Influence of Accelerometer Calibration Approach on MVPA Estimates for Adults

Charles E. Matthews[1], **Sarah Kozey Keadle**[1,2], **David Berrigan**[3], **John Staudenmayer**[4], **Pedro Saint Maurice**[1], **Richard P. Troiano**[5], and **Patty S. Freedson**[6]

[1]Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD

[2]Kinesiology Department, California Polytechnic State University, San Luis Obispo, CA

[3]Behavioral Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

[4]Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA

[5]Risk Factor Assessment Branch, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD

[6]Department of Kinesiology, University of Massachusetts, Amherst, MA

Abstract

Purpose—To compare estimates of moderate-vigorous physical activity (MVPA) duration derived from accelerometers calibrated only to walking and running activities to estimates from calibrations based on a broader range of lifestyle and ambulatory activities.

Methods—In a study of 932 older (50–74 yrs) adults we compared MVPA estimates from accelerometer counts based on three ambulatory calibration methods (Freedson 1952 counts/min (cpm); Sasaki 2690 cpm; activPAL 3+ METs) to estimates based on calibrations from lifestyle and ambulatory activities combined (Matthews 760 cpm; Crouter 3+ METs; Sojourn3× 3+ METs). We also examined data from up to 6 previous-day recalls describing the MVPA in this population.

Results—MVPA duration values derived from ambulatory calibration methods were significantly lower than methods designed to capture a broader range of both lifestyle and ambulatory activities ($p < 0.05$). MVPA (hrs/d) estimates in all participants were: Freedson (median [inter-quartile range]=0.35 [0.17–0.58]); Sasaki (0.91 [0.59–1.32]); and activPAL (0.97 [0.71–1.26]) compared to Matthews (1.82 [1.37–2.34]); Crouter (2.28 [1.72–2.82]); and Sojourn3× (1.85 [1.42–2.34]). Recall-based estimates in all participants were comparable (1.61 [0.89–2.57]) and indicated participation in a broad range of lifestyle and ambulatory MVPA.

Corresponding Author: Charles E. Matthews, PhD, Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive - 6E444, Bethesda, MD 20892-9704.

Conclusion—Accelerometer calibration studies that employ only ambulatory activities may produce MVPA duration estimates that are substantially lower than methods calibrated to a broader range of activities. These findings highlight the potential to reduce potentially large differences among device-based measures of MVPA due to variation in calibration study design by including a variety of lifestyle and ambulatory activities.

Keywords

Moderate-vigorous intensity; Epidemiology; Surveillance; Measurement; Behavior

INTRODUCTION

The use of accelerometers in epidemiologic studies has advanced our understanding of physical activity and public health, including providing the first objective estimates of moderate-vigorous intensity physical activity (MVPA) and sedentary time in the United States (1, 2), and more recent indications that light intensity physical activities also may have mortality benefits (e.g., (3, 4)). This progress was made with low tech "first generation" devices (e.g., vertical acceleration in 60 second epochs) worn at the waist, that were calibrated in the laboratory using ambulatory activities (i.e., walking and running) to identify activity count cut-points to estimate MVPA (e.g., (1, 5, 6)). The "next generation" of devices (e.g., 80Hz sampling in 3-axes) and associated calibration methods have moved on from use of activity count cut-points and are more refined and robust (e.g., (7–9)). New calibration methods for waist worn devices using densely sampled movement data and as many as 20–30 different physical activities to develop more advanced prediction models are being used to estimate both ambulatory and lifestyle activities (7), and machine learning methods have been used to identify specific activity types (10–12).

Much attention has focused on technical aspects of accelerometer use (e.g., epoch length, attachment site), the limitations of using activity count cut-points, and the value of using raw acceleration signals to minimize differences among monitors (e.g., (13, 14)). However, the design of calibration studies and the selection of activities included in the protocol, the key methodological step that translates acceleration signals into public health relevant metrics such as free-living MVPA values (14, 15), often receive less attention. Design choices may be critical because high quality studies using indirect calorimetry as a criterion have reported that MVPA values derived from monitors calibrated only to ambulatory activities are substantially lower than measured MVPA values (16, 17), and that methods designed to capture a broader range of lifestyle activities produced more accurate MVPA estimates at the group level (16). Similar results have also been noted in studies using direct observation as a criterion (8), when comparing devices calibrated only to ambulatory activities to those calibrated to a broad range of activities (18, 19), and in relation to previous-day recalls (20, 21).

Despite these known differences, new data processing methods for next generation devices worn on the wrist or waist continue to be developed using only ambulatory activities (e.g., (22)) in controlled and free-living settings, suggesting that it is not widely appreciated how much the choice of activities in calibration studies can affect the future estimates of free-

living MVPA. Furthermore, determining the magnitude of differences in group level MVPA estimates associated with different monitor calibration methods may provide insights into the potential for better harmonization of these studies to minimize future variation between monitors and data processing methods.

Accordingly, the purpose of this investigation was to compare mean MVPA estimates from three methods calibrated to only ambulatory activities to estimates from three methods calibrated to lifestyle and ambulatory activities. Two ambulatory-based methods relied on activity count cut-points (i.e., Freedson (5), Sasaki (24)), and the third relied on steps/minute to estimate MVPA (activPAL (25)). In contrast, three distinct methods were calibrated more broadly and included one method that used an activity count cut-point (Matthews (26)), another that integrated two-regression models to capture both ambulatory and lifestyle activity (Crouter (7)), and a third that used hybrid machine learning to estimate MVPA events (Sojourn 3× (8)). We also explored data from previous-day recalls to better understand the breadth of activities contributing to MVPA among our participants.

METHODS

Study Design

The Interactive Diet and Activity Tracking in AARP (iDATA) study was designed to evaluate a variety of diet and physical activity measures suitable for use in epidemiologic studies **of** middle-age and older adults. A convenience sample of AARP members (50–74 years) from Pittsburgh, PA who spoke English, had internet access, were not on a weight-loss diet, had a body mass index (BMI) < 40 kg/m^2, and were free of major medical conditions and mobility limitations were studied. The investigation protocol was approved by the NCI Special Studies Institutional Review Board. Consented participants visited the study center up to three times over 12-months and completed several diet and physical activity measurements, including up to 6 physical activity previous day recalls over a year and two accelerometer measurement periods (7 days each), six months apart. The ActiGraph (model GT3X), a triaxial accelerometer, was requested to be worn on the right hip from the time participants got out of bed for the day until they went to bed for the night. It was initialized to record in 1-s epochs (tri-axial acceleration) with the low-frequency extension enabled. The activPAL 3D (PAL Technologies, Glasgow, Scotland), an accelerometer worn on the mid–right thigh was also used. It estimates the duration of specific behavioral events including time spent lying, sitting, standing, stepping, and in non-wear time. Participants were asked to wear the activPAL for 24-hours and to record the wear date and time they got out of bed in the morning and into bed each night (i.e., waking day). To confirm wear dates and estimate the waking day for both monitors we used monitor wearing logs, and to estimate non-wear time during the waking day we used the Choi algorithm (9) for the ActiGraph, and the native non-wear time estimate from the activPAL. Further details of the iDATA study can be found elsewhere (21), and this report extends our prior investigation of the accuracy of the ActiGraph and previous-day recalls for estimating energy expenditure and total active and sedentary time (21). Table 1 describes key technical details (epoch length, axis measured) and the characteristics of the calibration study and prediction method

from which each MVPA estimate originated. For simplicity, we focused on estimating values of total accumulated MVPA at the group level.

Ambulatory Calibration Methods

Two ActiGraph-based estimates of MVPA derived from methods that only used walking and running (i.e., ambulatory) activities for calibration were calculated using vertical acceleration cut-points from Freedson (5) and vector magnitude cut-points of Sasaki (24). Time in MVPA recorded using activPAL was derived from ambulatory activities because this monitor is attached to the leg and uses a prediction equation that estimates MET values from step frequency, or cadence (i.e., MET-hours=$(1.4 \times$ duration [h]) + (4 − 1.4) × (cadence [steps/min]/120) × duration (25)).

Lifestyle + Ambulatory Calibration Methods

Estimates derived from calibration methods that employed both lifestyle and ambulatory activities were calculated using the 760 count/minute (cpm) cut-point of Matthews (20, 27) that was initially developed using 13 light and moderate intensity activities and refined using free-living indirect calorimetry measures (27). Two more sophisticated prediction methods that did not rely on activity count cut-points were also examined. The refined Crouter method (7) applies two-regression equations developed from 18 activities to estimate energy expenditure in both ambulatory and lifestyle activities. Sojourn 3× (Soj3×) (8) employs three-dimensional acceleration features from 1 second epochs and a hybrid machine-learning (neural network) model trained on approximately 30 activities (12) to classify behaviors and estimate energy expenditure.

Previous-day Recall Measures

Activities Completed Over Time in 24 Hours (ACT24) is an internet-based previous day recall designed to estimate total time (hrs/d) spent sleeping (in bed), sedentary (sitting or reclining), engaged in physical activity, and to estimate energy expenditure associated with these behaviors (MET-hrs/d) (21). To complete ACT24, respondents select from more than 200 individual activities, record the duration of each activity (typically 20–30 distinct behaviors), and answer pertinent follow-up questions including body position. Sedentary behaviors were defined as those performed during the waking day (out of bed) while sitting or reclining and that require little energy expenditure (typically < 1.8 METs) (36). Time spent in moderate-vigorous intensity (e., 3 METs) activities were also calculated. Similar previous-day recalls have been shown to provide useful estimates of MVPA (20, 26).

Accounting for Monitor Non-wear and Missing Time

To minimize differences between measures due to variations in non-wear or missing time, we used a simple imputation approach. Of the mean 15.8 hrs/d waking day reported on the monitor wearing logs, activPAL recorded a mean of 0.7 hrs/d of non-wear time, while the ActiGraph recorded 1.9 hrs/d of non-wear. For days with non-wear time, we imputed MVPA values for the non-wear time in proportion to time recorded in MVPA during the waking day (i.e., %MVPA) for each monitor. ACT24 missing time was minimal (0.005 hrs/d), so no adjustments were made.

Statistical Analysis

We described our study population using means and frequency counts overall and by sex. To examine differences in the MVPA estimates we calculated mean MVPA values using all valid days of observation and examined the distributions of each method (i.e., mean, SD, median, inter-quartile range) and plotted these values by sex. Paired t-tests were used to test for differences between methods. To better understand the specific moderate-vigorous intensity activities in which adults in this study engaged, we also evaluated 5,018 previous-day recalls from 923 participants with at least one recall and one valid day of accelerometer wear. We summed values for moderate and vigorous intensity activity duration over each recall day by activity domain, and ranked the individual activities reported. We then calculated the proportion of days in the overall population each activity was reported and calculated the mean duration of each activity on days of participation.

RESULTS

Average age and BMI of participants was 63 years and 28 kg/m^2, respectively, 30.9% of participants were obese, and the number of males and females was approximately equal. Participants reported spending 15.8 hrs/d out of bed (i.e., waking day) and the activPAL and ActiGraph estimates were derived from 9 to 10 days of data collection on average (Table 2).

Figure 1 shows the median and inter-quartile range of each estimate of MVPA duration in males and females. In both sexes combined, median values for methods calibrated only to ambulatory activities (Freedson, Sasaki, activPAL) ranged from 0.35 (Freedson) to 0.97 (activPAL) hrs/d. The MVPA estimate from Sasaki (0.91 hrs/d) derived from vector magnitude activity counts at the waist was not significantly different from activPAL (0.97 hrs/d) which used recorded steps/minute to estimate MVPA (see Table, Supplemental Digital Content 1, Mean values and mean differences between methods, by sex). Mean values and mean differences between methods, by sex). Overall, median values of MVPA derived from the ambulatory calibration methods alone were broadly similar to one another and all were much lower than the median values from the group of three methods calibrated to lifestyle and ambulatory activities (Matthews, Crouter, Soj30×; see Figure 1). Median values for the latter group were also broadly consistent with one another and ranged from 1.82 (Matthews) to 2.28 (Crouter) hrs/d. T-tests revealed that values for the ambulatory methods were significantly lower ($p < 0.05$) than the lifestyle/ambulatory calibration methods (see SDC Table 1). Within the lifestyle/ambulatory calibration methods we found small mean differences (< 2%) between Matthews and Soj3× that were significant in men (−0.05 hrs/d; p=0.002) but not women (0.004 hrs/d; p=0.78), and estimates from these two methods were lower than those from Crouter (p <0.05).

To gain insight into the type of moderate-vigorous intensity activities in which our population engaged, we examined data reported on up to six previous day recalls collected over 12-months. The distribution of overall MVPA duration reported in all participants was comparable to that observed using the lifestyle/ambulatory-based accelerometer methods (median=1.61 (IQR=0.89–2.57) hrs/d; see also Figure 1). Figure 2 describes mean values of <u>moderate</u> intensity activity by the major activity domains and lists the most common activities reported in each domain. Moderate intensity activities at work and in exercise

domains were often more ambulatory in nature, while household and lawn/garden activities were less rhythmic, involved more intermittent movement and often require upper-body movement (e.g., cleaning floors, sweeping, pick-up stick/debris in yard, trimming/weeding, etc.). Vigorous activities reported (data not shown) were primarily exercise and sports related activities, and carrying and lifting heavy loads at work or at home. The most common vigorous exercise activities reported were use of cardio machines, cycling, weight lifting, running, and swimming.

DISCUSSION

In this large study of older adults, we found MVPA estimates derived from accelerometers calibrated only to ambulatory activities to be similar to one another, but substantially lower than those derived from three distinct methods designed to capture a broader range of MVPAs. This finding was consistent in men and women and was independent of the epoch lengths (60, 10, 1 sec), axis of acceleration (vertical vs 3-dimensional), ActiGraph model type (7164, GT3X), and the modeling/prediction approach employed. Importantly, three methods examined were not limited by reliance on activity count cut-points; one in the ambulatory calibration category (activPAL) and two in the lifestyle and ambulatory category (Crouter, Soj3×). Furthermore, MVPA values from our previous-day recall were similar to those from our broadly calibrated accelerometer-based methods, and detailed analysis of commonly reported activities indicated that our participants engaged in a wide variety of moderate intensity lifestyle and ambulatory activities. Together, these findings support the notion that calibration method matters, and strongly suggest that reliance on ambulatory activities alone to calibrate accelerometers will not capture the full spectrum of MVPA accumulated in daily life. The present results also challenge our common understanding of the amount of MVPA adults may do each day, which in recent years has been quantified primarily using accelerometers calibrated using walking and running protocols in controlled settings.

Understanding the best approach to measuring MVPA with accelerometers has been plagued by many competing cut-points proposed for use with first generation devices and a general lack of rigorous free-living validation studies that would clarify differences between measures in their accuracy and precision. This study overcomes this limitation in part by examining two groups of methods to estimate MVPA (with and without cut-points) using a variety of accelerometer technology and prediction methods. Our results confirm earlier reports (16, 17, 20) and extend these findings to modern accelerometers (e.g., 3-dimensional acceleration) and more sophisticated calibration methods (e.g., Crouter, Soj3×). Importantly, the Soj3× method which reflects the current state-of-the-art for calibration study designs has been rigorously tested for validity in an independent study sample using a strong criterion measure (8). Lyden (8) reported that for MVPA duration Soj3× was within 1% of direct observation. By comparison, estimates from the Freedson method were lower and values from the Crouter two-regression method were somewhat higher. These results are also broadly consistent with at least two studies that reported a similar divergence in MVPA values by calibration method using indirect calorimetry as a criterion (16, 17), and mean MVPA values as high as 2.2 hrs/d were observed in a population-based sample from Iowa using the Sensewear device, a multi-sensor monitor that is calibrated to a broad range of

activities (19). Our results are also consistent with laboratory-based studies, for example, Kozey et al. reported that 7 self-paced moderate-vigorous intensity (3+ METs) lifestyle activities (sweeping, mopping, raking, painting, trimming, vacuuming and gardening) all had average activity counts values below commonly used ambulatory activity count cut-points for MVPA (28). We assume Soj3× to be our best estimate of MVPA, and find it notable that the other two methods in this category (Matthews, Crouter) produced broadly consistent estimates of MVPA duration. Although we were not able to test other accelerometer brands in this study, or wrist worn monitors, based on our results one might hypothesize that such instruments calibrated to capture only ambulatory activities could also capture lower MVPA estimates in large populations with heterogeneous activity patterns. Indeed, Leinonen et al. (29) recently reported MVPA values in middle-age adults of about 1.06 hrs/d as measured by the waist-worn Hookie AM20 device using mean amplitude deviation and raw acceleration data calibrated to only ambulatory activities (22); results quite similar to the 0.97 hrs/d of MVPA we observed using activPAL. Also, Menai et al. (30) reported that older adults (60–83 yrs) in the Whitehall II Study accumulated only 0.44 hrs/d of MVPA using a wrist-worn device collecting raw acceleration data and calibrated to ambulatory MVPAs (23).

Recommendations for calibration studies have encouraged inclusion of a broad range of daily activities (14, 15) but the impact of including only ambulatory activities vs. lifestyle and activities that primarily involve walking and running has not been clearly documented in large studies using more sophisticated acceleration signals. Our findings suggest that the magnitude of differences in MVPA that may result from use of divergent calibration study designs may be quite large. Although our results may appear obvious in hindsight, we suggest that this observation is not widely appreciated among general users of accelerometer data, and possibly by designers of calibration studies for new devices to predict behavioral metrics like MVPA from bodily movement at the wrist or waist (e.g., (22, 30, 31)). Anchored by the contrasting median values of MVPA from the activPAL (0.97 hrs/d) and Soj3× method (1.85 hrs/d), the present results suggest that adults may accumulate as much as 90% more MVPA each day than previously thought based on ambulatory estimates alone. These findings raise important questions about the interpretation of results from epidemiologic studies that use activity monitors calibrated only to ambulatory activities, or calibrated more broadly, such as: what is the optimal dose of accelerometer measured MVPA for greater longevity—is it 4–11 min/d or 50–86 min/d (3)?; do self-report instruments like the International Physical Activity Questionnaire over- (32) or under-estimate (33, 34) MVPA duration?; and, is the prevalence of meeting MVPA guidelines about 10%, or closer to 60% (35)? Notably, the Matthews cut-point was comparable to Soj3×, suggesting even data from "first generation" devices may provide useful population estimates of MVPA derived from both lifestyle and ambulatory movement. Given the large differences in MVPA by calibration method observed here, and the importance of the questions noted above, it is critical that we conduct high quality studies to identify accurate measures and better understand the background level of MVPA in the population.

Rigorous validation studies that might accomplish this task have several essential characteristics (14, 15). First, they must be conducted in a free-living setting and in a study sample that is independent from the study sample used to develop prediction models. Second, accurate and precise criterion measures are required, but implementing strong

measures of MVPA, such as direct observation (8) from video recordings, indirect calorimetry (16), or use of multi-sensor methods (18) can be logistically challenging and requires attention to appropriate sampling (e.g., time of day, day of week) to ensure capture of more representative daily behavior. Third, conduct of studies in different populations with heterogeneous patterns of activity is also key so that unique aspects of a one population don't obscure results. For example, testing populations that accumulate a large proportion of total MVPA in ambulatory activity (e.g., mail carriers) may not reveal the limitations of methods calibrated to capture only ambulatory MVPA.

The ability to access raw acceleration signals from devices has been an important advance that can reduce differences among monitors (36) and our findings suggest that a more concerted effort to enhance and harmonize future calibration efforts is another opportunity to reduce differences among device-based measures of MVPA. Freedson et al. noted the importance of including appropriate activities in training protocols for better prediction accuracy (12). Ideally, the activities used in a calibration study for model development should come from activities done in free-living settings reflective of a typical day, and this is increasingly possible with direct observation via video. The basic logic is that the activities employed in calibration studies should be a good match for the major activities encountered in daily life (14, 15), yet several new calibration studies have continued to rely primarily on ambulatory activities to develop MVPA estimates (e.g., (22, 31)). Figure 2 illustrates the breadth of moderate intensity activities accumulated in daily life, and our finding that recall-based estimates of MVPA were comparable to estimates from Soj3× and the Matthews cut-point suggest that short-term recall-based methods such as ACT24 or time-use surveys could help identify prevalent activities to include in calibration studies conducted in more controlled settings. Some resources to fulfill this purpose already exist from time-use surveys linked to the Compendium of Physical Activities (37, 38).

The strengths and limitations of our study should also be considered. The main strengths include a large sample size of middle-aged to older adults, use of distinct methods to estimate MVPA within each calibration-type category, as well as integration of previous-day recall estimates to gain more insight into the specific ambulatory and lifestyle activities reported in this population. It is important to note that due to differences in the time-frame of the previous-day recalls and the accelerometer monitoring periods, our description of moderate-vigorous intensity activities reported on the recalls should be considered a population estimate of these activities over the year of the study, rather than specific activities that may have been done while wearing the monitors. The main limitation of this study is the lack of an unequivocal gold standard measure of free-living MVPA that would clarify the accuracy and precision of each of our estimates at both the group and individual level. We are unable determine the most accurate and precise MVPA estimate, but evidence of validity for Soj3× (8), the general consistency of all our lifestyle and ambulatory estimates, and the unambiguous ambulatory nature of the activPAL MVPA values support our conclusions. Additional studies are needed to characterize the accuracy and precision of these methods at the individual level. Another limitation is that our study population is somewhat older than the study populations in which the accelerometers were originally calibrated. If the relation between activity counts and energy expenditure differs greatly by age, it is possible that an age by calibration methodology interaction may have affected our

results. However, studies conducted in younger participants (16, 17, 20) have reported findings that are consistent with ours, suggesting that our findings may be applicable to younger adults as well. Research in younger study populations should seek to replicate our findings. Additionally, our study sample was a relatively healthy group of older adults, so our results may not generalize to other populations, such as younger, lower SES, or rural adults whose activity patterns might differ from our sample, or people with underlying mobility and gait limitations, both healthy and unhealthy.

In conclusion, we observed that MVPA duration estimates derived from activity monitors calibrated with only ambulatory activities were substantially lower than estimates based on calibration methods designed to capture a broad range of both lifestyle and ambulatory activities. These findings have implications for the design of future monitor calibration studies, which may need to be conducted in free-living populations and/or include more representative daily activities. Importantly new methods should be rigorously tested for validity in free-living studies designed to demonstrate equivalence or superiority to existing methods (15). Our results also have implications for future application of accelerometers in etiologic studies of disease risk (39, 40), population surveillance, and evaluation of self-report measures of physical activity. We hope this research leads to conduct of more definitive studies that will help us understand how best to measure MVPA in future studies, and the underlying level of MVPA in the population.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical Activity in the United States Measured by Accelerometer. Med Sci Sports Exerc. 2008; 40(1):181–8. [PubMed: 18091006]

2. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors - United States 2003–2004. Am J Epidemiol. 2008; 167:875–81. [PubMed: 18303006]

3. Evenson KR, Wen F, Herring AH. Associations of Accelerometry-Assessed and Self-Reported Physical Activity and Sedentary Behavior With All-Cause and Cardiovascular Mortality Among US Adults. Am J Epidemiol. 2016; 184(9):621–32. [PubMed: 27760774]

4. Matthews CE, Keadle SK, Troiano RP, et al. Accelerometer-measured dose-response for physical activity, sedentary time, and mortality in US adults. Am J Clin Nutr. 2016; 104(5):1424–32. [PubMed: 27707702]

5. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc. 1998; 30:777–81. [PubMed: 9588623]

6. Colley RC, Tremblay MS. Moderate and vigorous physical activity intensity cut-points for the Actical accelerometer. J Sports Sci. 2011:29. [PubMed: 21077002]

7. Crouter SE, Kuffell E, Hass JD, Frongillo EA, Bassett DR. Refined Two-Regression Model for the ActiGraph Accelerometer. Med Sci Sports Exerc. 2010; 42(5):1029–37. [PubMed: 20400882]

8. Lyden K, Keadle SK, Staudenmayer J, Freedson PS. A method to estimate free-living active and sedentary behavior from an accelerometer. Med Sci Sports Exerc. 2014; 46(2):386–97. [PubMed: 23860415]

9. van Hees VT, Gorzelniak L, Dean Leon EC, et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PLoS One. 2013; 8(4):e61691. [PubMed: 23626718]

10. Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of Novel Techniques to Classify Physical Activity Mode Using Accelerometers. Med Sci Sports Exerc. 2006; 38(9):1626–34. [PubMed: 16960524]

11. Bonomi AG, Plasqui G, Goris AHC, Westerterp KR. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. J Appl Physiol. 2009; 107(3):655–61. [PubMed: 19556460]

12. Freedson PS, Lyden K, Kozey-Keadle S, Staudenmayer J. Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample. J Appl Physiol. 2011; 111(6):1804–12. [PubMed: 21885802]

13. van Hees VT, Thaler-Kall K, Wolf KH, et al. Challenges and Opportunities for Harmonizing Research Methodology: Raw Accelerometry. Methods Inf Med. 2016; 55(6):525–32. [PubMed: 27714396]

14. Bassett DR Jr, Rowlands A, Trost SG. Calibration and validation of wearable monitors. Med Sci Sports Exerc. 2012; 44(1 Suppl 1):S32–8. [PubMed: 22157772]

15. Welk GJ, McClain J, Ainsworth BE. Protocols for evaluating equivalency of accelerometry-based activity monitors. Med Sci Sports Exerc. 2012; 44(1 Suppl 1):S39–49. [PubMed: 22157773]

16. Crouter SE, DellaValle DM, Haas JD, Frongillo EA, Bassett DR. Validity of ActiGraph 2-Regression Model and Matthews and NHANES and Cut-Points for Assessing Free-Living Physical Activity. J Phys Act Health. 2013; 10(4):504–14. [PubMed: 22975460]

17. Strath SJ, Bassett DR Jr, Ham S, Swartz AM. Assessment of Physical Activity by Telephone Interview versus Objective Monitoring. Med Sci Sports Exerc. 2003; 35(12):2112–8. [PubMed: 14652510]

18. Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field Validation of the MTI Actigraph and BodyMedia Armband Monitor Using the IDEEA Monitor. Obesity. 2007; 15(4):918–28. [PubMed: 17426327]

19. Welk GJ, Beyler NK, Kim Y, Matthews CE. Calibration of Self-Report Measures of Physical Activity and Sedentary Behavior. Med Sci Sports Exerc. 2017; 49(7):1473–81. [PubMed: 28240704]

20. Matthews CE, Keadle SK, Sampson J, et al. Validation of a previous-day recall measure of active and sedentary behaviors. Med Sci Sports Exerc. 2013; 45(8):1629–38. [PubMed: 23863547]

21. Matthews CE, Keadle SK, Moore SC, et al. Measurement of Active & Sedentary Behavior in Context of Large Epidemiologic Studies. Med Sci Sports Exerc. 2018 Feb; 50(2):266–76. [PubMed: 28930863]

22. Vähä-Ypyä H, Vasankari T, Husu P, et al. Validation of Cut-Points for Evaluating the Intensity of Physical Activity with Accelerometry-Based Mean Amplitude Deviation (MAD). PLoS One. 2015; 10(8):e0134813. [PubMed: 26292225]

23. Hildebrand M, van Hees V, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. Med Sci Sports Exerc. 2014; 46(9):1816–24. [PubMed: 24887173]

24. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport. 2011; 14(5):411–6. [PubMed: 21616714]

25. PAL Technologies. activPAL Operating Guide. Appendix A - Technical Description. 2010:15–7.

26. Matthews CE, Ainsworth BE, Hanby C, et al. Development and testing of a short physical activity recall questionnaire. Med Sci Sports Exerc. 2005; 37(6):986–94. [PubMed: 15947724]

27. Matthews CE. Calibration of Accelerometer Output for Adults. Med Sci Sports Exerc. 2005; 37(11):S512–S22. [PubMed: 16294114]

28. Kozey SL, Lyden K, Howe CA, Staudenmayer JW, Freedson PS. Accelerometer output and MET values of common physical activities. Med Sci Sports Exerc. 2010; 42(9):1776–84. [PubMed: 20142781]

29. Leinonen A-M, Ahola R, Kulmala J, et al. Measuring Physical Activity in Free-Living Conditions—Comparison of Three Accelerometry-Based Methods. Front Physiol. 2017; 7(681)

30. Menai M, van Hees VT, Elbaz A, Kivimaki M, Singh-Manoux A, Sabia S. Accelerometer assesed moderate-to-vigorous physical activity and successful aging: results from teh Whitehall II study. Sci Rep. 2017; 7(45772)

31. Guiry JJ, van de Ven P, Nelson J, Warmerdam L, Riper H. Activity recognition with smartphone support. Med Eng Phys. 2014; 36(6):670–5. [PubMed: 24641812]

32. Timperio A, Salmon J, Rosenberg M, Bull FC. Do logbooks influence recall of physical activity in validation studies? Med Sci Sports Exerc. 2004; 36(7):1181–6. [PubMed: 15235322]

33. Dyrstad SM, Hansen BH, Holme IM, Anderssen SA. Comparison of Self-reported versus Accelerometer-Measured Physical Activity. Med Sci Sports Exerc. 2014; 46(1):99–106. [PubMed: 23793232]

34. Hagstromer M, Ainsworth BE, Oja P, Sjostrom M. Comparison of a subjective and an objective measure of physical activity in a population sample. J Phys Act Health. 2010; 7(4):541–50. [PubMed: 20683097]

35. Watson KB, Carlson SA, Carroll DD, Fulton JE. Comparison of accelerometer cut points to estimate physical activity in US adults. J Sports Sci. 2014; 32(7):660–9. [PubMed: 24188163]

36. Rowlands AV, Yates T, Davies M, Khunti K, Edwardson CL. Raw Accelerometer Data Analysis with GGIR R-package: Does Accelerometer Brand Matter? Med Sci Sports Exerc. 2016; 48(10): 1935–41. [PubMed: 27183118]

37. Tudor-Locke C, Johnson WD, Katzmarzyk PT. Frequently Reported Activities by Intensity for U.S. Adults: The American Time Use Survey. Am J Prev Med. 2010; 39(4):e13–e20. [PubMed: 20837277]

38. Tudor-Locke C, Washington TL, Ainsworth BE, Troiano RP. Linking the American Time Use Survey (ATUS) and the Compendium of Physical Activities: methods and rationale. J Phys Act Health. 2009; 6(3):347–53. [PubMed: 19564664]

39. The German National Cohort: aims, study design and organization. Eur J Epidemiol. 2014; 29(5): 371–82. [PubMed: 24840228]

40. Doherty A, Jackson D, Hammerla N, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS One. 2017; 12(2):e0169649. [PubMed: 28146576]
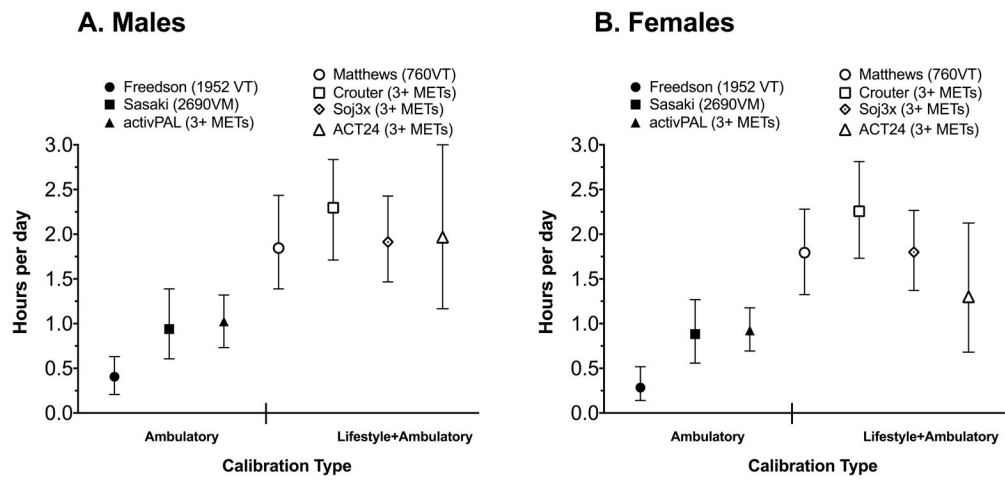
Figure 1.
Duration of MVPA by calibration type, measurement method, and sex.

Values are median (25th – 75th percentiles).

See text for results of statistical testing.

Moderate=1.67 hrs/d (SD=1.36)

□ Other
walking for transportation (10%; 0.45 hrs)
miscellaneous/unspcified (4%; 1.33 hrs)
leisure/unspecified (2%; 1.89 hrs)
transport/unspecified (2%; 1.08 hrs)
volunteer work (1%; 2.63 hrs)

■ Household chores
cleaning floors (11%; 0.84 hrs)
walking dog (6%; 0.77 hrs)
child care (4%; 0.92 hrs)
sweeping (3%; 0.74 hrs)
elder care (2%; 1.28 hrs)

□ Work
standing/much walking (7%; 3.97 hrs)
walking longer distances (2%; 0.67 hrs)

□ Lawn/garden/repairs
mowing lawn (4%; 1.28 hrs)
pick-up sticks/debris (4%; 0.72 hrs)
haul brush/trimmings (3%; 1.10 hrs)
digging (3%; 1.67 hrs)
trim/weed with machines (2%; 0.68 hrs)
raking (2%; 0.98 hrs)
sweeping drive-way/side-walk (2%; 0.45 hrs)
interior remodeling (2%; 3.3 hrs)

□ Exercise/fitness
walking for exercise (18%; 0.92 hrs)
weight lifting (4%; 0.70 hrs)
exercise/sports unspecified (2%; 1.39 hrs)
aerobics (2%; 1.08 hrs)
golf (1%; 3.61 hrs)

0.24 hrs/d (14%)
0.35 hrs/d (21%)
0.33 hrs/d (20%)
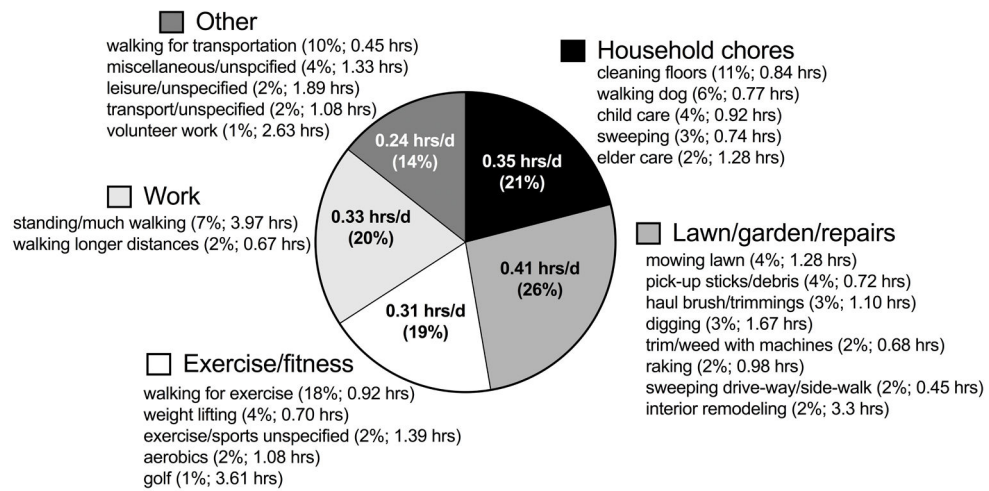0.41 hrs/d (26%)
0.31 hrs/d (19%)

Figure 2.
Domain specific moderate intensity activity duration (hrs/d) reported via previous-day recalls and the most common activities reported in each domain.

Table 1

Description of monitor-based estimates of moderate-vigorous intensity physical activity

Calibration type	Method	Technical Details (monitor)			Calibration Study Details		
		AG Model*/Epoch length	Axis measured	Number/type of activities	Modeling/prediction approach	Intensity Classification
Ambulatory						
	Freedson	**7164**/60 sec	Vertical (VT)	3 / walking, running	Linear regression	1952 VT cts/min
	Sasaki	**GT3X**/60 sec	3-dimensions; Vector Magnitude (VM)	3 / walking, running	Linear regression	2690 VM cts/min
	activPAL	Not applicable/20 Hz	3-dimensions	Stepping (cadence)	Linear regression	3+ METs
Lifestyle + Ambulatory						
	Matthews	**7164**/60 sec	Vertical (VT)	13 / lifestyle and walking activities; refined vs. free-living indirect calorimetry	None	760 VT cts/min
	Crouter	**7164**/10 sec	Vertical (VT)	18 / lifestyle + walking, running activities	2 regression equations (moving average)	3+ METs
	Sojourn3×	**GT3X**/1 sec	3-dimensions	30 / lifestyle and walking, running activities	Hybrid machine learning and neural networks	3+ METs

*
AG = ActiGraph

Table 2

Description of participant characteristics and relevant monitor administration details

Participant Characteristics	Overall (N=932)	Males (n=461)	Females (n=471)
Age (yrs)	63.2 (5.9)	64.2 (5.7)	62.3 (6.0)
Height (cm)	169.4 (9.3)	176.1 (6.8)	162.8 (6.1)
Weight (kg)	80.9 (16.3)	88.4 (14.8)	73.5 (14.3)
Body mass index (kg/m^2)	28.1 (4.7)	28.5 (4.3)	27.7 (5.0)
Obese (% BMI 30+)	30.9%	30.1%	31.3%
Monitor administration			
Reported waking day (hrs)	15.8 (1.0)	15.9 (0.9)	15.6 (1.0)
activPAL total wear days (#)	10.3 (3.6)	10.1 (3.6)	10.4 (3.5)
activPAL non-wear (hrs)	0.7 (0.5)	0.7 (0.5)	0.7 (0.5)
ActiGraph total wear days (#)[1]	9.4 (3.5)	9.2 (3.5)	9.5 (3.4)
ActiGraph non-wear (hrs)[1]	1.9 (0.9)	2.0 (0.9)	1.7 (0.8)

[1] ActiGraph N=924 overall; n=457 men and n=467 women