

BIG DATA INSTITUTE

Li Ka Shing Centre for Health Information and Discovery



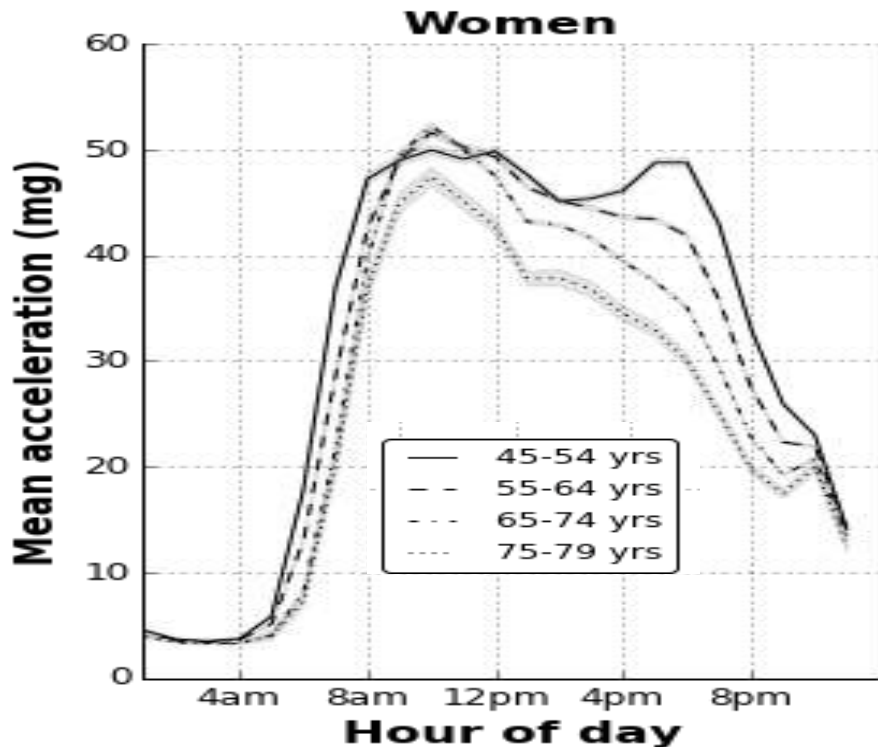
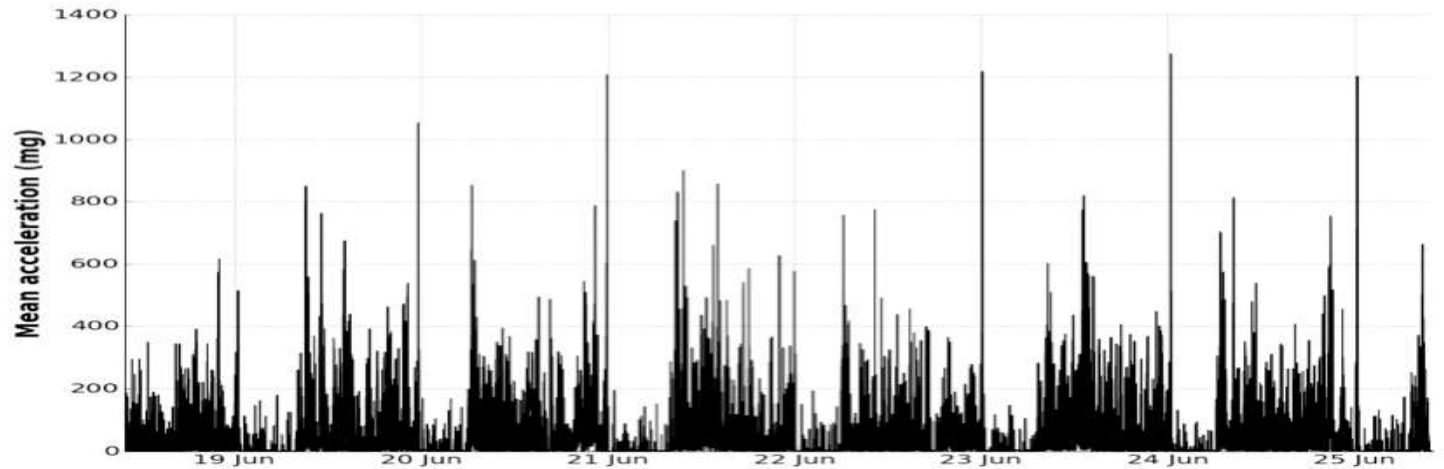
Statistical machine learning of sleep duration and physical activity

Aiden Doherty

*Nuffield Department of Population Health
NIHR Oxford Biomedical Research Centre*

With thanks to: Cecilia Lindgren, Chris Holmes, Michael Holmes, Sara Pulit, Teresa Ferreira, Karl Smith-Byrne, Matthew Willetts, Louis Aslett, Nick Wareham, Soren Brage, Rory Collins, Gil McVean, Martin Landray, and many others...

UK Biobank accelerometer resource

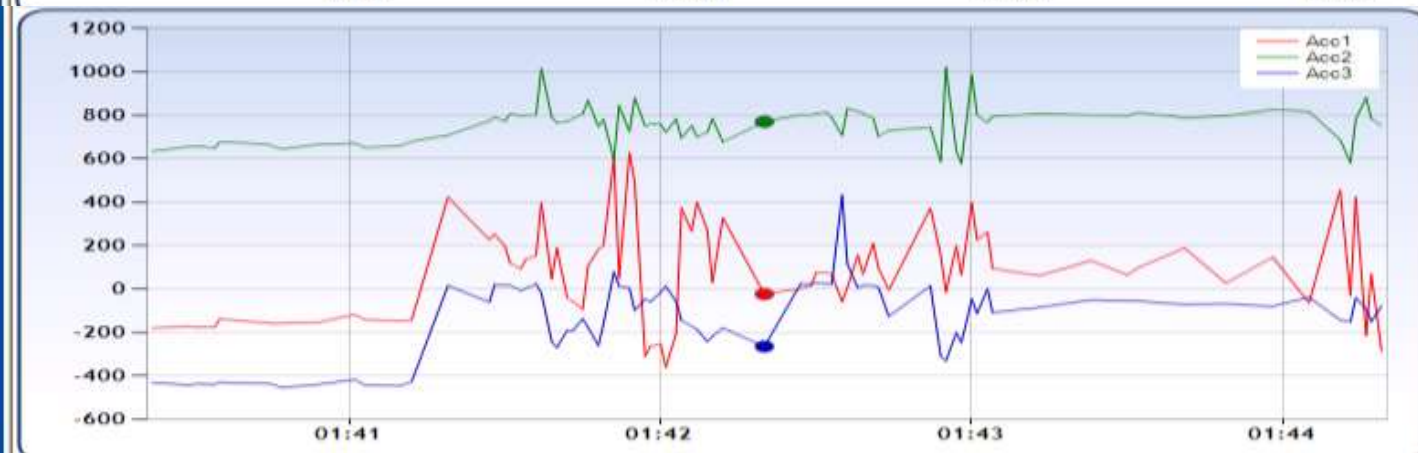


- Physical activity associated with obesity and incident vascular disease
- Circadian rhythm associated with mood disorders

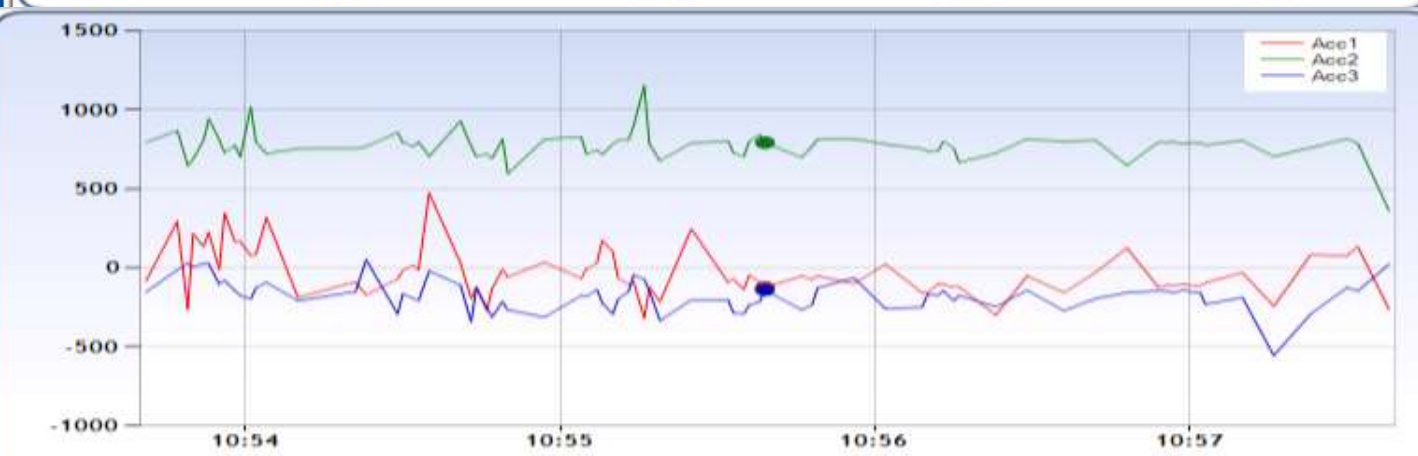
Machine learning & physical activity



Sitting or Standing



Walking



Driving

Wearable camera video

<https://ajpmonline.wordpress.com/2013/04/15/using-wearable-cameras-in-your-research/>

<http://www.sciencedirect.com/science/article/pii/S0926641007493797/44/3>

<https://www.youtube.com/watch?v=KJ2g>



Machine learning of behaviours from sensor data



150 people – activity monitors + cameras

Behaviour – machine learning results

- 159,504 mins annotated behaviour from 132 people
- Kappa = 0.79 (accuracy = 87%)

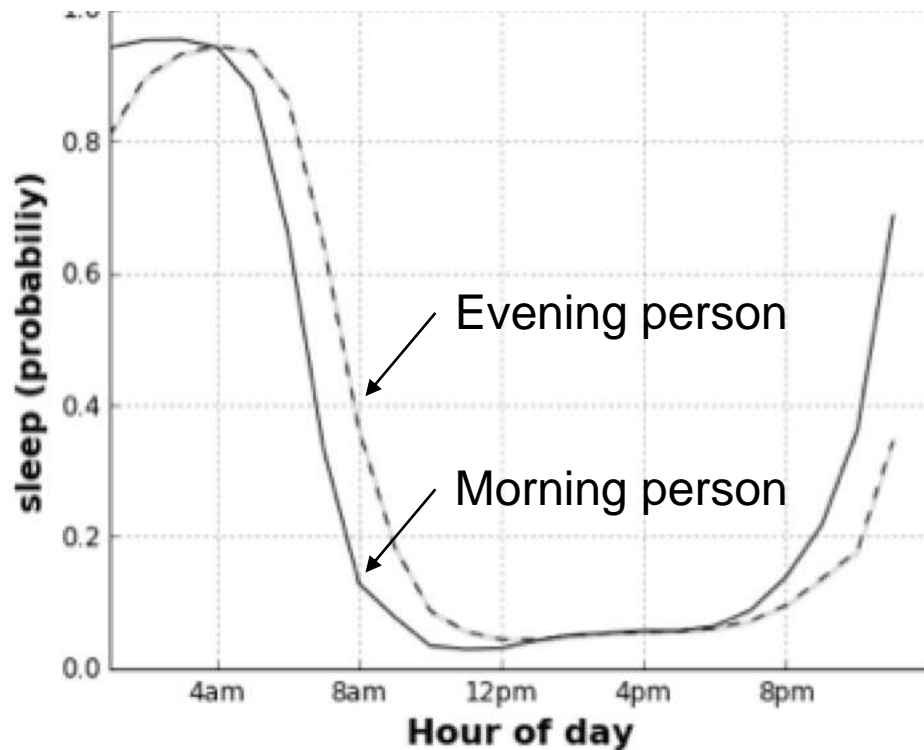
Table 1. Percentage of machine-learned behaviours automatically classified from wrist-worn accelerometer data. Confusion matrix after leave-one-out validation on 84,616 labelled minutes of human activity in free-living environments: the CAPTURE-24 study 2014-2015 (n = 132).

Prediction→ Ground truth↓	Sleep	Sit/stand	Vehicle	Walking	Mixed- activity	Bicycling
Sleep	97%	3%	<1%	<1%	1%	<1%
Sit/stand	3%	89%	1%	3%	3%	<1%
Vehicle	<1%	13%	74%	3%	9%	<1%
Walking	1%	11%	2%	71%	15%	1%
Mixed-activity	1%	20%	2%	19%	57%	1%
Bicycling	1%	1%	1%	12%	14%	71%

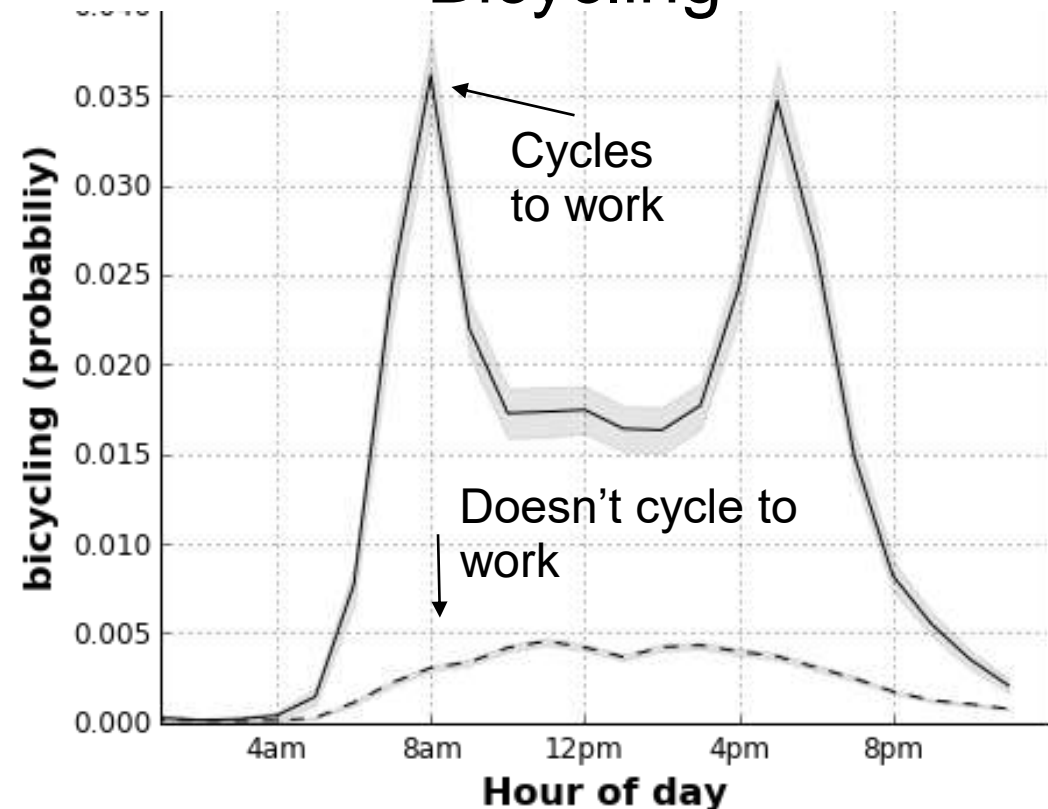
Face validity of behaviour model

Variation in accelerometer-measured behaviour types (2013–2015) across the day by participant characteristics (measured 2007–2010): the UK Biobank study.

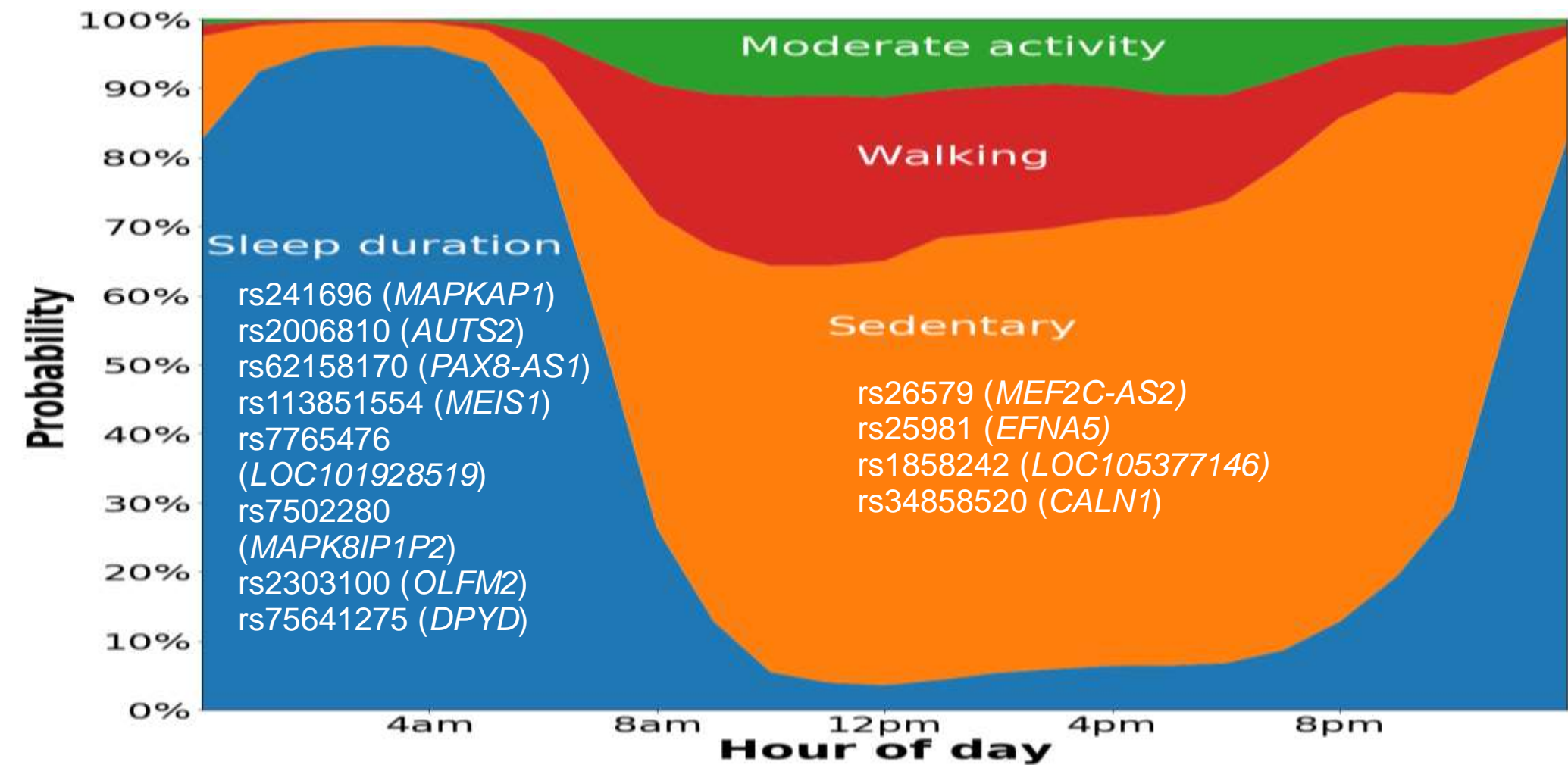
Sleep



Bicycling

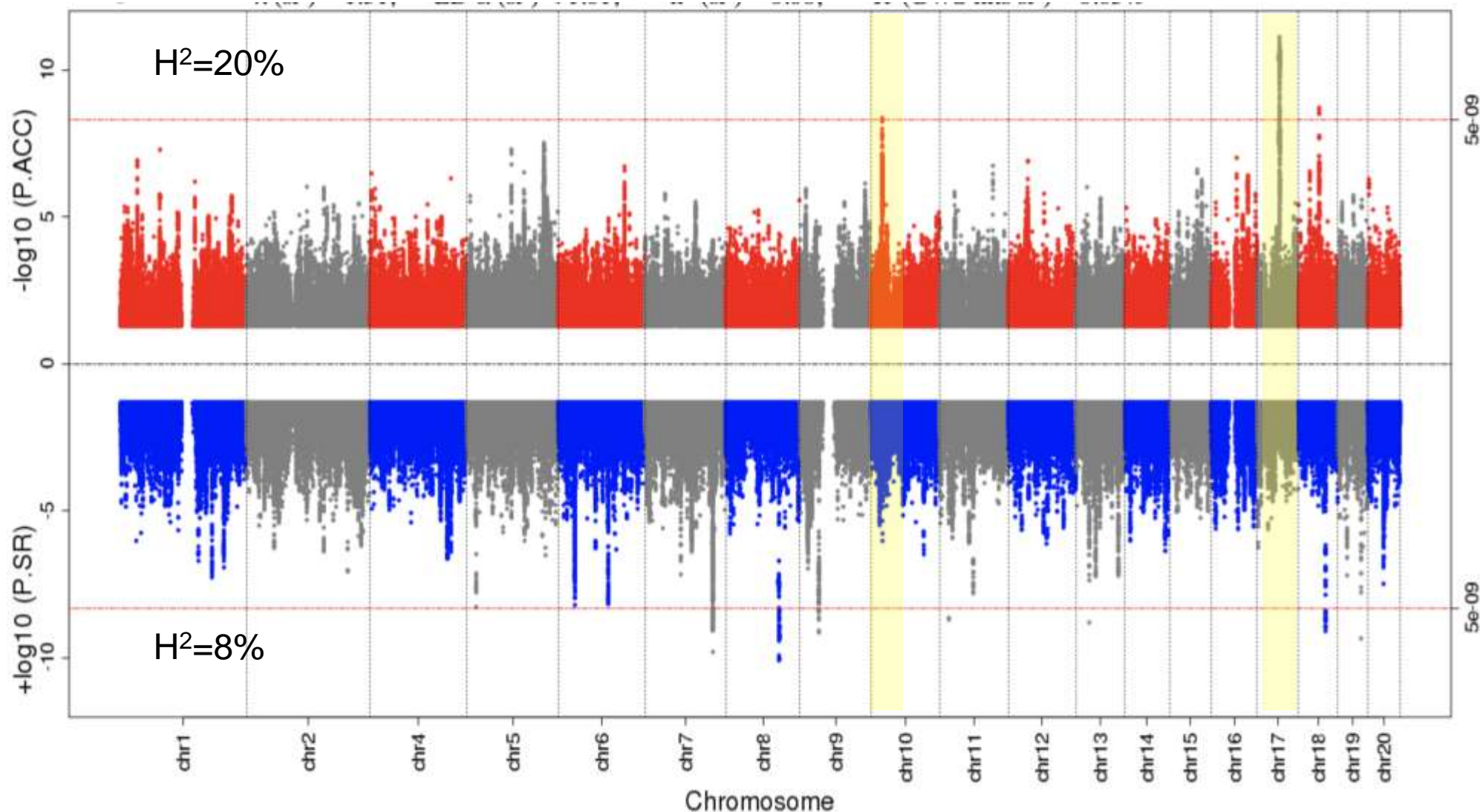


Objectively measured activity behaviours in UK Biobank (n=91,105)

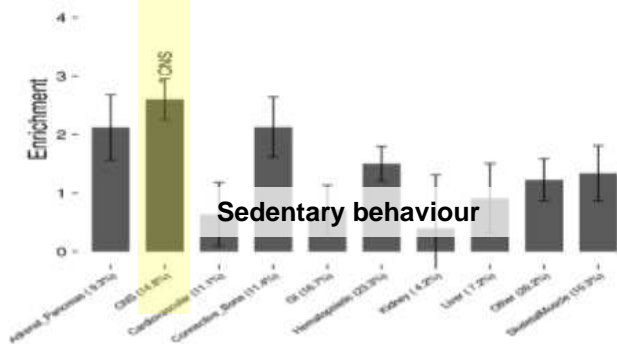
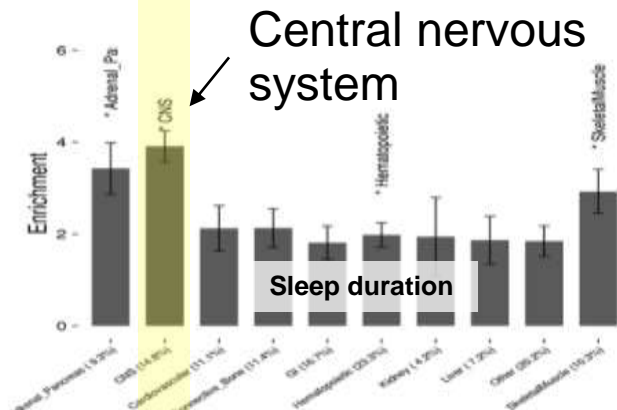
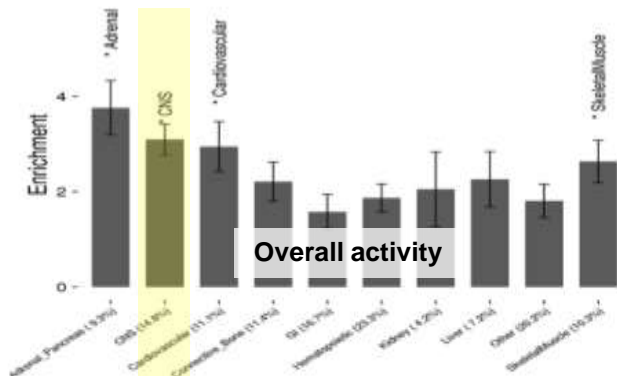


Precise measures of phenotype complement traditional methods

Summary Miami plot of accelerometer (n=91k) and self-report (n=351k) measures of physical activity in UK Biobank.



Biological relevance of accelerometer measures



Mendelian randomization (MR) suggests that increased overall activity reduces:

- diastolic blood pressure

(β /SD: -0.95 , $SE=0.19$, $p=9.5 \times 10^{-7}$)

- odds of hypertension

(OR/SD: 0.85 , $SE=0.03$, $p=4.5 \times 10^{-7}$)

MR suggests longer sleep duration increases:

- odds of hypertension

(OR/SD: 1.13 , $SE=0.03$, $p=7.8 \times 10^{-5}$)

UK Biobank accelerometer resource

- Summary metrics of physical activity intensity available for all researchers
- Sleep duration, sedentary behaviour, and walking phenotypes will soon be available
- Exciting possibilities to further enhance the resource via semi- and unsupervised statistical machine learning

With thanks to: Cecilia Lindgren, Chris Holmes, Michael Holmes, Sara Pulit, Teresa Ferreira, Karl Smith-Byrne, Matthew Willetts, Louis Aslett, Nick Wareham, Soren Brage, Rory Collins, Gil McVean, Martin Landray, and many others...

