnhanesdata)

UK Biobank accelerometry at a glance



Ranking predictors of time to death in the UK

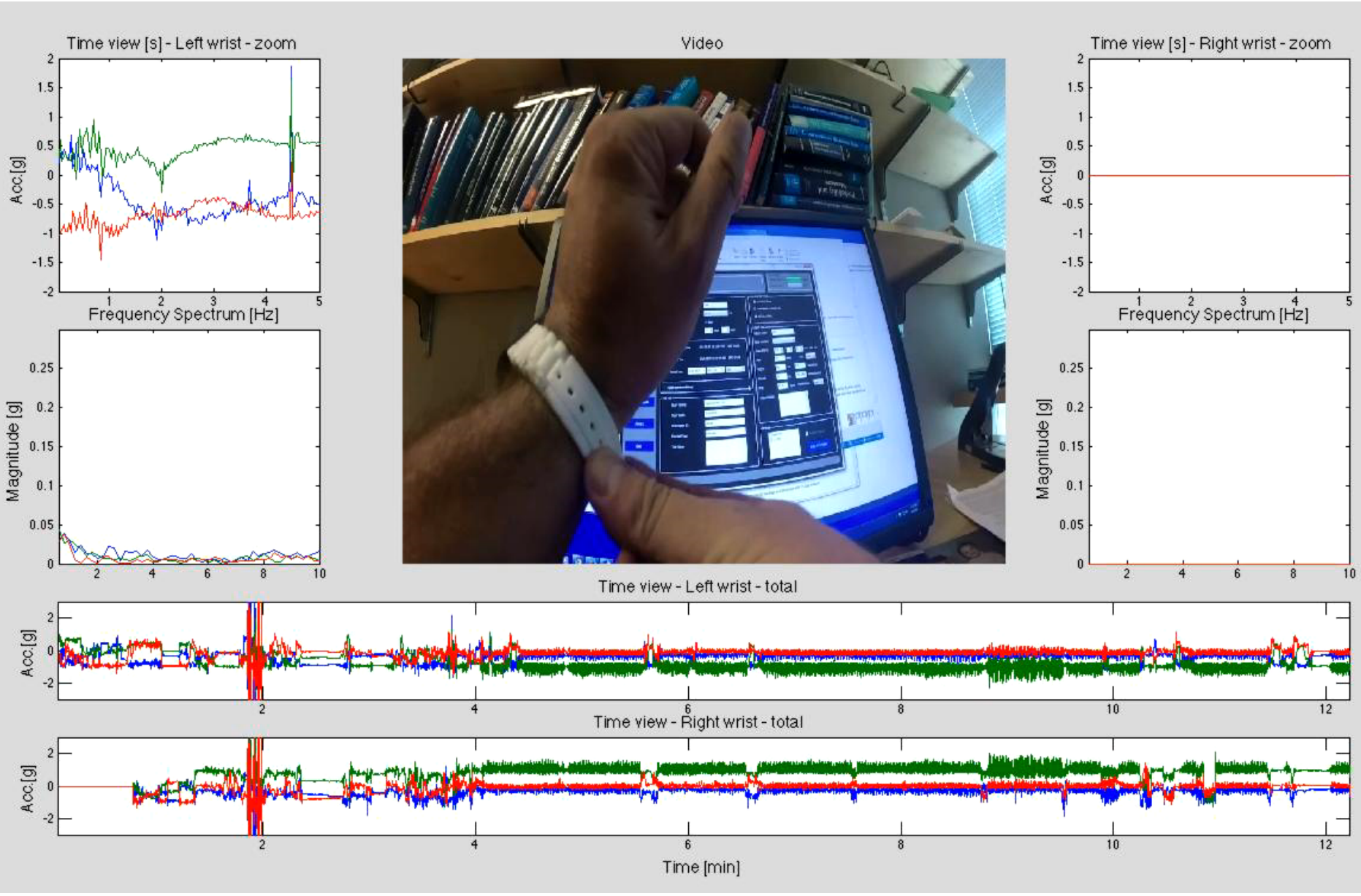
Rank	Variable	C	Rank	Variable	C
1	Age	0.681	**16**	Gender	0.587
2	**MVPA**	**0.661**	**17**	SR Health	0.585
3	**TA**	**0.661**	**18**	**$TLA_{10am-12pm}$**	**0.585**
4	**RA**	**0.644**	**19**	**TLA_{8-12pm}**	**0.584**
5	**M10**	**0.641**	**20**	**LIPA**	**0.578**
6	**TLA_{4-6pm}**	**0.640**	21	Smoking	0.577
7	**TLA_{6-8pm}**	**0.629**	22	**TLA_{8-10am}**	**0.575**
8	**TLA**	**0.628**	23	UW Speed	0.574
9	**ST**	**0.627**	24	BMI	0.565
10	**TLA_{2-4pm}**	**0.617**	25	**L5**	**0.563**
11	**TLA_{12-2pm}**	**0.609**	26	High BP	0.562
12	**ASTP**	**0.599**	27	**TLA_{2-4am}**	**0.557**
13	**SATP**	**0.595**	28	**TLA_{6-8am}**	**0.549**
14	SR Disability	**0.590**	29	Alcohol Cons.	0.545
15	**DARE**	**0.589**	30	**TLA_{4-6am}**	**0.543**

UK Biobank, age: 50+, total: 82,304, cases: 1959, follow-up: 439,707 py

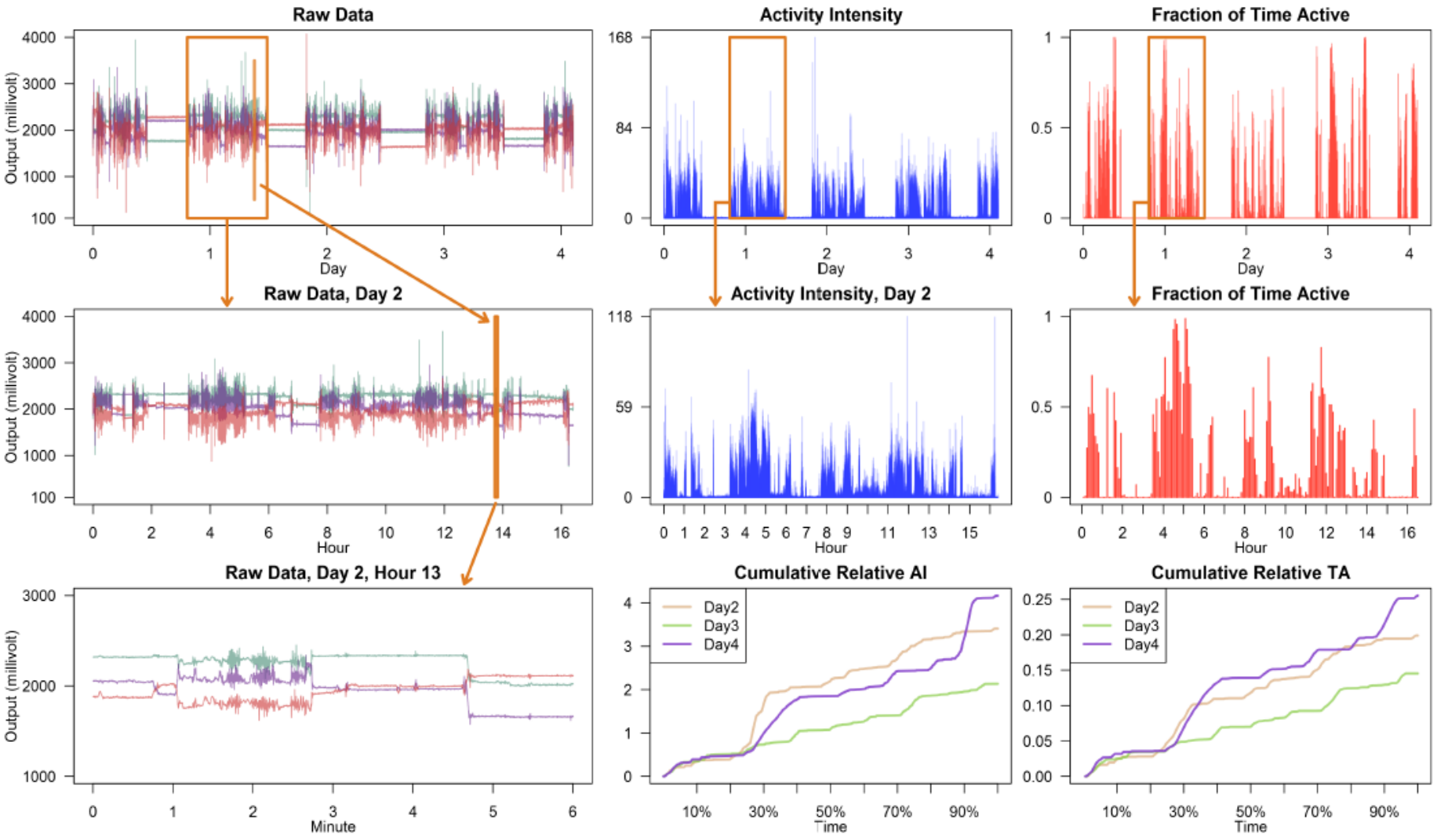
How much does activity add to known mortality risk factors?

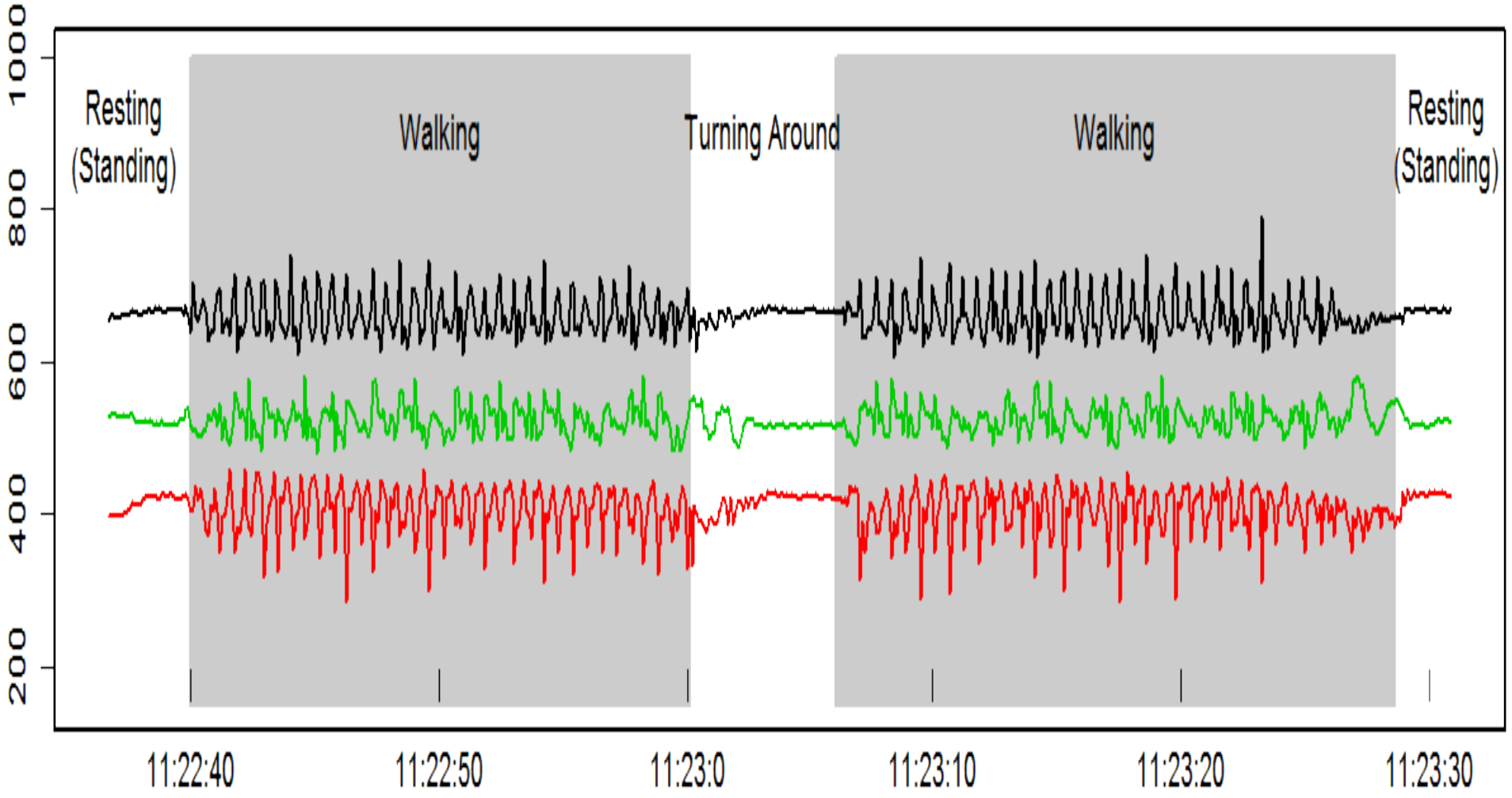
Stopping Rule: $\delta C \geq 0.001$			
Variable	Cumulative Concordance	δC	$\hat{\beta} \pm 2\text{SE}(\hat{\beta})$
Age	0.669	0.669	0.077 (0.065, 0.089)
Self-reported overall health	0.701	0.032	
Excellent			-0.071 (-0.278, 0.136)
Fair			0.178 (0.001, 0.355)
Poor			0.531 (0.244, 0.819)
Cigarette Smoker	0.714	0.013	
Former			0.122 (-0.026, 0.270)
Current			0.851 (0.642, 1.059)
Gender (male)	0.723	0.009	0.295 (0.149, 0.440)
Longstanding illness/disability	0.730	0.006	0.300 (0.144, 0.456)
Cancer	0.733	0.003	0.406 (0.208, 0.603)
High blood pressure	0.735	0.002	0.175 (0.029, 0.322)
Injury/illness within past 2 years	0.737	0.002	0.319 (0.123, 0.515)
Relative amplitude (RA)	0.758	0.021	-0.276 (-0.341, -0.212)
TLA 4-6PM	0.760	0.002	-0.219 (-0.297, -0.141)

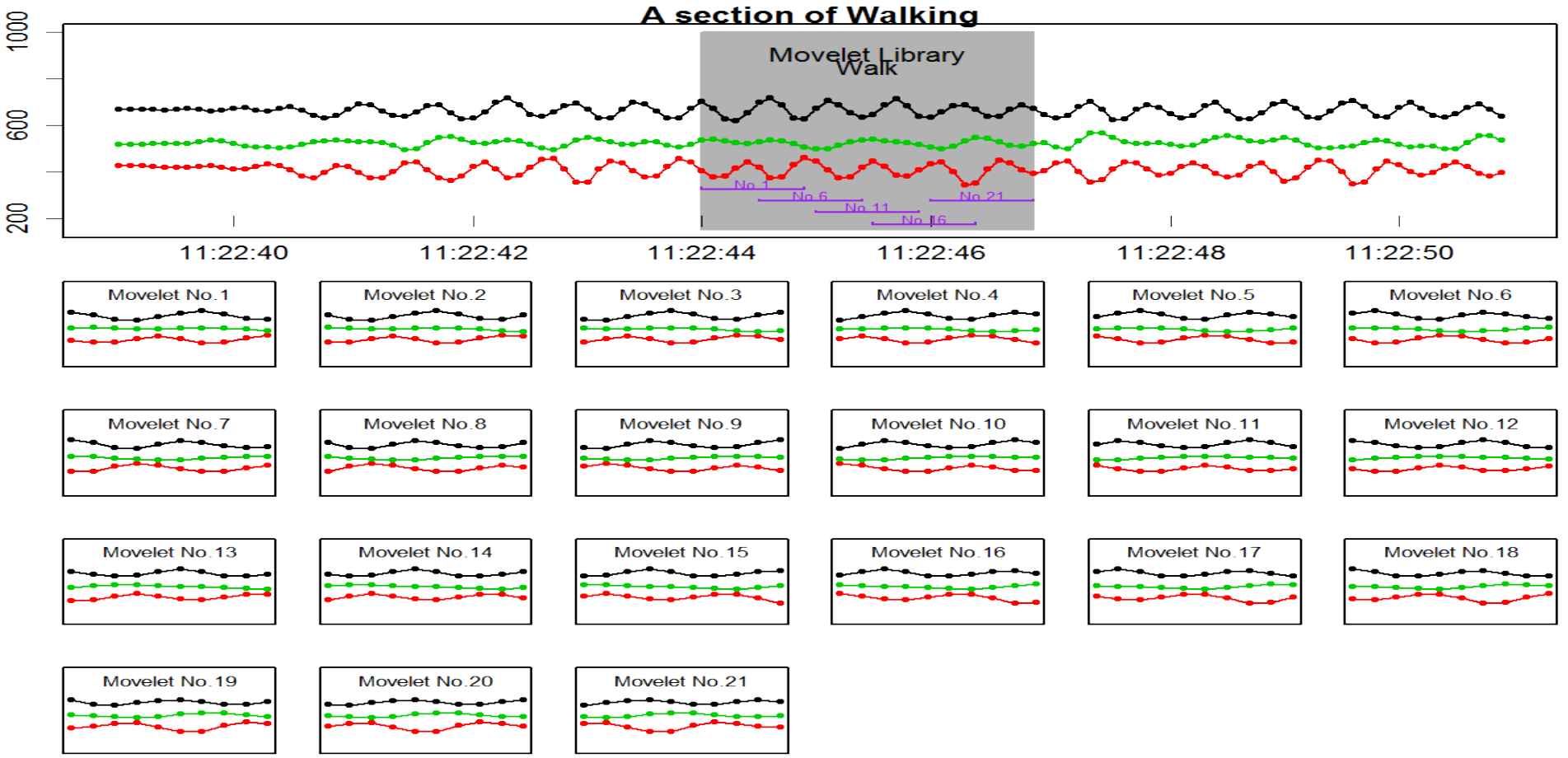
What kind of sensors?


Time view [s] - Left wrist - zoom
Acc.[g]
Frequency Spectrum [Hz]
Magnitude [g]
Video
Time view [s] - Right wrist - zoom
Acc.[g]
Frequency Spectrum [Hz]
Magnitude [g]
Time view - Left wrist - total
Acc.[g]
Time view - Right wrist - total
Acc.[g]
Time [min]

Micro- and macro-level data

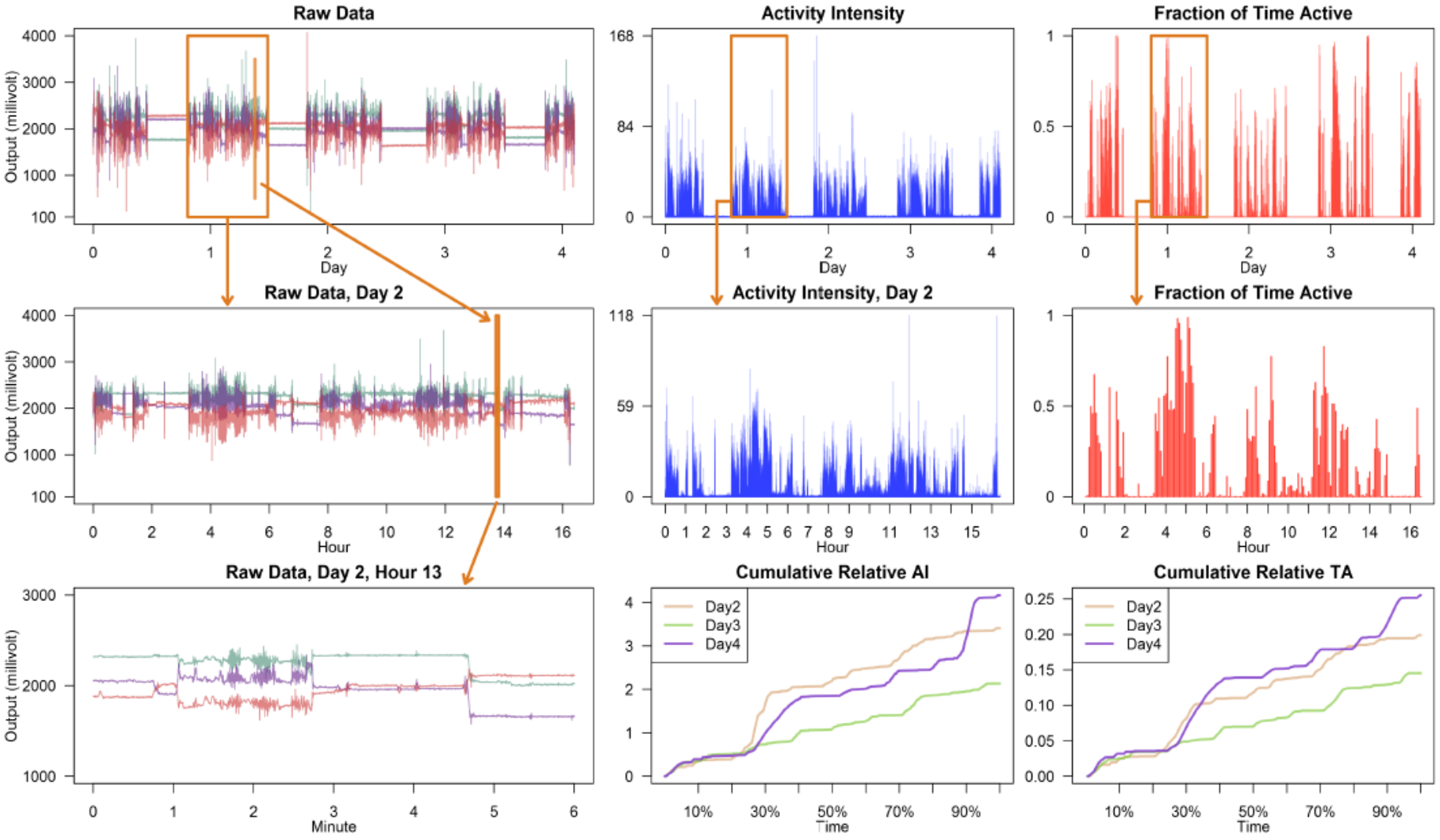



Resting
(Standing)
Walking
Turning Around
Walking
Resting
(Standing)
1000
800
600
400
200
11:22:40
11:22:50
11:23:0
11:23:10
11:23:20
11:23:30

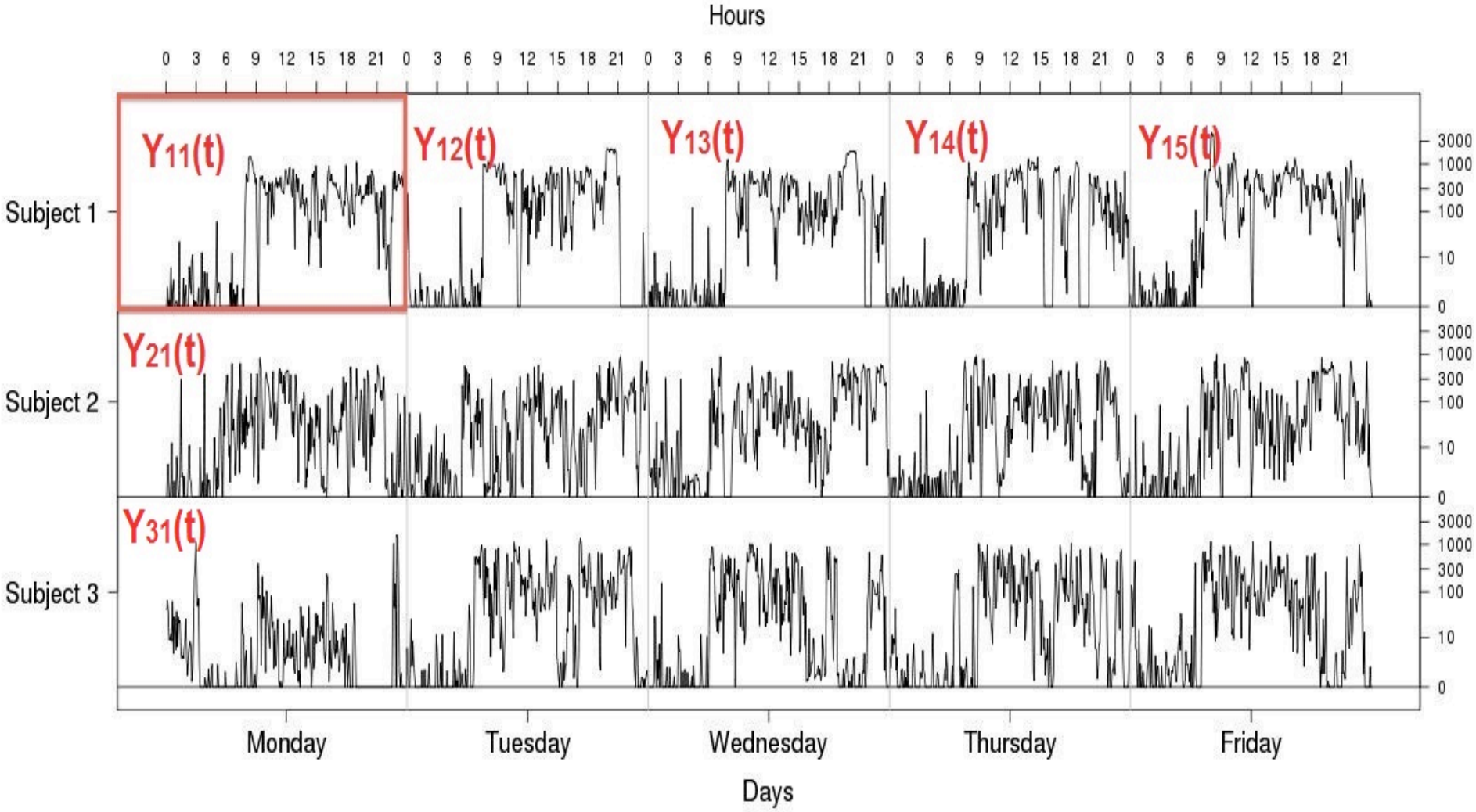


Chapter	Activity	Movelets
$\mathcal{C}_1$	Activity 1	$\{M_i(t) : L_i(t) = \text{Act}_1\}$
$\mathcal{C}_2$	Activity 2	$\{M_i(t) : L_i(t) = \text{Act}_2\}$
$\vdots$	$\vdots$	$\vdots$
$\mathcal{C}_A$	Activity A	$\{M_i(t) : L_i(t) = \text{Act}_A\}$

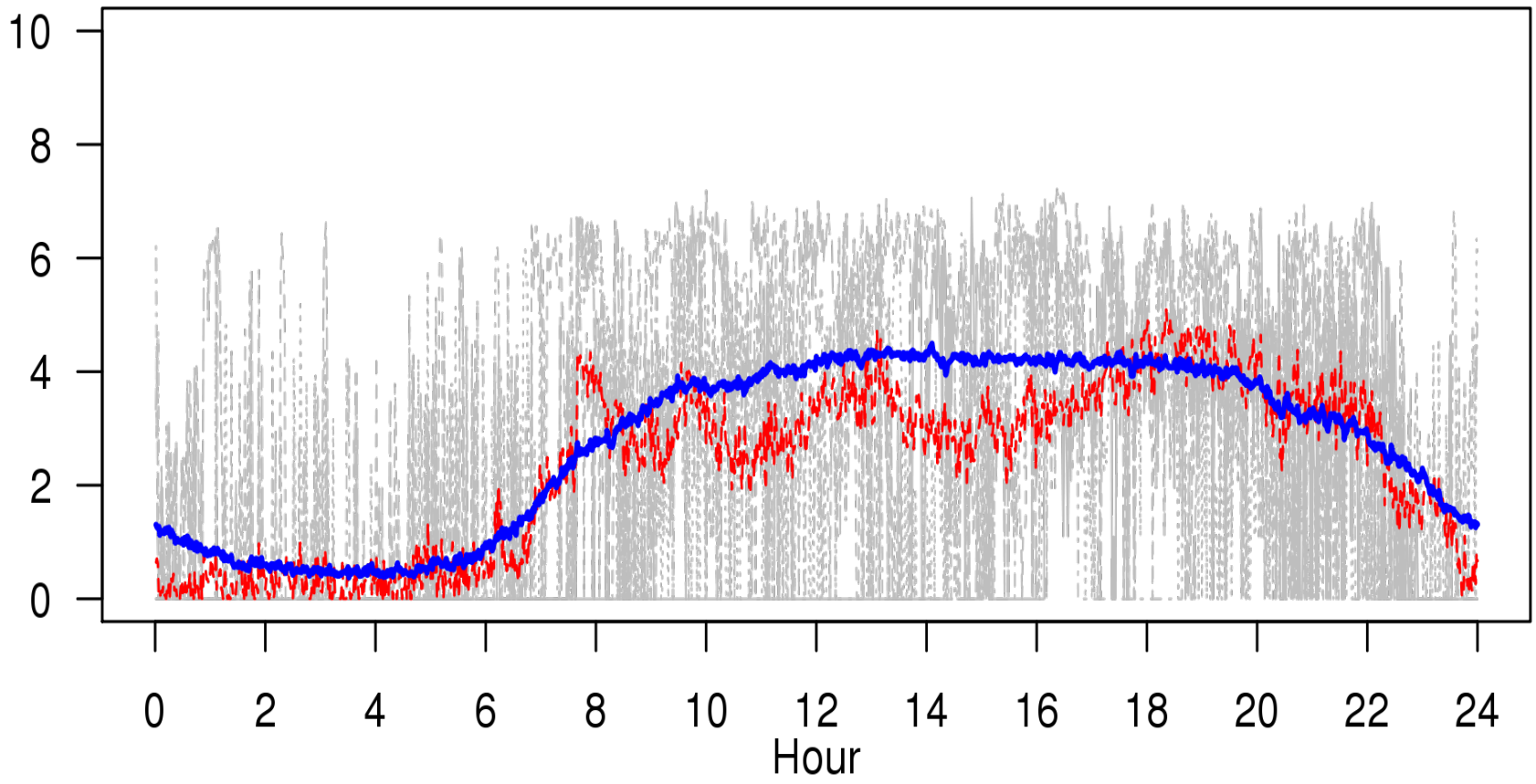
Activity intensity (counts, steps, vector magnitude)



Daily patterns of activity counts



Data: one subject + subject mean + group mean



Baltimore Longitudinal Study of Aging (BLSA)

WIT: organized the BLSA data to the 1440+ standard

- Subjects : 773 (394 females, 379 males): i
- Average number of days/subjects : 7 : j
- Daily profile :1440 minutes : t
- Age : between 31 and 96 : x
- Data set : 5478 by 1440

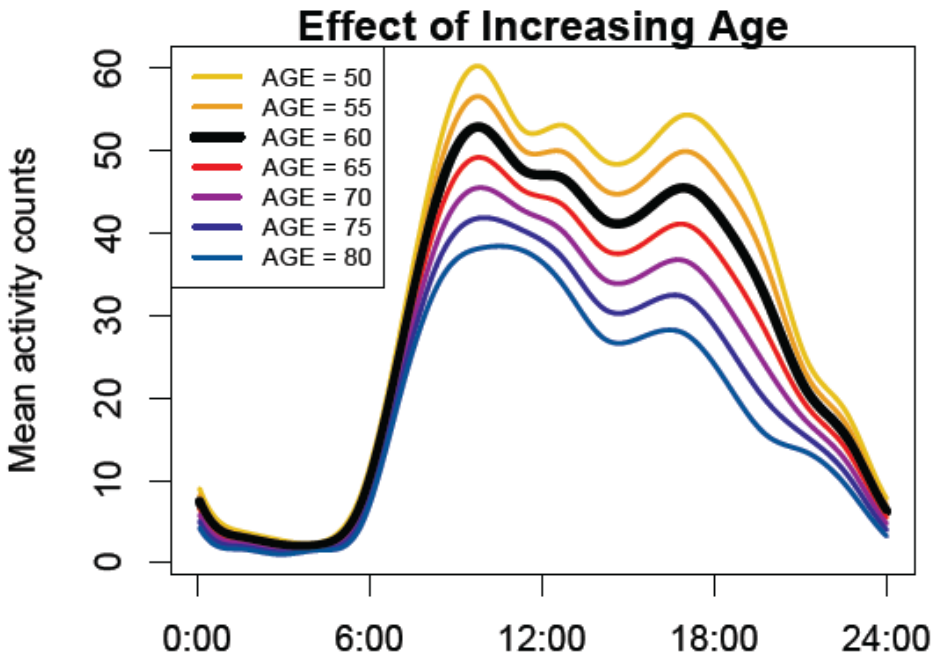
A macro level of the activity data

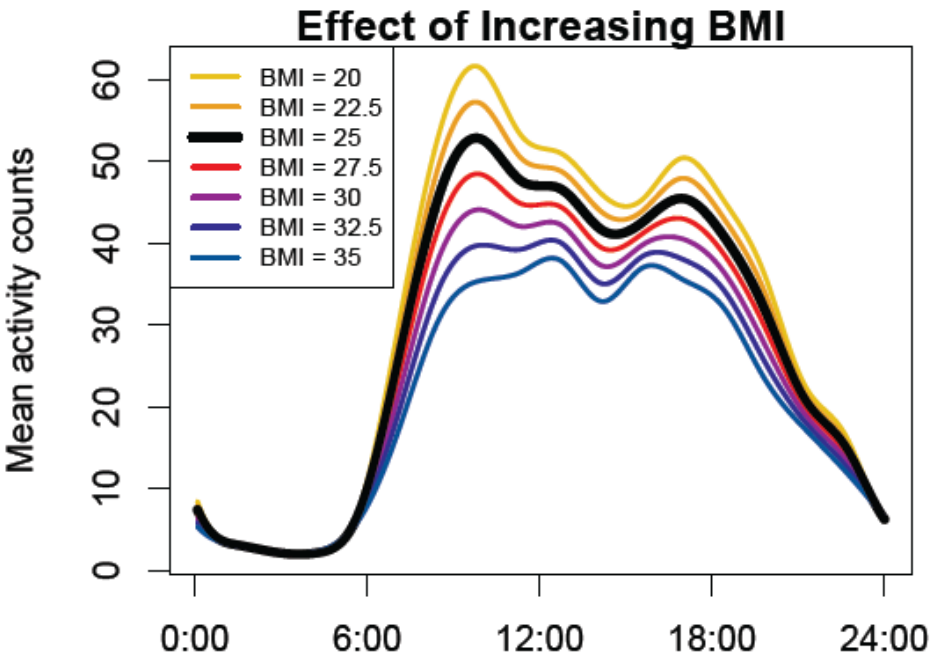
- $Y_{ij}(t)$ = "activity counts" for subject i, on day j, at minute t
- Interested in the time varying effect of age and BMI on activity

$$Y_{ij}(t) = age_i\, \beta(t) + BMI_i\, \gamma(t) + W_{ij}(t)$$

- Use penalized splines to fit $\beta(t)$, $\gamma(t)$
- Account for functional correlation within subjects
- For inference
  - bootstrap of subjects
  - structured functional decompositions (e.g. MFPCA, SFPCA)

Structured-function-on-scalar regression





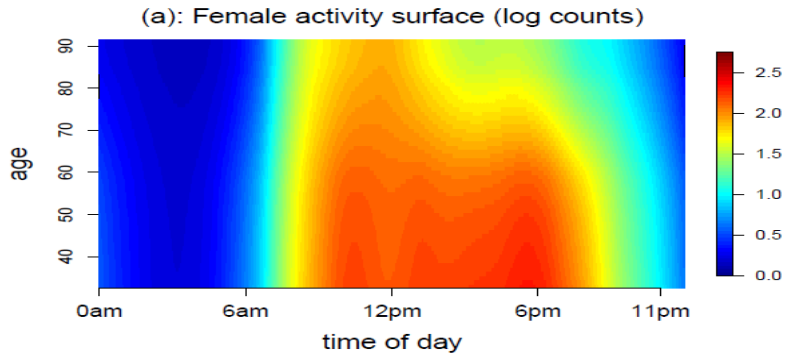
Generalized Multilevel Function-on-Scalar Regression and Principal Component Analysis (2014), Goldsmith, Zipunnikov, Schrack, Biometrics

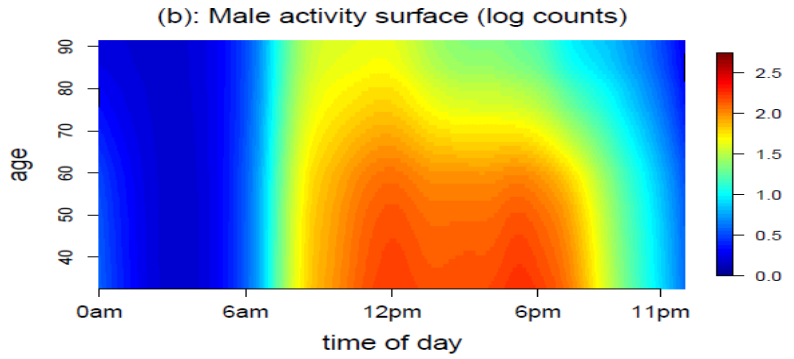
High dimensional bi/tri-variate smoothing (BLSA)

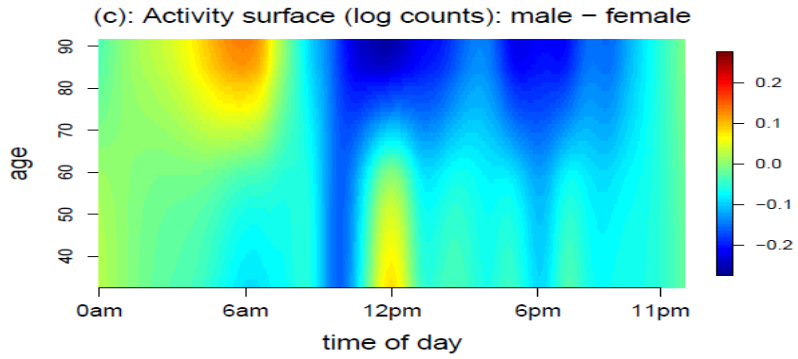
$$Y_{ij}(t)=m(t,x_i)+U_i(t,x_i)+V_{ij}(t,x_i)+\varepsilon_{ij}(t)$$

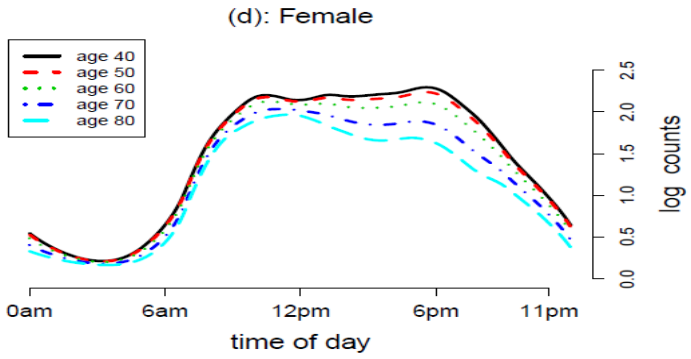
- Requires:
  - fast new smoothers (Luo Xiao's penalty)
  - leave-one-subject-out CV (one-time data pass)

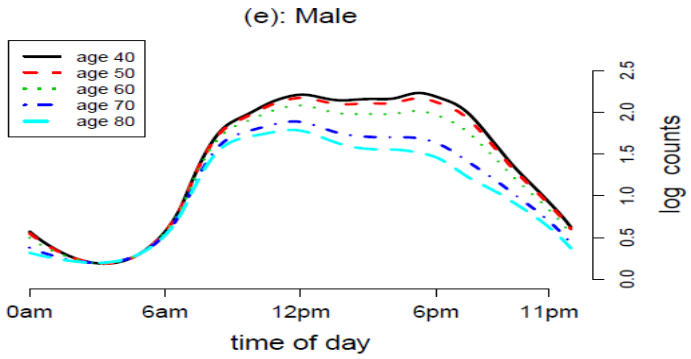
BLSA

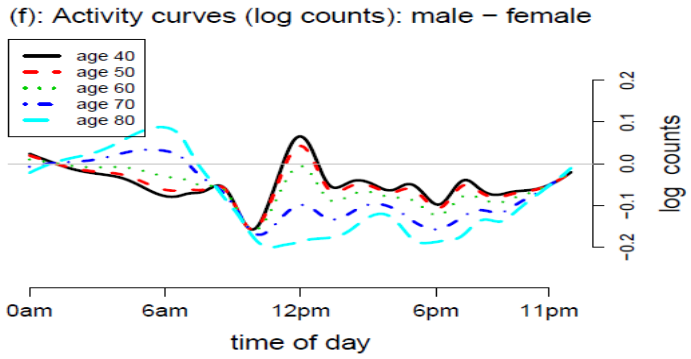

(a): Female activity surface (log counts)
age
time of day
0am
6am
12pm
6pm
11pm


(b): Male activity surface (log counts)
age
time of day
0am
6am
12pm
6pm
11pm

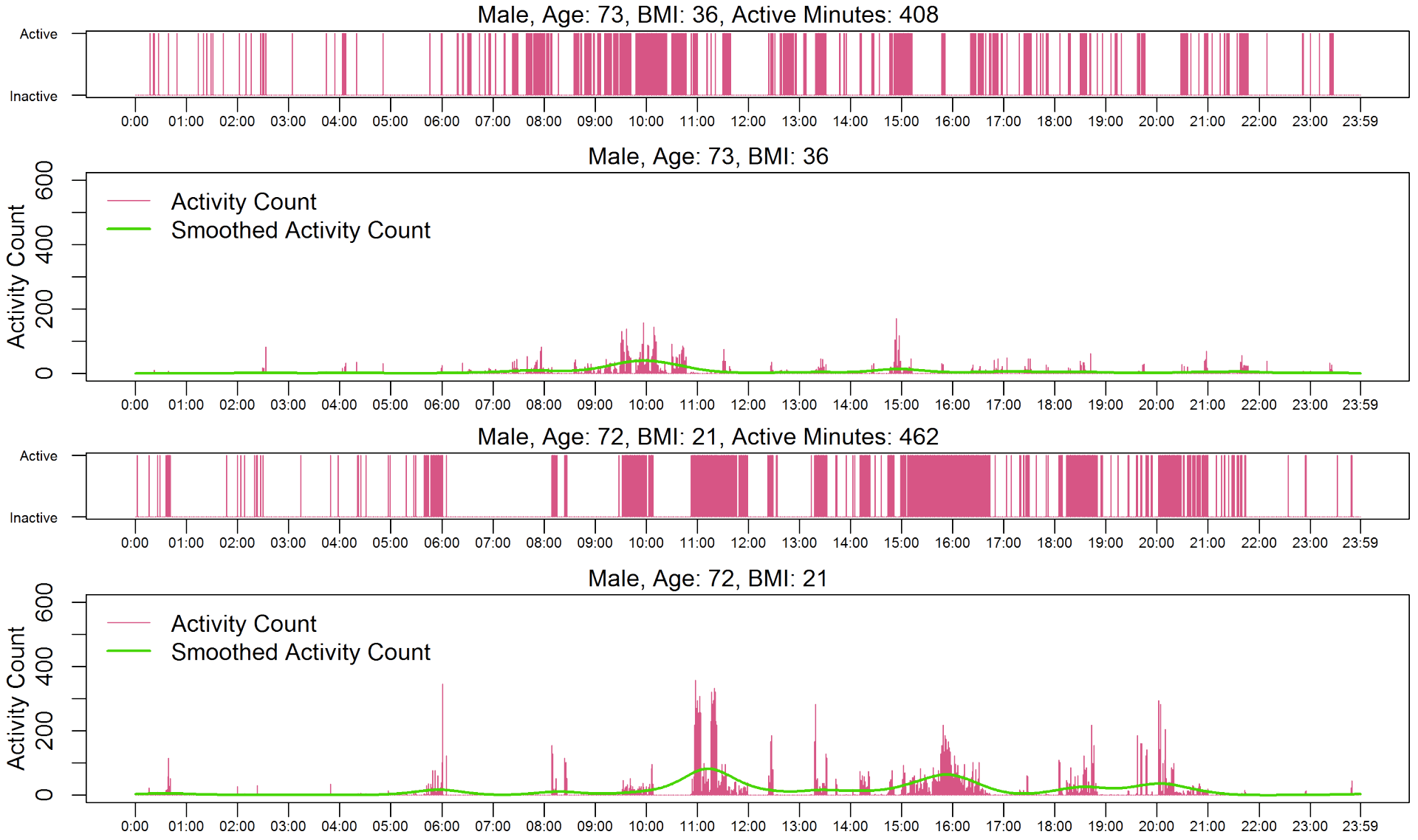

(c): Activity surface (log counts): male − female
age
time of day
0am
6am
12pm
6pm
11pm


(d): Female
age 40
age 50
age 60
age 70
age 80
log counts
time of day
0am
6am
12pm
6pm
11pm

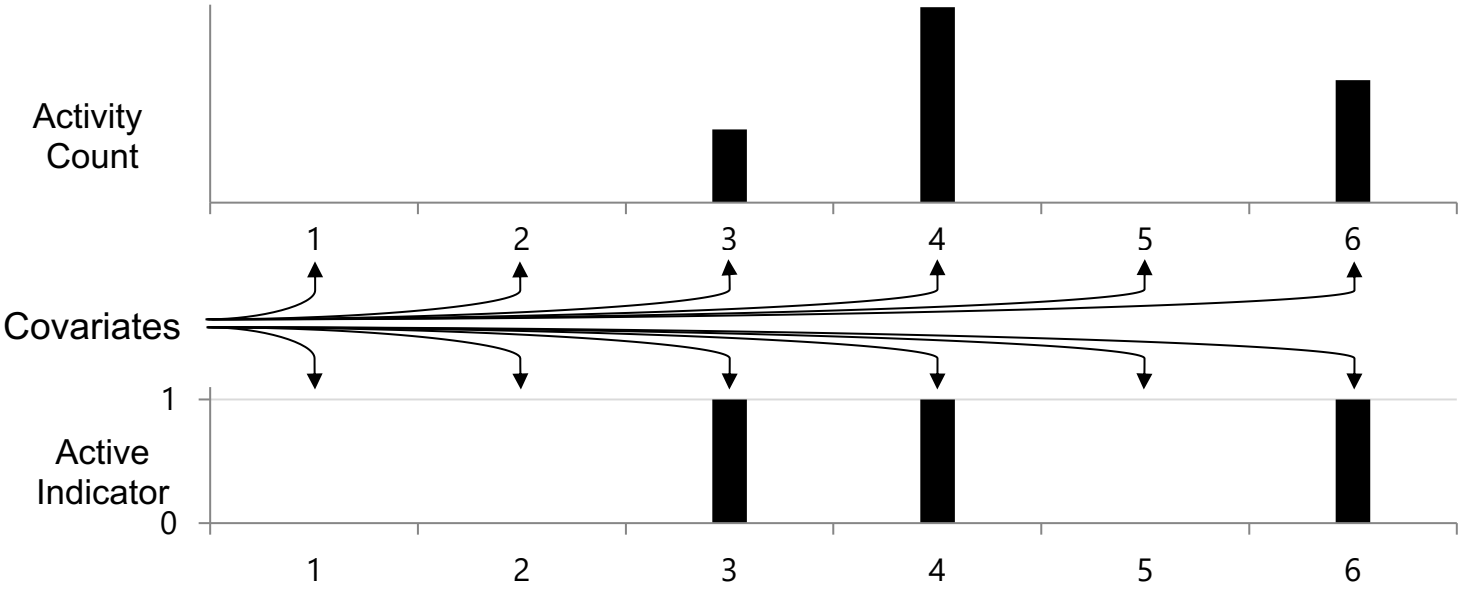

(e): Male
age 40
age 50
age 60
age 70
age 80
log counts
time of day
0am
6am
12pm
6pm
11pm


(f): Activity curves (log counts): male − female
age 40
age 50
age 60
age 70
age 80
log counts
time of day
0am
6am
12pm
6pm
11pm

Dynamic modeling (active/inactive & activity intensity)



Conceptual two-stage dynamic model

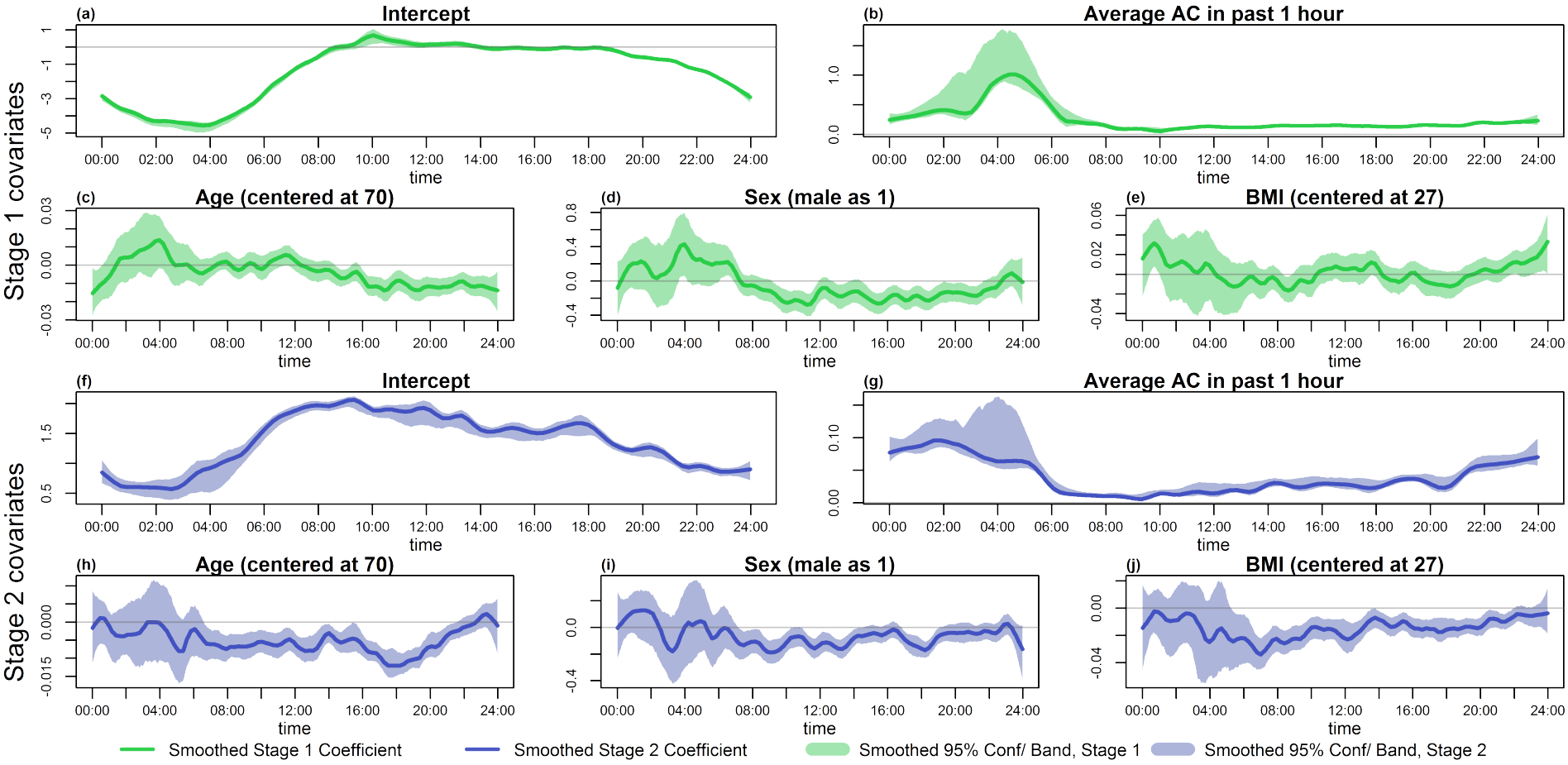


Two-stage dynamic model

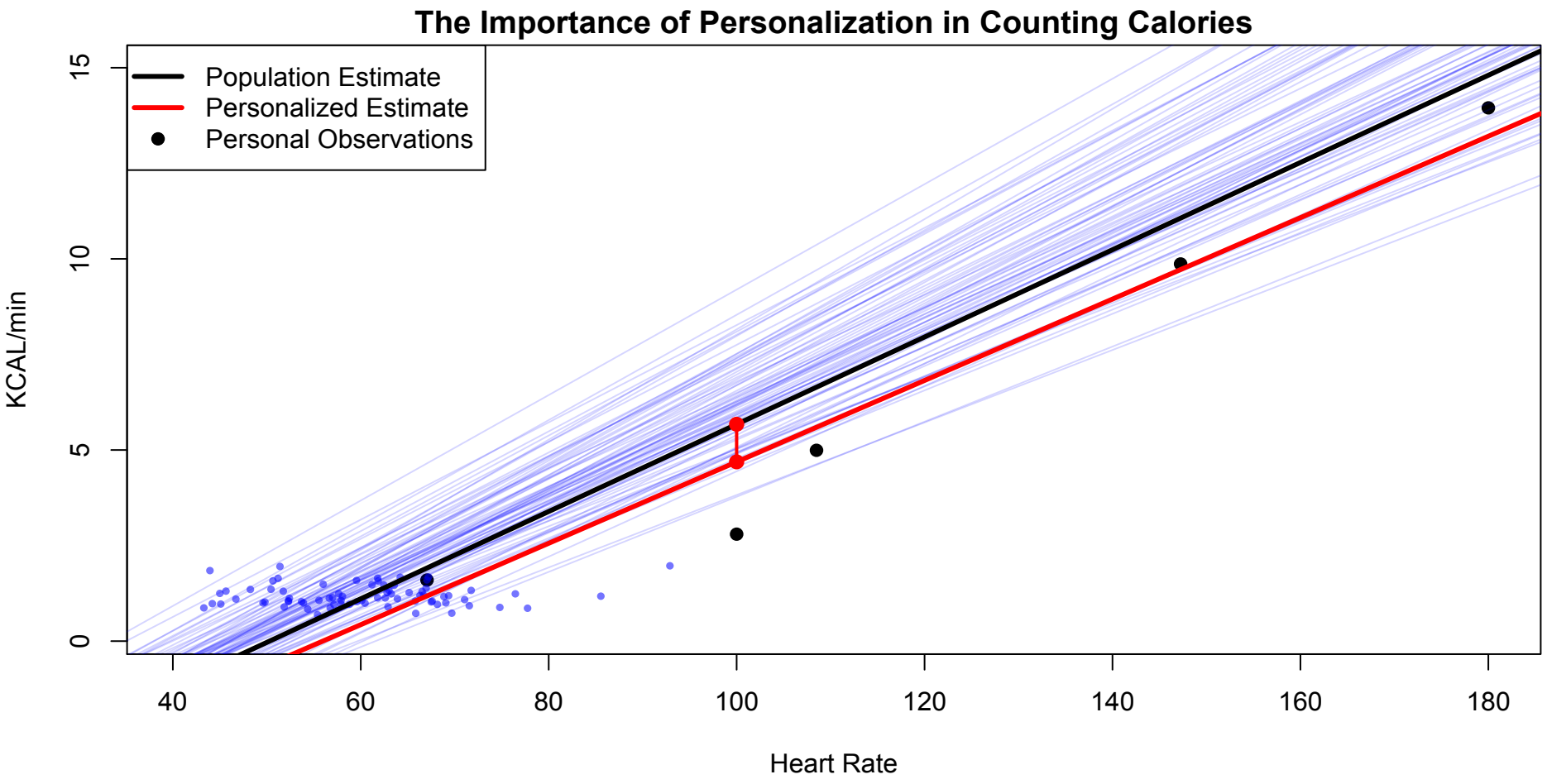
$$\text{logit}[P\{A_i(t)=1\}] = \beta_0(t) + H_i^T(t)\,\beta_1(t)$$

$$\log\{Y_i(t) \mid A_i(t)=1\} = \gamma_0(t) + H_i^T(t)\,\gamma_1(t) + \varepsilon_i(t)$$

Some results



Do not trust your calorie counter



500 Calories	750 Calories	1000 Calories
43%	23%	9%