Large scale population assessment of physical activity using wrist worn accelerometers: the UK Biobank study

Aiden Doherty[1,2]*, Dan Jackson[3], Nils Hammerla[3], Thomas Plötz[3], Patrick Olivier[3], Malcolm H Granat[4], Tom White[5], Vincent T van Hees[6], Michael I Trenell[6], Christoper G Owen[7], Stephen J. Preece[4], Rob Gillions[8], Simon Sheard[8], Tim Peakman[8], Soren Brage[5†], Nicholas J Wareham[5†]

[1]Big Data Institute, Nuffield Department of Population Health, BHF Centre of Research Excellence, University of Oxford, U.K. [2]Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, U.K. [3]Open Lab, Newcastle University, U.K. [4]School of Health Sciences, University of Salford, Manchester, U.K. [5]MRC Epidemiology Unit, University of Cambridge, U.K. [6]MoveLab, Institute of Cellular Medicine, Newcastle University, U.K. [7]Population Health Research Institute, St George's University of London, U.K. [8]UK Biobank, Stockport, U.K.

* Corresponding author. E-mail: aiden.doherty@ndph.ox.ac.uk

† Joint last authors.

ABSTRACT

Background: Physical activity has not been objectively measured in prospective cohorts with sufficiently large numbers to reliably detect associations with multiple health outcomes. Technological advances now make this possible. We describe the feasibility of accelerometer measured physical activity in over 100,000 participants of the UK Biobank study, and its variation by age, sex, day, time of day, and season.

Methods: Participants were approached by email to wear a wrist-worn accelerometer for seven days that was posted to them. Physical activity information was extracted from 100Hz raw triaxial acceleration data after calibration, removal of gravity and sensor noise, and identification of wear / non-wear episodes. We report age- and sex-specific wear-time compliance and accelerometer measured physical activity, overall and by hour-of-day, week-weekend day and season.

Results: 103,712 datasets were received (44.8% response), with a median wear-time of 6.9 days (IQR:6.5–7.0). 96,600 participants (93.3%) provided valid data for physical activity analyses. Vector magnitude, a proxy for overall physical activity, was 7.5% (2.35m*g*) lower per decade of age (Cohen's d=0.9). Women had a higher vector magnitude than men, apart from those aged 45-54yrs. There were major differences in vector magnitude by time of day (d=0.66). Vector magnitude differences between week and weekend days (d=0.12 for men, d=0.09 for women) and between seasons (d=0.27 for men, d=0.15 for women) were small.

Conclusions: It is feasible to collect and analyse objective physical activity data in large studies. The summary measure of overall physical activity is lower in older participants and age-related differences in activity are most prominent in the afternoon and evening. This work lays the foundation for studies of physical activity and its health consequences. Our summary variables are part of the UK Biobank dataset and can be used by researchers as exposures, confounding factors or outcome variables in future analyses.

INTRODUCTION

Low physical activity is associated with an increased risk of morbidity and mortality [1]. However previous studies are predominantly based on self-reported participation in leisure time activity [2] from which it is difficult to quantify total physical activity across different domains [3]. This uncertainty makes it difficult to convert epidemiological association results into public health recommendations about the minimum level of physical activity required for health and the benefits of engaging in different durations of activity of different intensity. The development of objective methods for assessing physical activity has provided an opportunity to quantify the dose-response relationship of activity with health as a complement to the subjective assessment of self-reported participation in specific activities.

Accelerometry is the most widely used method for objective assessment of physical activity in population studies [4], and large studies from the UK [5,6], US [7], and Canada [8] indicate age gradients and differences between men and women; time-of-day and day-of-week differences in physical activity. These studies all used accelerometers which were worn around the waist and during awake-time only, a protocol which can result in relatively large amounts of missing data [9]. Therefore, wrist-worn accelerometers are becoming more widely used as an objective measure of physical activity in cohorts in the UK [10], US [9], and Brazil [11]. These devices are water-proof and worn continuously day and night, resulting in higher levels of participant compliance [9,10].

Cohort studies which include hundreds of thousands of participants followed up over time are required in order to describe the relationship between physical activity and health outcomes that have a number of potential lifestyle, environmental, and genomic causes [12]. Objective assessment of physical activity in such large population-based cohorts has previously not been undertaken because of the challenges of cost and the feasibility of collecting, processing and analysing data on this large scale. In this paper we demonstrate the feasibility of assessing physical activity by wrist-worn accelerometry in the UK Biobank cohort study and report the variation in activity in more than 100,000 participants by age and sex.

METHODS

Study Population

UK Biobank is a large prospective study with 500,000 participants aged 40-69 years when recruited in 2006-2010 [12]. The study has collected, and continues to collect, extensive phenotypic and genotypic detail about its participants, with ongoing longitudinal follow-up for a wide range of health-related outcomes. Only de-identified data are provided to researchers, who must sign a material transfer agreement, undertaking not to attempt to identify any participant, to keep the data secure, and to use it only for the purposes of the approved research [12]. Between February 2013 and December 2015, participants who had provided a valid email address were sent an email invitation to wear an accelerometer for seven days. The participant email addresses were chosen randomly, with the exception of the North West region which was excluded for much of the project due to participant burden concerns, as this area had been used to trial new projects. From June 2013, participants were sent devices in order of acceptance. This study was covered by the general ethical approval for UK Biobank studies from the NHS National Research Ethics Service on 17th June 2011 (Ref 11/NW/0382). None of the authors had direct contact with the study participants.

Accelerometer & Data Collection

For objective assessment of physical activity, we used the Axivity AX3 wrist-worn triaxial accelerometer (see **Fig 1**), a commercial version of the Open Movement AX3 open source sensor (https://github.com/digitalinteraction/openmovement) designed by Open Lab, Newcastle University. This device demonstrated equivalent signal vector magnitude output on multi-axis shaking tests [13] to the GENEActiv accelerometer used in the Whitehall II [10], Fenland [14] and Pelotas cohorts [11]. The Axivity device facilitates transparent data processing analysis due to its open-source firmware platform and unforced sampling of raw measurement data. We set up the Axivity accelerometers to start at 10am two working days after postal dispatch, and capture triaxial acceleration data over a seven day period at 100Hz with a dynamic range of +-8*g*.

Fig 1. UK Biobank triaxial accelerometer and processing steps to extract physical activity information.

Axivity AX3 triaxial accelerometer worn on dominant hand as used in UK Biobank (top left). Time series trace of processed accelerometer values after one week of wear (top right). Overview of process to extract proxy physical activity information from raw accelerometer data (bottom).

Participants were informed in the invitation email and device mail-out letter that the accelerometer should be worn continuously and that they should carry on with their normal activities. Participants were asked to start wearing the accelerometer immediately after receiving it in the post and to wear the monitor on their dominant wrist. They were also informed that the device was configured to automatically turn itself on soon after its arrival and off seven days later. Finally, participants were asked to mail the device back to the co-ordinating centre, in a pre-paid envelope, after the seven day monitoring period.

Data Processing

To ensure different devices provided a similar output under similar conditions we calibrated the acceleration signals to local gravity using the procedure described by van Hees and colleagues [14]. Briefly, we identified stationary periods in ten second windows where all three axes had a standard deviation of less than 13.0 m*g.* These stationary periods were then used to optimise the gain and offset for each axis (9 parameters) to fit a unit gravity sphere using ordinary least squares linear regression. If insufficient data were available to conduct calibration for a given participant (where any of the three sensor axes did not have values outside a +- 300 m*g* range), we used the calibration coefficients from the previous (or if unavailable, the next) accelerometer record from the same device worn by a different participant. Clipped values, which occur when the sensor's dynamic range of +-8*g* is exceeded, were flagged before and after calibration. Recording errors and 'interrupts', which could have occurred for example if participants tried to plug their accelerometer device into a computer, were also logged. Valid data were then resampled to 100 Hz using linear interpolation, except for interrupts lasting longer than 5 seconds which were set to missing. We calculated the sample level Euclidean norm of the acceleration in x/y/z axes, and removed machine noise using a fourth order Butterworth low pass filter with a cutoff frequency of 20Hz. In order to separate out the activity-related component of the acceleration signal, we removed one gravitational unit from the vector magnitude, with remaining negative values truncated to zero [10,11].

To describe the overall level and distribution of physical activity intensity, we combined the sample level data into five second epochs for summary data analysis, maintaining the average vector magnitude value over the epoch. To represent the distribution of time spent by an individual in different levels of physical activity intensity, we generated an empirical cumulative distribution function from all available five second epochs [11,15]. We removed non-wear time, defined as consecutive stationary episodes lasting for at least 60 minutes where all three axes had a standard deviation of less than 13.0 m*g* [10,16]. We imputed non-wear data segments using the average of similar time-of-day vector magnitude and intensity distribution data points with one minute granularity on different days of the measurement, as in previous studies [10,16]. This imputation accounts for potential wear time diurnal bias where, for example, if the device was systematically not worn during sleep in an individual, the crude average vector magnitude during wear time would be a biased overestimate of the true average. We then constructed a physical activity outcome variable by averaging all worn and imputed values. Our analysis is freely available and hosted as an open source software project at https://github.com/activityMonitoring/biobankAccelerometerAnalysis

Data Analysis

For process evaluation we generated descriptive statistics on the number of participants and devices used. We recorded the number of participants who had insufficient data for calibration. We also noted the percentage of data recording errors caused by interrupts and clipped values, both before and after calibration. Furthermore, we described the number of participants who provided different amounts of wear time. We then excluded individuals with less than three days (72 hours) of wear data or who did not have wear data in each one-hour period of the 24-hour cycle. We defined this criteria after finding 72 hours of wear were needed to be within 10% of a complete seven day measure (using intraclass correlation coefficients) in missing data simulations on 29,765 participants who had perfect wear time compliance (see S1 – minimum wear time criterion).

Descriptive statistics were used to report device wear time compliance in hours and accelerometer measured physical activity in milli-gravity units (m*g*). Age groups were categorised into decade bands from ages 45-79 years. Age and seasonal (with Spring starting on 1st March) differences in device

wear-time were examined using the Kruskal-Wallis test, while sex differences were examined using the Wilcoxon-Mann Whitney test. Differences in wear-time distribution were examined using the Friedman test for time-of-day (six hour quadrants, e.g. 00:00-05:59, 06:00-11:59, etc.) and Wilcoxon signed ranks test for days (weekdays versus weekend days), within individuals for men and women separately. Mean acceleration vector magnitude differences by age group were investigated using one-way repeated measures ANOVA for time-of-day (six hour quadrants) and days (weekdays versus weekend days), within individuals for men and women separately. Seasonal differences in mean acceleration vector magnitude were investigated using two-way ANOVA between age groups, for men and women separately. We used R to perform all statistical analysis [17]. Given the size of this dataset, almost all of our findings show robust statistical significance ($p<0.001$). We therefore do not report such small p-values.

RESULTS

A total of 236,519 UK Biobank participants were approached, of whom 106,053 agreed to wear a physical activity monitor (44.8%). The median time between each participant being invited to take part and being sent a device was 113 days (IQR: 73 – 137 days). Fig 2 shows that 103,712 datasets were received for data analysis. 123 participants were excluded as they were aged less than 45 years. Eleven participants were excluded from further analysis; eight because the calibration by the preceding or subsequent measurement was not possible due to insufficient data; and three participants due to unreliable device data. A total of 4043 devices were used on a median number of 27 occasions (IQR: 8 – 39). The median time between each device being posted was 17.0 days (IQR: 15.8 – 19.8) with a median of 832 devices (IQR: 629 – 994) posted each week.

Fig 2. Participant flow chart; the UK Biobank study 2013-2015 (n= 103,712).

Calibration of the data to local gravity greatly reduced the error in the assessment of acceleration with the root mean square error of stationary points falling from an average of 81.8 m*g* (95% CI: 81.6 – 82.1) to an average of 2.6 m*g* (95% CI: 2.6 – 2.6). However, 2.9% (n=3049) of participants had insufficient stationary data to inform the calibration. These individual records were calibrated using stationary episodes from the previous (n=2887) or next (n=154) use of the same device by different

participants. The influence of clips (readings beyond the sensor's dynamic range of +-8*g*) before (median: 160, IQR: 62 – 393) and after (median: 169, IQR: 67 – 410) calibration, interrupts (median: 0, IQR: 0 – 0), and errors such as clips or missing readings (median: 200, IQR: 66 – 355) was negligible, with respect to the median of 58.6 million data readings (IQR: 56.0 – 60.1 million).

Fig 3 illustrates that 80.6% of participants wore the device for at least 150 hours out of a scheduled 168 hours. Men wore the device for a median of 166.3 hours (IQR: 157.7 – 168.0) and were slightly more compliant than women who wore the device for a median of 165.6 hours (IQR: 156.7 – 167.0). Table 1 shows that older age groups had marginally higher levels of compliance than younger age groups. Analysis of wear time compliance by age on a linear scale shows that on average there was a difference of 2 hours 18 minutes (1.6%) for each decade. In addition, Table 1 indicates minimal differences in the wear time compliance by time-of-day and week-weekend day. No wear-time differences were found by season. We removed 6978 (6.7%) participants who had insufficient wear data for our remaining analyses on accelerometer measured physical activity.

Table 1. Wear-time compliance and acceleration vector magnitude by age, day, time of day, and season, stratified by sex: The UK Biobank study 2013-2015 (n=103,578).

	Wear time [median (IQR) hours]		**Acceleration vector magnitude [mean +- stdev m*g*]**	
	Women	**Men**	**Women**	**Men**
Age (yrs) [A]				
45-54	164.9 (152.4 – 167.0) (n=12,586)	165.4 (149.5 – 168.0) (n=8655)	31.2 +- 8.7 (n=11,572)	31.1 +- 9.7 (n=7838)
55-64	165.4 (156.0 – 167.0) (n=21,322)	165.8 (156.5 – 168.0) (n=14,410)	29.1 +- 8.0 (n=19,890)	28.8 +- 8.8 (n=13,362)
65-74	165.6 (159.1 – 168.0) (n=22,821)	166.8 (160.8 – 168.0) (n=20,595)	26.6 +- 7.1 (n=21,489)	25.6 +- 7.7 (n=19,385)
75-79	165.6 (158.9 – 167.0) (n=1494)	166.8 (162.6 – 168.0) (n=1695)	23.9 +- 6.5 (n=1436)	22.9 +- 6.8 (n=1628)
p value	p<0.001	p<0.001	p<0.001	p<0.001
Time of day [B]				
0 - 5.59 $_{am}$	40.9 (36.0 – 42.0) (n=58,223)	42.0 (36.6 – 42.0) (n=45,355)	4.4 +- 3.1 (n=54,387)	4.9 +- 4.4 (n=42,213)
6 - 11.59 $_{am}$	41.0 (38.9 – 42.0) (n=58,223)	42.0 (39.0 – 42.0) (n=45,355)	38.6 +- 14.9 (n=54,387)	37.4 +- 16.4 (n=42,213)
12 - 5.59 $_{pm}$	42.0 (40.3 – 42.0) (n=58,223)	42.0 (40.3 – 42.0) (n=45,355)	44.3 +- 13.8 (n=54,387)	42.9 +- 16.0 (n=42,213)
6- 11.59 $_{pm}$	42.0 (39.5 – 42.0) (n=58,223)	42.0 (40.2 – 42.0) (n=45,355)	26.4 +- 10.4 (n=54,387)	24.9 +- 11.5 (n=42,213)
p value	p<0.001	p<0.001	p<0.001	p<0.001
Day [C]				
Weekday	23.7 (22.5 – 24.0)	23.8 (22.6 – 24.0)	28.5 +- 8.2	27.5 +- 9.0

	(n=58,223)	(n=45,355)	(n=54,387)	(n=42,213)
Weekend	24.0 (22.9 – 24.0)	24.0 (23.3 – 24.0)	28.0 +- 9.4	27.1 +- 10.8
	(n=58,223)	(n=45,355)	(n=54,387)	(n=42,213)
p value	p<0.001	p<0.001	p<0.001	p<0.001
Season [D]				
Spring	165.6 (156.2 - 167.5)	166.1 (157.4 - 168.0)	28.8 +- 8.0	28.1 +- 9.1
	(n=13,365)	(n=10,224)	(n=12,480)	(n=9,469)
Summer	165.4 (156.2 - 168.0)	166.3 (157.4 - 168.0)	28.8 +- 8.1	28.2 +- 8.7
	(n=15,450)	(n=11,943)	(n=14,353)	(n=11,016)
Autumn	165.6 (157.2 - 167.0)	166.3 (158.2 - 168.0)	28.3 +- 8.0	27.3 +- 8.7
	(n=17,213)	(n=13,506)	(n=16,157)	(n=12,633)
Winter	165.6 (156.7 - 168.0)	166.3 (157.9 - 168.0)	27.7 +- 7.8	26.3 +- 8.4
	(n=12,195)	(n=9,682)	(n=11,397)	(n=9,095)
p value	p=0.289	p=0.104	p<0.001	p<0.001

[A] *Age:* Kruskal-Wallis test used to compare wear-time distributions, and one-way analysis of variance test used to compare acceleration vector magnitude means. Sum wear time hours for week displayed (max = 168.0).

[B] *Time of day:* Friedman test used to compare wear-time distributions within individuals, and repeated one-way analysis of variance test used to compare acceleration vector magnitude means within individuals and between age groups. Sum wear time hours for time quadrant over a week displayed (max = 168.0).

[C] *Day:* Wilcoxon test used to compare wear-time distributions within individuals, and repeated one-way analysis of variance test used to compare acceleration vector magnitude means within individuals and between age groups. Average wear time hours for day displayed (max = 24.0).

[D] *Season (Spring starting on 1st March):* Kruskal-Wallis test used to compare wear-time distributions, and two-way analysis of variance test used to compare acceleration vector magnitude means between age groups. Sum wear time hours for week displayed (max = 168.0).

Fig 3. Cumulative distribution function of accelerometer wear time compliance; the UK Biobank study 2013-2015 (n= 103,578).

Table 1 describes the variation in mean vector magnitude, the summary measure of accelerometer measured physical activity, by age and sex in the sub-group of 96,600 participants who had good wear time compliance. Vector magnitude was higher in women than men, apart from those aged 45-54 years (p = 0.98). The mean effect size for these sex differences was small (0.09), ranging from 0.01 for 45-54 years to 0.15 for 75-79 years. There was strong evidence of accelerometer measured physical activity differing by age group in both men and women. The mean physical activity in the age group 45-54 years was 31.17 m*g* (SD 9.10) and was, on average 7.5% or 2.35 m*g* lower per decade. The mean effect size for these age differences was large, at 0.89 for women and 0.9 for men. Fig 4 shows the distribution of the data within age and sex strata, highlighting that although there appears to

be an overall decline in average physical activity with increasing age, there is considerable overlap in the distributions with many older participants being more active than those in the youngest age category.

Fig 4. Acceleration vector magnitude by sex and age; the UK Biobank study 2013-2015 (n= 96,600).

Fig 5 shows the mean physical activity level by hour of day averaged across the whole measurement period by age and sex. It shows that the effect size for physical activity differences between age groups are most apparent in the afternoon (0.74 for women and 0.69 for men) and evening (1.06 for women, 1.12 for men) with smaller differences by age group in the morning (0.56 for women, 0.46 for men). Weekdays and weekend days differed, with vector magnitude higher at weekdays except for those aged 45-54 years. However, the mean effect size for these day differences was small (0.10), ranging from 0.04 to 0.15 across female age groups and 0.11 to 0.18 for male age groups (see Fig 6). Seasonality also differed, with vector magnitude lower during winter months except for women aged 75-79. However, the mean effect size for these season differences was small (0.21), ranging from 0.09 to 0.18 across age groups in women and 0.17 to 0.41 across age groups in men (see Fig 6).

Fig 5. Variation in mean acceleration across the day by age and sex: the UK Biobank study 2013-2015 (n= 96,600).

Shading bounds represent two standard errors.

Fig 6. Acceleration vector magnitude by day of the week (top), season (bottom), age, and sex: the UK Biobank study 2013-2015 (n= 96,600).

To illustrate time spent at different physical activity intensities, Fig 7 plots the empirical cumulative distribution function of the five second sample values for each subgroup. The bottom part of this figure shows sex differences in the distribution of physical activity intensity, for each age group. For example, men spend more time at or below 25 m*g* than women (122.6 versus 119.3 hours), but also slightly more time above 225 m*g* than women (2.18 versus 2.09 hours).

Fig 7. Cumulative time spent in various acceleration categories by sex and age (top), and sex differences by age and intensity level (bottom); the UK Biobank study 2013-2015 (n=96,600)

DISCUSSION

Developments in the technology supporting objective assessment of physical activity have now made it possible to consider assessing this behaviour objectively in large scale population-based cohort studies as an adjunct to more traditional assessment of self-reported participation in activities within different domains of life. However, even with those technological developments, it has previously been unknown whether it would be possible for this approach to be acceptable to participants and whether it would prove to be feasible to collect, analyse and interpret data from over one hundred thousand participants. This report from the UK Biobank study shows that 45% of participants who were invited to wear a monitor accepted the invitation. It also shows that measuring activity with a wrist worn device is highly acceptable to participants as manifest by the very high proportion of people in whom the data were of high quality and completeness. By necessity in the UK Biobank Study participants were invited to wear the monitor some time after recruitment to the baseline visit. As with all add-on measurements that are conducted on a different occasion, there will be participants who don't accept the invitation to participate. Other studies in which wrist worn accelerometers are part of the protocol for a baseline visit, rather than a separate add-on, will be likely to achieve higher participation levels.

We have shown that mean vector magnitude in this population was greater in women than men, apart from those aged 45-54 years. Whether these results indicate true differences in physical activity between sexes or are a function of a between-sex difference in the relationship between wrist acceleration and true activity remains to be investigated. In addition, we observed a marked overall difference in the summary measure of physical activity by age, with older participants having levels of activity that are, on average, 7.5% lower for each 10 year age difference. These differences by age group are similar to other population-based studies [5,7] that have used hip worn accelerometers. Our findings also suggest that men spend more time than women in what might be considered low or sedentary levels of physical activity, while women spend more time in moderate levels of activity.

Older participants are much less active than younger participants during afternoons/evenings than in the morning, which mirrors previous findings in older UK adults using hip worn accelerometers [6]. There were small differences between weekday and weekend day physical activity, and also small seasonal differences in activity. We have not generalised the overall descriptive findings to the UK population since the UK Biobank was established as an aetiological study rather than one aimed at population surveillance [5,7].

We have extracted objective physical activity information from 103,578 participants aged 45-79, who were asked to wear accelerometers for seven days on their dominant wrist. The strengths of this study include its use of objective measures of physical activity, excellent participant compliance, unprecedented scale, and use of reproducible methods. For example, >93% of participants provided more than 72 hours of wear time with no missing data bias by time of day. The overall levels of participant compliance in the UK Biobank mirror findings in other studies that have used wrist-worn accelerometers in thousands of participants [9–11]. Wrist-worn accelerometers are not only highly acceptable to participants, but are also valid. Laboratory-based studies have demonstrated that the signal from wrist-worn devices correlates with physical activity energy expenditure as well as traditional waist-worn devices (R=0.83 vs. R=0.87) [18]. However, more robust validation studies of physical activity information from raw wrist-worn accelerometer data are needed to enhance the interpretation of this signal.

Even though we used relatively simple summary measures in these analyses, their derivation still involved several critical data processing decisions, the alteration of which would have large effects on the derived physical activity variables [14]. For example, there is uncertainty on how to address negative values during the gravity removal process. Furthermore, it is not possible to perfectly separate static and dynamic acceleration (for example gravity and physical activity) from the measurement of triaxial acceleration alone. Therefore, we produced summary statistics of vector magnitude which do not attempt this separation. In addition, absolute and truncated Euclidian norm minus one and high-pass filtered vector magnitude (all of which attempt to separate out activity) were generated too. We found that while the magnitude of these variables changes, their correlation was very strong (>0.95), which provides confidence in our chosen metric for association studies.

Uncertainty also exists on the best method to identify non-wear episodes, and the size of epoch on which to base distributions of physical activity intensity. While the derived factors are only the most basic variables that can be extracted from the raw 100Hz triaxial acceleration data, future projects will be able to build on this foundation to derive additional parameters describing other aspects of physical activity, sedentary behaviour and sleep.

In conclusion, the collection and processing of this large accelerometer dataset in a prospective cohort study lays the foundation for studies of physical activity and its health consequences. The summary variables that we have constructed are now part of the UK Biobank dataset and can be used by researchers as exposures, confounding factors or outcome variables in future analyses.

ACKNOWLEDGEMENTS

We would like to thank all UK Biobank participants for agreeing to volunteer in this research. We would also like to acknowledge the use of the University of Oxford Advanced Research Computing (ARC) facility in carrying out this work. http://dx.doi.org/10.5281/zenodo.22558

REFERENCES

1. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet. 2012;380: 219–29. doi:10.1016/S0140-6736(12)61031-9

2. Chief Medical Officers of England, Scotland, Wales, and Northern Ireland. Start Active, Stay Active: A report on physical activity for health from the four home countries. U.K. Department of Health; 2011.

3. Steene-Johannessen J, Anderssen SA, van Der Ploeg HP, Hendriksen IJM, Donnelly AE, Brage S, et al. Are Self-report Measures Able to Define Individuals as Physically Active or Inactive? Med Sci Sport Exerc. 2016;48. Available: http://journals.lww.com/acsm-msse/Fulltext/2016/02000/Are_Self_report_Measures_Able_to_Define.8.aspx

4. Strath SJ, Kaminsky LA, Ainsworth BE, Ekelund U, Freedson PS, Gary RA, et al. Guide to the assessment of physical activity: Clinical and research applications: a scientific statement from the American Heart Association. Circulation. 2013;128: 2259–79. doi:10.1161/01.cir.0000435708.67487.da

5. Craig R, Mindell J, Hirani V. Health Survey for England - 2008: Physical activity and fitness. National Centre for Social Research; 2008.

6. Sartini C, Wannamethee SG, Iliffe S, Morris RW, Ash S, Lennon L, et al. Diurnal patterns of objectively measured physical activity and sedentary behaviour in older men. BMC Public Health. 2015;15: 609. doi:10.1186/s12889-015-1976-y

7. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sport Exerc. 2008;40: 181–188.

8. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. Heal reports. 2011;22: 7–14. Available: http://europepmc.org/abstract/med/21510585

9. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med. 2014;48: 1019–23. doi:10.1136/bjsports-2014-093546

10. Sabia S, van Hees VT, Shipley MJ, Trenell MI, Hagger-Johnson G, Elbaz A, et al. Association Between Questionnaire- and Accelerometer-Assessed Physical Activity: The Role of Sociodemographic Factors. Am J Epidemiol. 2014;179: 781–790. doi:10.1093/aje/kwt330

11. da Silva IC, van Hees VT, Ramires V V, Knuth AG, Bielemann RM, Ekelund U, et al. Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. Int J Epidemiol. 2014;43: 1959–68. doi:10.1093/ije/dyu203

12. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. PLoS Med. Public Library of Science; 2015;12: e1001779. doi:10.1371/journal.pmed.1001779

13. Ladha C, Ladha K, Jackson D, Olivier P. Shaker Table Validation Of Openmovement Ax3 Accelerometer. 3rd International Conference on Ambulatory Monitoring of Physical Activity and Movement. Amherst, MA, USA; 2013. pp. 69–70. Available: http://www.umass.edu/sphhs/sites/default/files/ICAMPAM Poster Session Abstracts 4-24.pdf

14. van Hees VT, Fang Z, Langford J, Assah F, Mohammad A, da Silva ICM, et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. J Appl Physiol. 2014;117: 738–44. doi:10.1152/japplphysiol.00421.2014

15. Hammerla NY, Kirkham R, Andras P, Ploetz T. On preserving statistical characteristics of accelerometry data using their empirical cumulative distribution. Proceedings of the 17th annual international symposium on International symposium on wearable computers - ISWC '13. New York, New York, USA: ACM Press; 2013. p. 65. doi:10.1145/2493988.2494353

16. van Hees VT, Renström F, Wright A, Gradmark A, Catt M, Chen KY, et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. PLoS One. 2011;6: e22922. doi:10.1371/journal.pone.0022922

17. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria; 2016. Available: https://www.r-project.org/

18. Esliger DW, Rowlands A V, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA Accelerometer. Med Sci Sport Exerc. 2011;43: 1085–1093.

402

403 **SUPPORTING INFORMATION**

404

405 **S1 – Minimum wear time criterion**

fig1

Click here to download Figure fig1.tiff

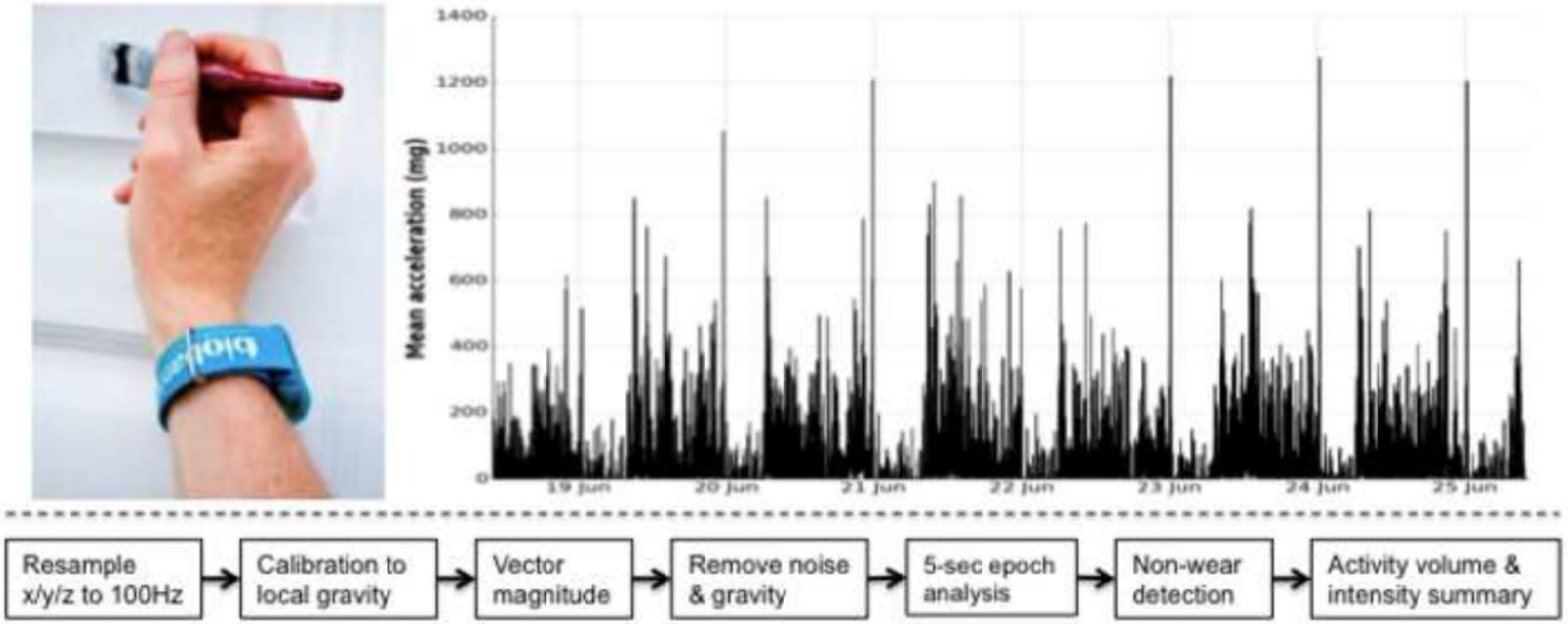

fig2

Click here to download Figure fig2.tiff

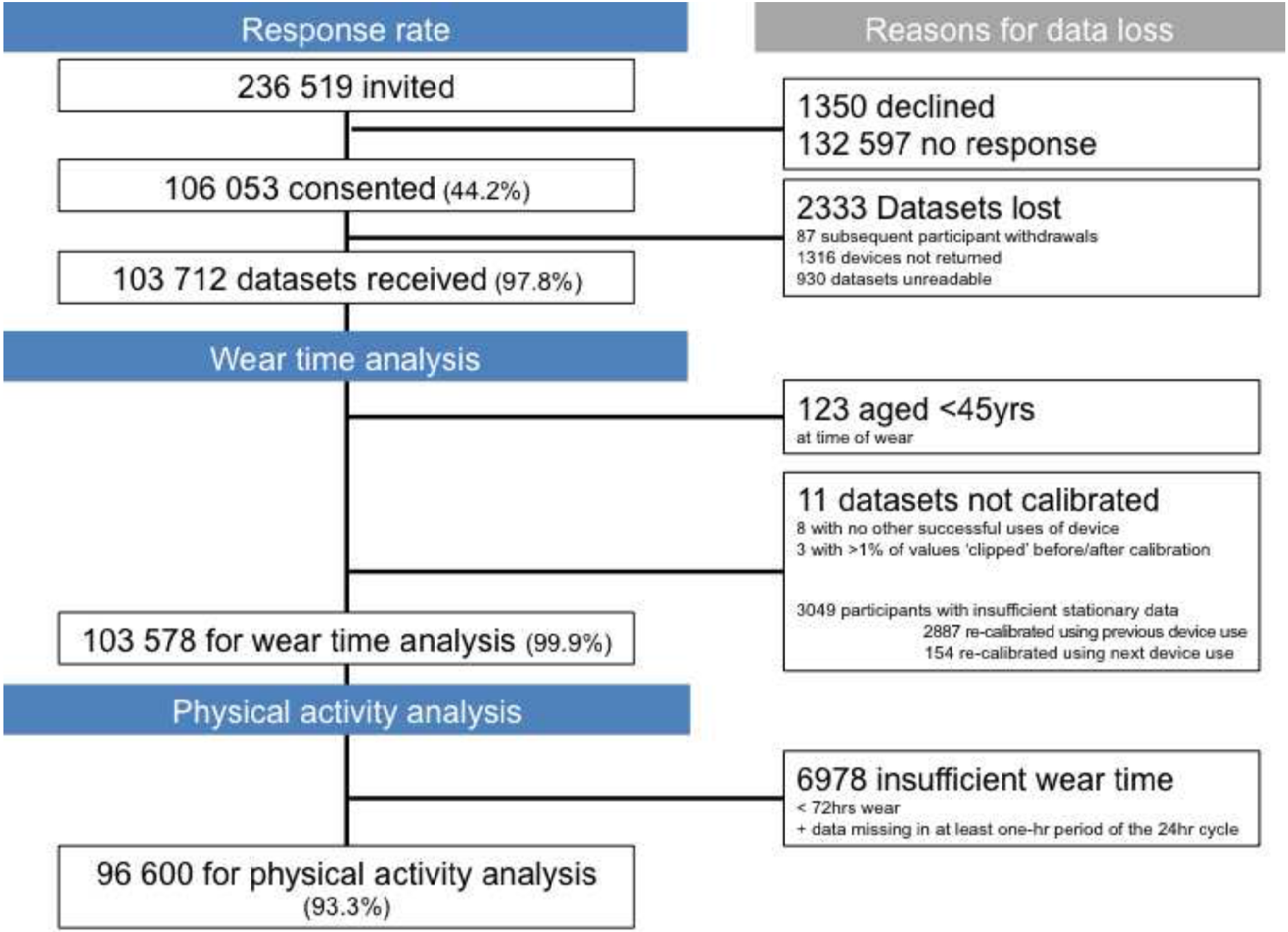

fig3

Click here to download Figure fig3.tiff

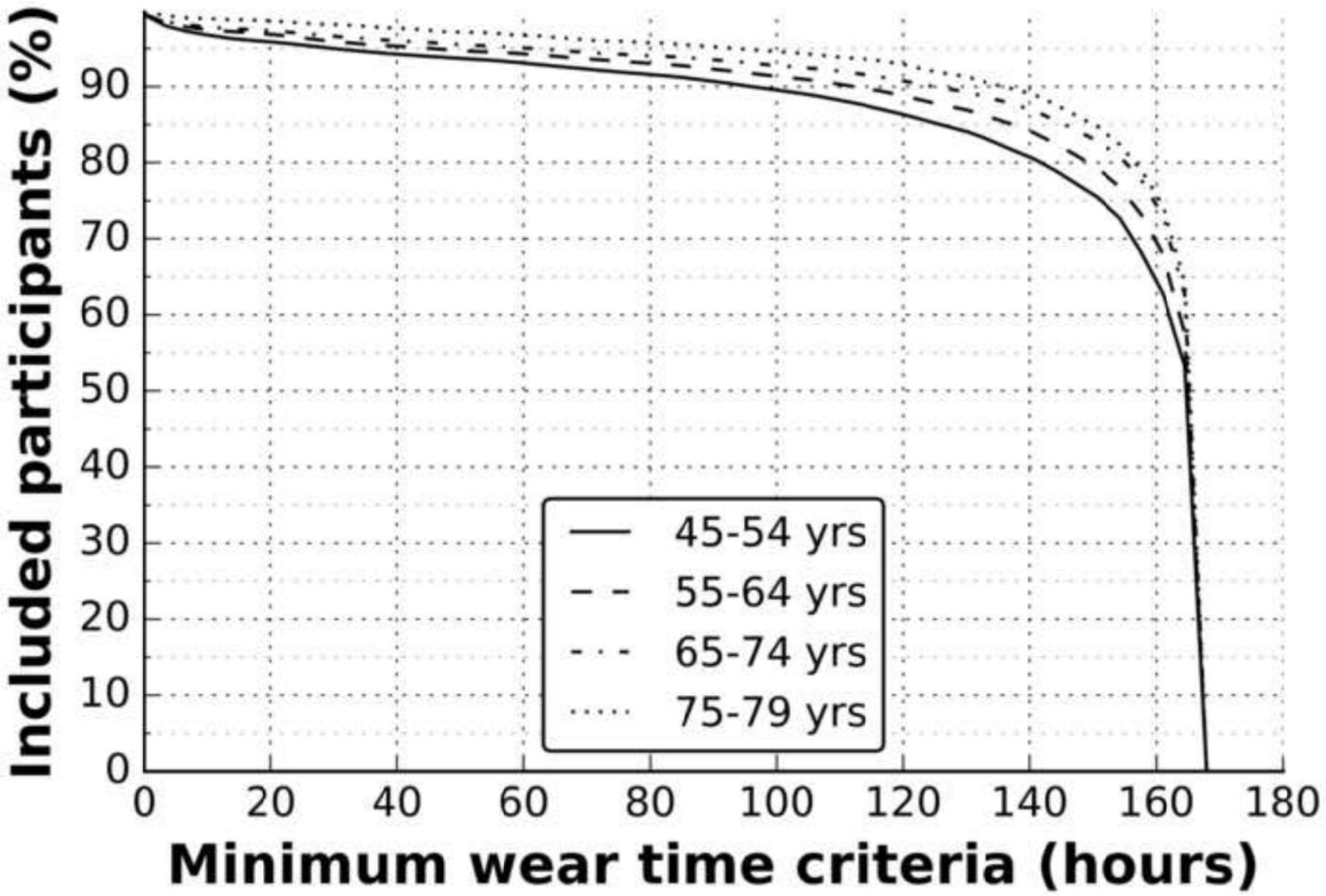

fig4

Click here to download Figure fig4.tiff

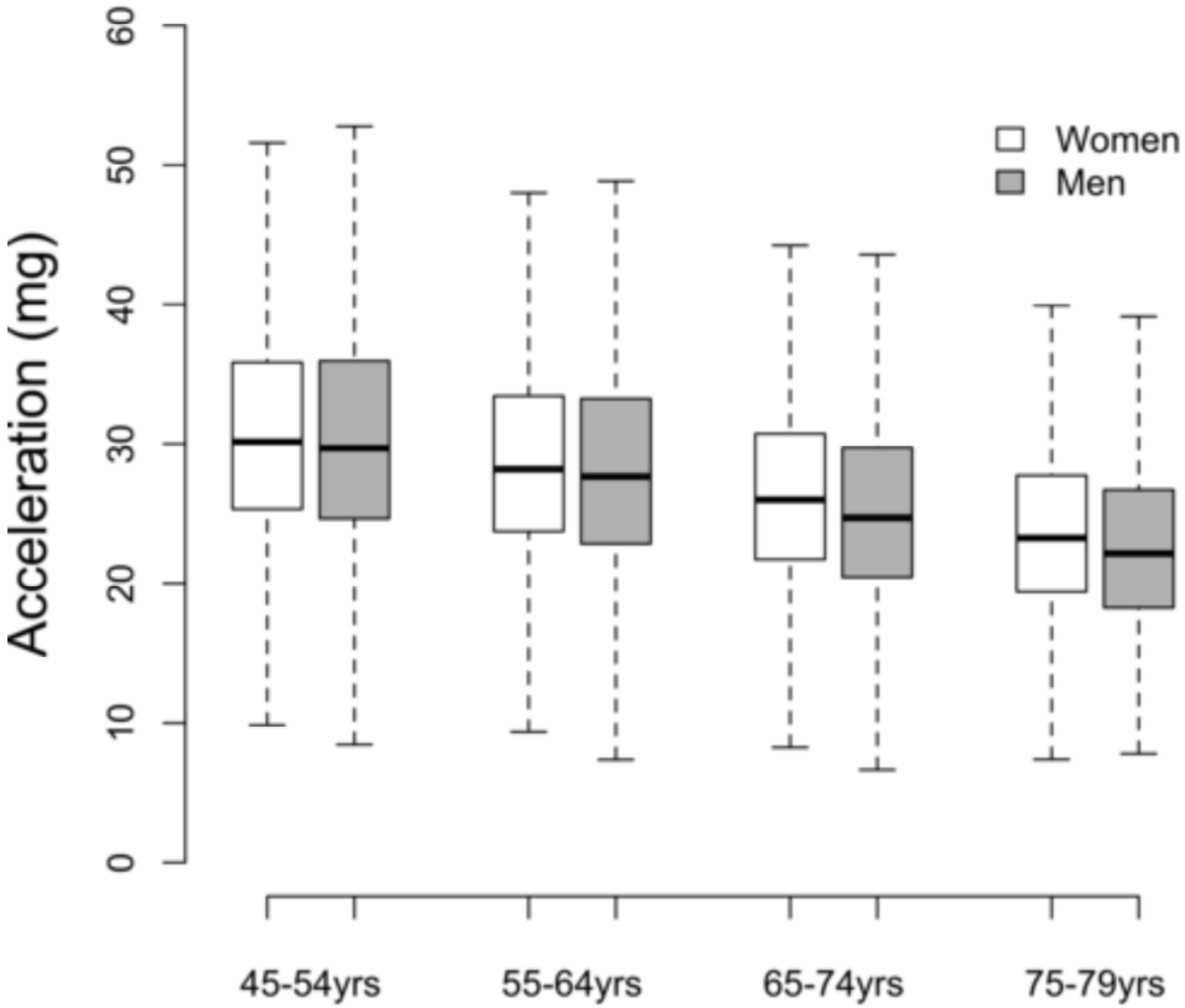

fig5

Click here to download Figure fig5.tiff

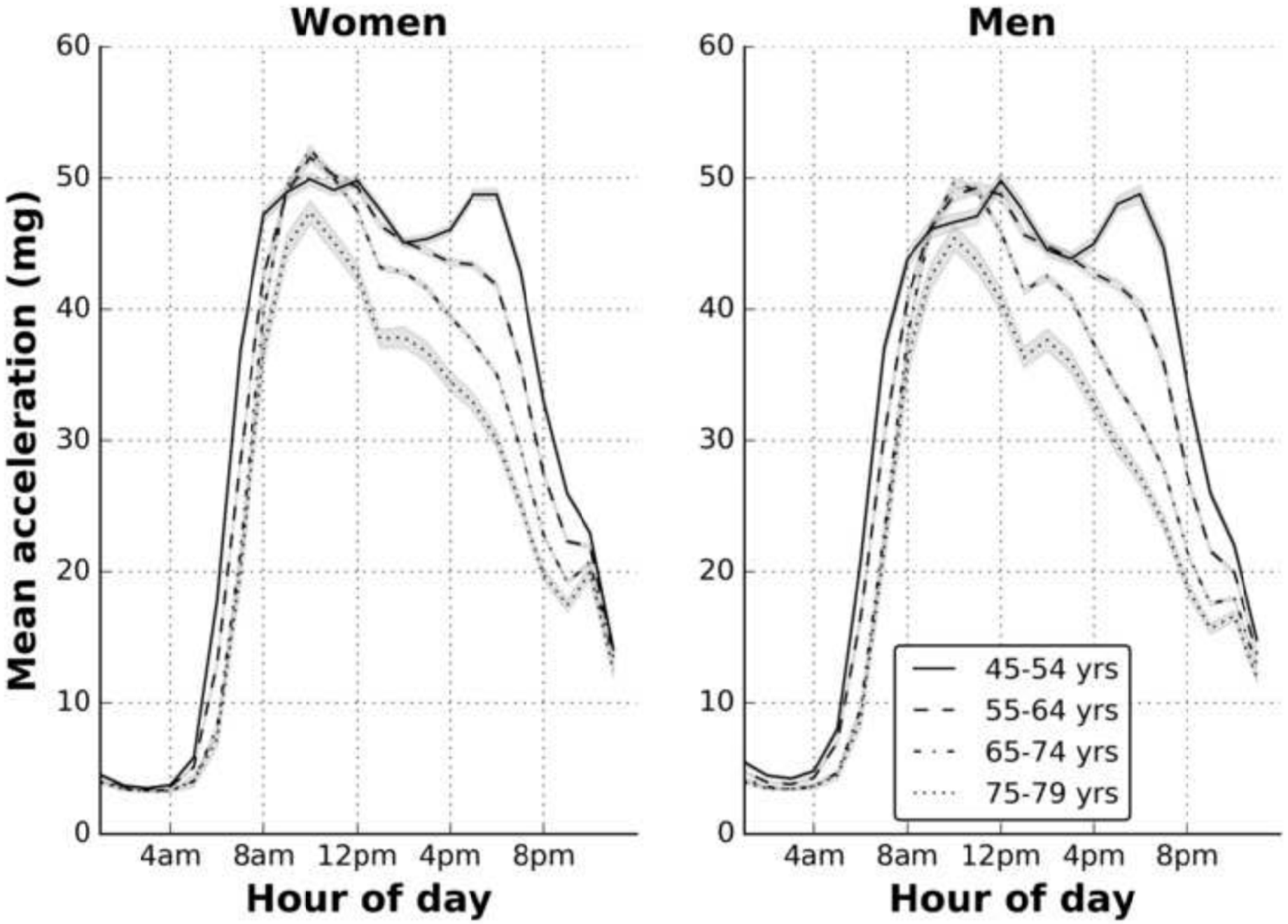

fig6

Click here to download Figure fig6.tiff

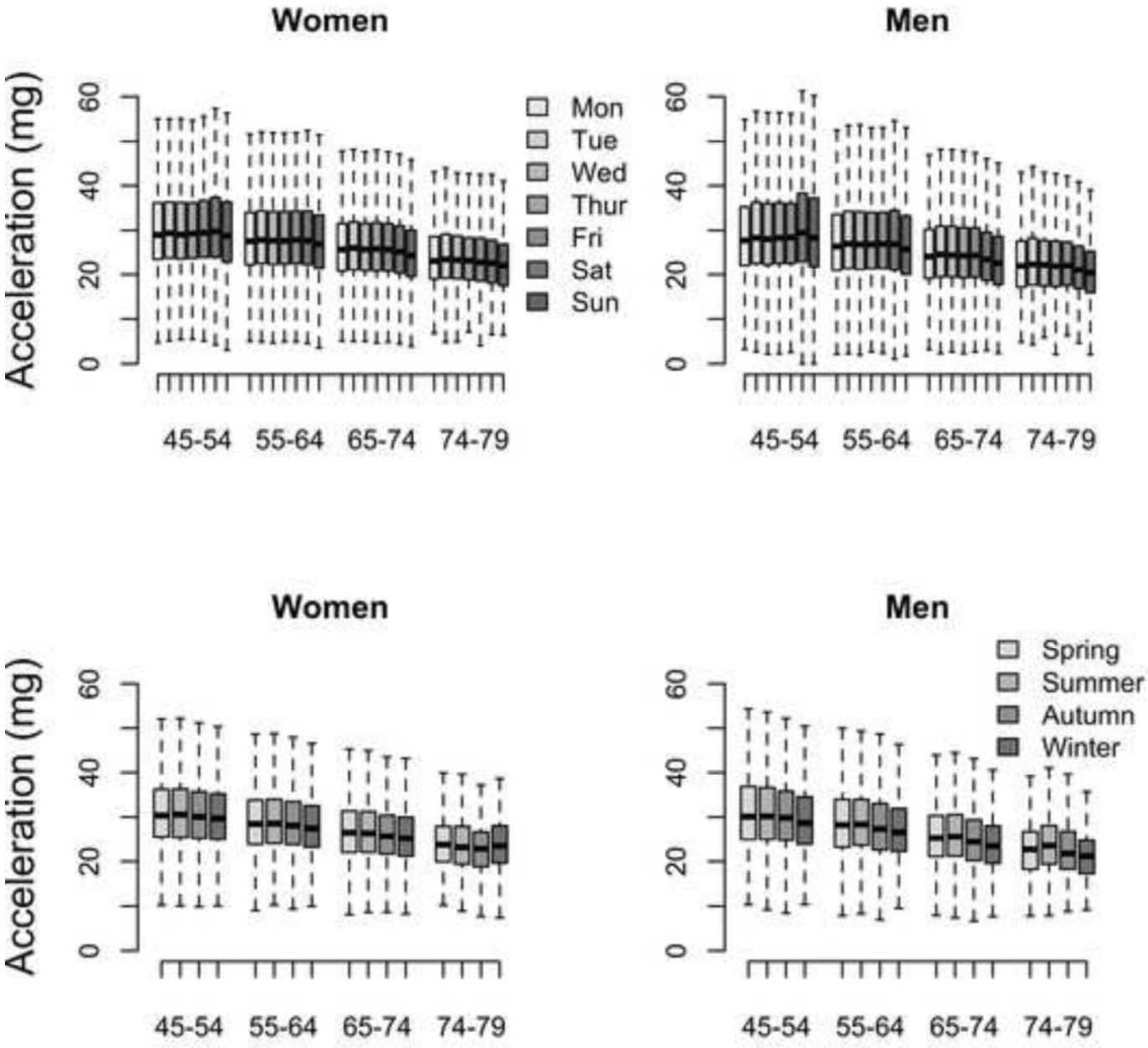

fig7

Click here to download Figure fig7.tiff