



Published in final edited form as:

Br J Sports Med. 2012 May ; 46(6): 436–442. doi:10.1136/bjism.2010.079699.

Identifying sedentary time using automated estimates of accelerometer wear time

Elisabeth A.H. Winkler¹, Paul A. Gardiner¹, Bronwyn K. Clark¹, Charles E. Matthews, (FACSM)², Neville Owen^{1,3}, and Genevieve N. Healy^{1,3}

¹The University of Queensland, School of Population Health, Brisbane, Australia

²Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, USA

³Baker IDI Heart and Diabetes Institute, Melbourne, Australia

Abstract

Purpose—We evaluated the accuracy of three automated accelerometer wear-time estimation algorithms against self-report. Direct effects on sedentary time (<100 counts per minute; cpm) and indirect effects on moderate-to-vigorous physical activity time (MVPA, 1952 cpm) were examined.

Methods—A sub-sample from the 2004/05 Australian Diabetes, Obesity and Lifestyle Study (n=148) completed activity logs and wore accelerometers for a total of 987 days. A published algorithm that allows movement within non-wear periods (Algorithm 1) was compared to one that allows less movement (Algorithm 2), or no movement (Algorithm 3). Implications for population estimates were examined using 2003/04 US National Health and Nutrition Examination Survey data.

Results—Mean difference per day between the criterion and estimated wear time was negligible for all three algorithms (11 minutes), but 95% limits of agreement (LOA) were wide (± 2 hours). Respectively, the algorithms (1, 2 and 3) misclassified sedentary time as non-wear on 31.9%, 19.4% and 18.0% of days and misclassified non-wear time as sedentary on 42.8%, 43.7%, and 51.3% of days. Use of Algorithm 2 (compared with 1) affected population estimates of sedentary time (higher by 20 minutes per day) but not MVPA time. Agreement between Algorithms 1 and 2 was good for MVPA time (mean difference -0.08 , LOA: -2.08 , 1.91 minutes), but not for wear time or sedentary time.

Conclusion—Accelerometer wear time can be estimated accurately on average; however, misclassification can be substantial for individuals. Algorithm choice affects estimates of sedentary time. Allowing very limited movement within non-wear periods can improve accuracy.

Keywords

physical activity; algorithms; validity; AusDiab; NHANES

Corresponding author: Elisabeth Winkler, PhD Cancer Prevention Research Centre School of Population Health The University of Queensland Herston Road Herston, QLD 4006 Australia Telephone: +61 7 336 55526 Fax: +61 7 3365 5540.

Competing interests There are no competing interests to declare.

INTRODUCTION

Accelerometers are increasingly used to provide valid, objective assessments of sedentary time [1] and physical activity [2, 3] in free-living populations, including in large-scale population monitoring.[4, 5] Good correlation or agreement between accelerometer output and activity intensity or energy expenditure has been established,[6] however, other issues need to be considered, particularly how best to determine accelerometer wear time. Study protocols typically specify wearing the accelerometer during waking hours only, and removing for any water-based activities. Thus, wear time varies between days and between participants. Collecting self-report data can assist in determining wear time, but adds to participant and researcher burden. Automated wear-time estimation algorithms are thus especially desirable for large-scale studies, but are not standardised [7] and their accuracy remains to be established.

Automated estimations classify prolonged periods of non-movement (e.g. 60 minutes or 20 minutes at zero intensity) as non-wear time. However, non-moving periods could be either non-wear time or sedentary time. Estimations cannot detect non-wear periods that are shorter than the minimum length of the algorithm criteria (e.g. <60 or <20 minutes) or of a higher intensity than the algorithm allows (e.g. where failure of the accelerometer filtration process or external movement has occurred). To overcome this latter problem, spurious data can be defined and removed [11]; alternatively, some wear-time estimation algorithms [1] allow a limited amount of movement (non-zero counts) to occur within a block of non-wear time.

In addition to the direct impact on sedentary time estimates, achieving the criteria for valid data can be affected by wear time estimation. Thus, all accelerometer measures can be impacted, even if they do not directly include very low-intensity activities in their calculation (e.g. moderate to vigorous physical activity [MVPA] time). Studies often employ a minimum daily wear-time criterion (typically 10 hours) and often require a minimum number of valid days (commonly 4). Potentially, bias can be introduced if the automated estimation process erroneously excludes days on which participants are most sedentary (and possibly also least physically active) and/or excludes the participants who are most sedentary (and least physically active).

Wear-time estimation algorithms produce varied results.[7, 11] Defining non-wear time as all blocks of non-movement 20 minutes, relative to 60 minutes, leads to lower estimates of wear time, lower sedentary time, and higher average counts (a marker of overall activity intensity).[11] One study concluded that 60 minutes was preferable to the 20 minutes criterion, but lacked a referent assessment.[11] No studies that assessed automated estimation against a criterion method in free-living persons were found in database searches of PubMed and ProQuest (June 2010).

We examined the validity and compared the performance for adults of three automated accelerometer wear-time estimation algorithms. We also examined whether misclassification varied with socio-demographic characteristics or bodyweight. Finally, we examined the net impact of misclassification on population estimates of sedentary time, and also on MVPA time.

METHODS

Automated estimation algorithms

All accelerometer data were analysed in SAS 9.1 via a program developed by the National Cancer Institute.[4] The program's wear-time component (Algorithm 1) has been used to

generate US population estimates of sedentary time [1] and physical activity. [5] The program was adapted to produce Algorithms 2 and 3. In view of the apparent advantage of longer (60 minutes) over shorter (20 minutes) criteria for identifying non-wear periods, [11] we focused on algorithms that use the 60 minutes criterion, and compared the effect of the extent of interruptions they allowed within the non wear-period. More/fewer or no interruptions means less/more spurious data but also more/less discarding of true sedentary time. The interruptions allowed by the algorithms were:

- <100 cpm in intensity, with no more than two occurring consecutively (Algorithm 1);
- <50 cpm in intensity, with no more than two occurring per non-wear period (Algorithm 2);
- No interruptions allowed, 0 cpm only (Algorithm 3).

Data sources

We used data from a sub-study of the 2004/05 Australian Diabetes Obesity and Lifestyle Study (AusDiab) [12] and the 2003/04 United States (US) National Health and Nutrition Examination Surveys (NHANES; www.cdc.gov/nchs/nhanes.htm).[13] Both studies obtained ethical approval from relevant parties and written informed consent from participants. For both studies, accelerometers (Actigraph model 7164; Actigraph LLC, Fort Walton Beach, Florida) were set to record in one-minute epochs and participants were instructed to wear the accelerometer on the right hip for seven consecutive days during all waking hours, unless doing water-based activities.

AusDiab accelerometer sub-study

Detailed methods for the AusDiab sub-study are reported elsewhere.[12] Participants (n=202) recorded the times that they wore the accelerometer (on/off times) in an activity log, plus times that they removed it for a period of 15 minutes or more. Data were available for analyses from 148 participants with 987 matching days for accelerometer wear and the activity log. Data were excluded if the accelerometer failed (n=6) or the participant withdrew (n=7), and from all observed days where the accelerometer was not worn or the activity log was poorly completed (n=41 participants, n=336 days). Specifically, data were excluded that were suspicious (e.g. all on/off times occurred on the hour), ambiguous (e.g. cannot be certain that the times recorded referred to AM or PM), or missing time or date information.

Determining non-wear time from activity logs and accelerometer

The beginning and end-times of periods classed as non-wear by the algorithms were extracted and compared to the activity logs. Self-report was used as a criterion of whether participants were wearing the accelerometer or not at any given time; however, allowances were made for imprecision in self-reported times (of up to 30 minutes). The beginning and end times of our criterion non-wear periods were defined in one of three ways depending on the discrepancies between self report times from the log and those derived from the algorithms (Figure 1). If the discrepancy was less than 30 minutes (assumed to be imprecise time reporting), then the times identified by whichever algorithm most closely matched the activity log were used. If the discrepancy was 30 minutes or more, the algorithm was assumed to have failed and the times reported in the activity log were used. Times from the activity log were also used if the algorithms did not detect a reported non-wear period.

We calculated the overall agreement between the algorithms and the criterion measure in their assessment of each epoch as non-wear/wear. Although agreement with our imperfect

criterion should not be interpreted in terms of diagnostic accuracy, we reported our results in terms of sensitivity and specificity instead of the traditionally-used Kappa, as these statistics can distinguish between types of misclassification and are unaffected by the amount of time that is non-wear according to the criterion. The best combination of sensitivity and specificity was judged by highest Youdin's J (sensitivity + specificity - 1). In view of the non-independence of observations, we used a cluster bootstrap method to assess 95% confidence intervals (STATA v11). To explore the consistency of the performance of the algorithms, we also calculated sensitivity and specificity for each day, and report the range observed.

We descriptively compared Algorithms 2 and 3 with Algorithm 1 in terms of the percentage of days in which misclassification occurred and the amount of misclassification for days in which misclassification occurred. Due to skewness, the latter were reported as medians and ranges. Because of the direct substitution between non-wear and sedentary time, we report wear time that was misclassified as non-wear and sedentary time that was misclassified as non-wear. Similarly, a failure to detect non-wear time was reported as non-wear time that was misclassified as sedentary. Using Bland-Altman analysis [14] we also looked at mean differences and 95% limits of agreement (LOA) between estimated and criterion wear time for all three algorithms. To explore whether misclassification was systematic, we looked at bivariate associations of sociodemographic characteristics (age, gender, employment, education, income), and body mass index (BMI, $\text{kg}\cdot\text{m}^{-2}$) with the amount of sedentary time misclassified as non-wear (0, < 1 hr, 1-<2hrs, 2-<3 hrs or 3 hrs+) using Generalised Estimating Equations (in view of the repeated measures).

National Health and Nutrition Examination Surveys (NHANES)

NHANES used a complex, multi-stage design, and its 2003–2004 cycle included an accelerometer component [13] for which all ambulatory participants at least six years of age who attended the Mobile Examination Centre (MEC) were eligible. Since we examined algorithm validity for adults, we focus only on data for adults ($n = 4,741$ MEC participants aged ≥ 20 years). To estimate the potential impact of algorithm choice on population estimates, we compared the algorithm that has to date been used on the NHANES data (Algorithm 1) [1, 5] with the algorithm that agreed most with the criterion. Daily values (minutes) and valid averages were calculated for: wear time, total sedentary time (worn time of intensity <100 cpm), [1] and MVPA time (worn time of intensity ≥ 1952 cpm). [15] As with other analyses of the NHANES accelerometer data, [1, 5] valid averages include only data from monitors that were returned in calibration and from days with ≥ 10 hours of wear-time. We further removed days where excessively high counts were encountered ($> 20,000$ cpm) as these may indicate unreliable data.

The algorithms were compared descriptively in terms of the population averages estimated, and the valid sample of participants and days from which these averages were derived. Population figures were calculated using linearized methods and appropriate sample weights. [16] Population averages were based on all available valid data. Bland-Altman analysis was used to examine agreement between the algorithms in their estimates of: wear-time (minutes per day), MVPA (minutes per day), and sedentary time (minutes per day, percentage of time worn, and as minutes per day with correction for wear time by the residuals method [17]). Agreement was examined for the 3,078 participants with sufficient data for reliable estimates; [18] that is, at least four valid days according to both estimation algorithms.

RESULTS

AusDiab subsample: validity of automated estimation against diaries

Participants from the AusDiab accelerometer sub-study were aged between 30 and 87 years (mean=54.2 years, SD 12.0). The sample included men and women diverse in age range, weight status and socio-demographic characteristics (Table 1).

Agreement between the algorithms and the criterion was excellent for all three algorithms. On average, all three algorithms had high sensitivity and specificity (all >95%); however, the values observed for the days with least sensitivity (0% to 43%) indicate that performance was not consistently good (Table 2). Algorithm 1 had the most sensitivity and the least specificity, Algorithm 3 had the least specificity and the most sensitivity, while Algorithm 2 showed the best balance of both, by a very small amount. When misclassification occurred, sedentary time was misclassified as non-wear for 72 to 78 minutes on average whereas only 44 to 50 minutes on average of non-wear time was misclassified as sedentary time.

Misclassification of sedentary time as non-wear occurred on nearly one third (31.9%) of all observed days using Algorithm 1. Algorithm 2 reduced this to 19.4% without altering the percentage of days on which non-wear time was misclassified as sedentary time. Algorithm 3 also reduced the occurrence of misclassified sedentary time (to 18.0%) but at the same time increased misclassification of non-wear time as sedentary time (51.3% versus 42.8% for Algorithm 1).

For all algorithms, the mean differences between estimated and criterion wear time were negligible (11 minutes) but LOA were wide, spanning approximately two to three hours. The LOA were wider for Algorithm 1 than the others (Table 2). The Bland-Altman plots (Figure 2) further indicated that all three algorithms tended to have more underestimation of wear time than overestimation, and that the most extreme underestimation occurred at lower values of wear time.

Correlates of misclassification

Significantly more misclassification occurred among overweight or obese participants compared with those of normal or underweight BMI (Table 2). Based on the crude percentages (with differences of 5% considered noteworthy) and the GEE analysis, there was otherwise very little evidence that the amount of misclassification of sedentary time by the original method (Algorithm 1) varied across socio-demographic groups (Table 3). Noteworthy, but non-significant differences in crude percentages were seen only for participants who were working full time and those in the 40–49 year age bracket, when compared with their respective counterparts.

NHANES: Effect of algorithm choice on US population estimates

Table 4 compares estimates for the adult US population that result from using Algorithm 2 versus 1. Compared to Algorithm 1, Algorithm 2 generated higher estimates of wear time on average and consequently classified 581 more observed days as valid, and more participants as having valid data (using either a one-day or four-day criterion). Mean population estimates of sedentary time were higher when using Algorithm 2 than Algorithm 1, even when correcting for wear time or when examining sedentary time as a percentage of worn time. The magnitude of the difference in estimates was modest: approximately twenty minutes or one percent. By contrast, estimates of average time spent in MVPA were not affected by choice of wear-time algorithm, with the median [minimum, maximum] for Algorithms 1 and 2 being 18.0 [0, 215.5] and 17.6 [0, 215.5] minutes, respectively.

Figure 3 shows the Bland-Altman plots of agreement between Algorithms 1 and 2 for wear time, sedentary time, and MVPA. Agreement between the algorithms was poor for wear time and for all measures of sedentary time. For wear time, sedentary time and percentage sedentary time, some heteroscedasticity was evident (i.e. the amount of misclassification increased with the mean), so the Bland-Altman plots are displayed for the log transformed data. The back-transformed mean difference (1.02) and LOA (0.94, 1.11) indicate that Algorithm 2 produced estimates of wear time that were 2% higher on average than Algorithm 1, and anywhere between 6% lower and 11% higher for 95% of people. Relative to Algorithm 1, Algorithm 2 also generated higher estimates of sedentary time (+4%, LOA -8%, +17%) and percentage worn time spent sedentary (+2%, LOA -3%, +7%). For corrected sedentary time, the mean difference was 20.4 minutes (LOA: -12.7, 53.5). While mean differences were small, the wide LOA and large outliers show the two algorithms do not yield equivalent estimates of sedentary time or wear time. In contrast, there was good agreement for MVPA estimated by the two algorithms (mean difference -0.08, LOA: -2.08, 1.91 minutes).

DISCUSSION

All three wear-time estimation algorithms showed excellent agreement with the criterion. However, for some days, the algorithms had poor sensitivity and specificity and generated estimates that were incorrect by several hours. In addition to problems with short periods of spurious data, long bouts of time were often misclassified. Reducing the amount of movement permitted within non-wear time (Algorithm 2) reduced the misclassification of sedentary time as non-wear without affecting the detection of non-wear time. By contrast, allowing no movement within non-wear periods (Algorithm 3) had a similar effect on misclassification of sedentary time, but at a cost of failing to detect 'true' non-wear time. One problem for all algorithms was true non-wear periods shorter than 60 minutes, which commonly occurred when the accelerometer was removed after 11pm. Overall, allowing very limited interruptions (i.e. <50 cpm, no more than two per non-wear period) appeared optimal, although the benefit over no interruptions (i.e. 0cpm) was only slight.

It is possible that misclassification resulting from wear-time estimation algorithms is differential; however, our analysis of this issue was limited by the small sample. The association of misclassification with BMI suggests that the algorithms may perform better for normal weight adults than for adults who are overweight or obese. One possible explanation is under-detection of movement for overweight or obese persons on whom accelerometers tend to sit at the wrong angle.[19]

Algorithm 1 was developed for surveillance purposes to estimate population means, primarily MVPA, in large-scale population studies, such as NHANES.[4] The use of this versus Algorithm 2, which has also been used in analysing the NHANES data,[20] affected the estimates of total wear time, and thus the number of valid days and participants included in analysis. This did not translate into an impact on MVPA time, and the impact on population estimates of sedentary time was modest. However, agreement in sedentary time estimates across algorithms was poor, even when wear time was supposedly 'controlled' either via the residuals method [17] or by conversion to percentages. Overall, the implications for research studies are that algorithm choice may be of little importance when obtaining descriptions of population levels, but studies aiming to examine factors associated with sedentary time or to detect within-person change may be affected by misclassification.

This study adds to the relevant research literature [7, 11] by using a referent assessment method in free-living participants. Our findings complement those of a recent laboratory study, which established that several automated estimations with long minimum durations

(60 minutes, or 90 minutes) perform adequately, particularly when allowing interruptions. [20] There is no gold-standard criterion for free living populations and although our criterion was not ideal, it was unlikely to favour any particular algorithm and was likely adequate for comparison purposes. The sample was not fully population representative,[12] thus, generalizability is not certain. Other populations, including children, adolescents, young adults, and people from various racial and ethnic backgrounds should also be examined. Given the rapid advances in accelerometer technology (which includes the collection of large amounts of raw data), appropriate algorithms also need to be validated for different epoch lengths and also for accelerometers using dual-axis or triaxial modes. This study used three 'cut-points' for allowable interruptions; further exploration may reveal better algorithms.[21]

Conclusions and recommendations

Automated accelerometer wear time estimation has acceptable validity for adults for many purposes, with the better results achieved by allowing non-wear periods to contain very limited movement (Algorithm 2) rather than extensive interruptions (Algorithm 1). However, further achievable improvements are needed, particularly when accurate sedentary time measures are necessary, such as identifying and removing spurious data [11] and reducing the failure to detect short non-wear periods (<60 minutes) by allowing non-wear bouts to continue past midnight.[2] Estimation algorithms are a time-efficient and feasible option for large-scale population monitoring, but associated measurement error in sedentary time is substantial and needs consideration.

Statement

“The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in BJSM editions and any other BMJ PGL products to exploit all subsidiary rights, as set out in our licence (<http://group.bmj.com/products/journals/instructions-for-authors/licence-forms/>).”

Acknowledgments

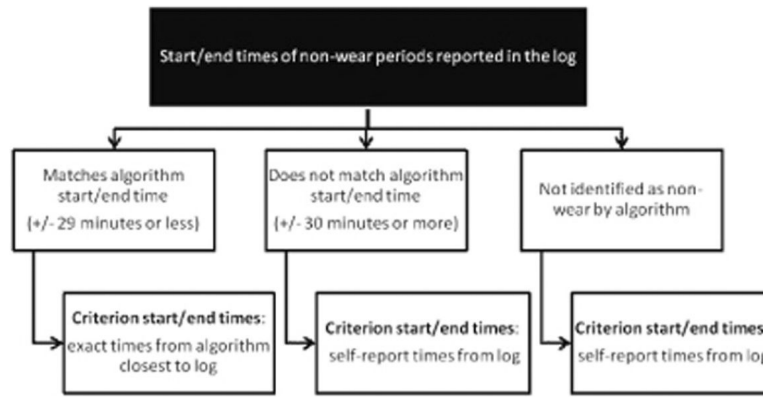
Clark, Gardiner, Healy, Winkler and Owen are supported by a Queensland Health Core Research Infrastructure grant and by NHMRC Program Grant funding (#569940). Gardiner is supported by a Heart Foundation of Australia (# PP 06B 2889). Clark is supported by an Australian Post-graduate Award. Healy is also supported by a NHMRC (#569861) / National Heart Foundation of Australia (PH 374 08B 3905) Postdoctoral Fellowship. Data from the AusDiab study were used (for full acknowledgments of the many funding sources, see [12]). Data used in this study (NHANES) were collected by the National Center for Health Statistics, Centers for Disease Control and Prevention.

Funding This study was supported by funding from a Queensland Health Core Research Infrastructure grant and the National Health and Medical Research Council of Australia (#569940)

References

1. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, et al. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol.* 2008; 167:875–81. [PubMed: 18303006]
2. Matthews CE. Calibration of accelerometer output for adults. *Med Sci Sports Exerc.* 2005; 37:S512–22. [PubMed: 16294114]
3. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc.* 2005; 37:S523–30. [PubMed: 16294115]
4. National Centre for Health Statistics. Centres for Disease Control and Prevention. NHANES 2003–2004 Data Documentation Exam Component: Physical Activity Monitor (PAXRAW_C). US

- Department of Health and Human Services; Hyattsville, MD: 2006. http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/paxraw_c.pdf
5. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc.* 2008; 40:181–8. [PubMed: 18091006]
 6. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc.* 2005; 37:S501–11. [PubMed: 16294113]
 7. Masse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, et al. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Med Sci Sports Exerc.* 2005; 37:S544–54. [PubMed: 16294117]
 8. Hamilton MT, Healy GN, Dunstan DW, Zderic TW, Owen NT. Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior. *Current Cardiovascular Risk Report.* 2008; 2:292–8.
 9. Owen N, Healy GN, Matthews CE, Dunstan DW. The population health science of sedentary behavior: too much sitting. *Exerc Sport Sci Rev.* 2010; 38:105–13. [PubMed: 20577058]
 10. Katzmarzyk PT, Church TS, Craig CL, Bouchard C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc.* 2009; 41:998–1005. [PubMed: 19346988]
 11. Evenson KR, Terry JW Jr. Assessment of differing definitions of accelerometer nonwear time. *Res Q Exerc Sport.* 2009; 80:355–62. [PubMed: 19650401]
 12. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Objectively measured light-intensity physical activity is independently associated with 2-h plasma glucose. *Diabetes Care.* 2007; 30:1384–9. [PubMed: 17473059]
 13. Centers for Disease Control and Prevention (CDC). National Center for Health Statistics (NCHS). Centers for Disease Control and Prevention; 2003–2004, 2005–2006. U.S. Department of Health and Human Services; Hyattsville, MD: National Health and Nutrition Examination Survey Data. [updated 2003–2004, 2005–2006]; Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention]. <http://www.cdc.gov/nchs/nhanes.htm>
 14. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986; 1:307–10. [PubMed: 2868172]
 15. 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]
 16. National Centre for Health Statistics. Centres for Disease Control and Prevention. Analytic and Reporting Guidelines: The National Health and Nutrition Examination Survey (NHANES). US Department of Health and Human Services; Hyattsville, MD: 2005. http://www.cdc.gov/nchs/data/nhanes/nhanes_03_04/nhanes_analytic_guidelines_dec_2005.pdf
 17. Willett WC, Howe GR, Kushi LH. Adjustment for total energy intake in epidemiologic studies. *Am J Clin Nutr.* 1997; 65:1220S–8S. discussion 9S–31S. [PubMed: 9094926]
 18. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc.* 2005; 37:S531–43. [PubMed: 16294116]
 19. Corder K, Brage S, Ekelund U. Accelerometers and pedometers: methodology and clinical application. *Curr Opin Clin Nutr Metab Care.* 2007; 10:597–603. [PubMed: 17693743]
 20. Healy GN, Matthews CE, Dunstan D, Winkler EAH, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06. *Eur Heart J.* Published Online First 4 Nov 2010. doi 10.1093/eurheartj/ehq451.
 21. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. *Med Sci Sports Exerc.* Published Online First 23 June 2010. doi:10.1249/MSS.0b013e3181ed61a3.



* Note: periods identified by the algorithms as non-wear that are not reported in the log are not treated as part of the criterion non-wear time

Figure 1.
Determination of start and finish times of criterion non-wear periods

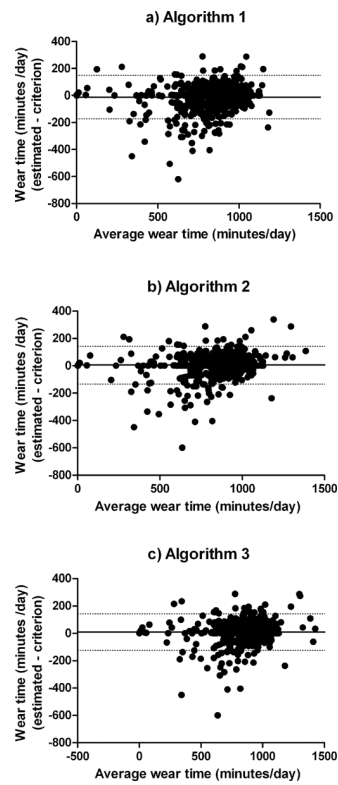


Figure 2. Bland Altman plots of agreement between estimated and criterion wear time for Algorithm (a) 1, (b) 2, and (c) 3 among AusDiab participants (n=148)

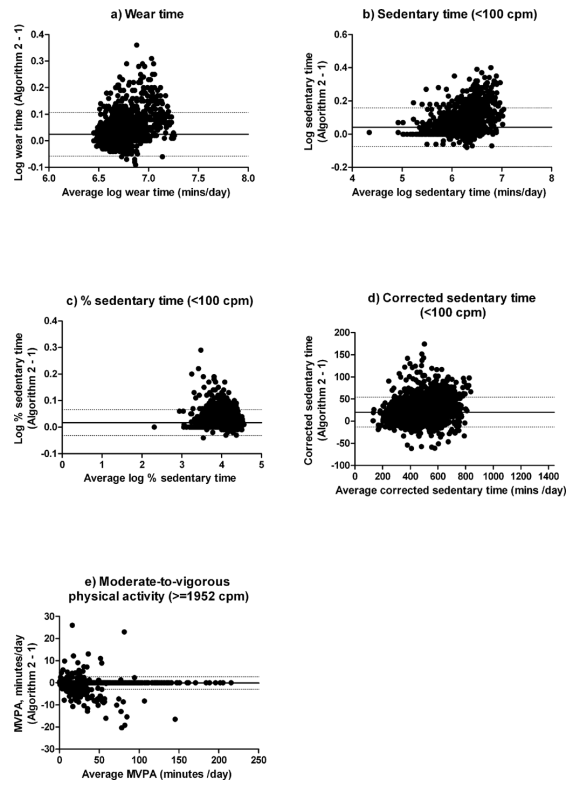


Figure 3. Bland-Altman plots of agreement between Algorithms 1 and 2 in (a) wear time, (b–d) sedentary time, and (e) MVPA time (NHANES 2003–2004, n=3,078)

Table 1

Characteristics of participants in the accelerometer sub-study of the Australian Diabetes, Obesity and Lifestyle (AusDiab) study

	AusDiab Sub-study Sample (n=148)% (n)	
Age (years)		
	30–39	11.5% (17)
	40–49	26.4% (39)
	50–59	32.4% (48)
	60+	29.7% (44)
Gender		
	women	58.1% (86)
	men	41.9% (62)
Body Mass Index (kg/m²)		
	<25	30.4% (45)
	25 to <30	49.3% (73)
	30+	19.6% (29)
Employed full-time		
	Yes	54.7% (81)
	No	45.3% (67)
Education		
	High school or less	47.3% (70)
	Post school training	52.7% (78)
Household Income		
	\$1500 per week (p.w.)	35.1% (52)
	<\$1500 p.w. or missing	64.9% (96)
Current smoking		
	Yes	2.0% (3)
	No	98.0% (145)

Table 2

Comparison of automated wear-time estimation algorithms¹ in classifying time as sedentary or non-wear against activity log

	Algorithm 1	Algorithm 2	Algorithm 3
Sensitivity (%)			
Overall (95% CI)	95.61 (95.16, 96.05)	95.43 (94.96, 95.89)	95.10 (94.64, 95.56)
Worst day	43.16	0	0
Best day	100	100	100
Specificity (%)			
Overall (95% CI)	95.61 (95.14, 96.09)	97.40 (96.99, 97.81)	97.67 (97.32, 98.03)
Worst day	19.93	20.11	19.93
Best day	100	100	100
Sedentary time classed as non-wear			
Days with misclassification (%)	31.9%	19.4%	18.0%
Minutes, Median (Min, Max) ²	78 (6, 620)	75 (26, 600)	72 (1,600)
Non-wear time classed as sedentary			
Days with misclassification (%)	42.8%	43.7%	51.3%
Minutes, Median (Min, Max) ²	49 (1,329)	50 (1,338)	44 (1,329)
Mean difference (95% limits of agreement) estimated – criterion wear time	-11 (-172, 149)	5 (-134, 143)	9 (-124, 142)

¹The algorithms define non-wear periods as non-movement (i.e. 0cpm and any allowed interruptions) of 60 minutes duration: Algorithm 1 permits 2 consecutive active counts of maximum intensity 99 cpm; Algorithm 2 permits 2 active counts of maximum intensity 49 cpm; and, Algorithm 3 allows no interruptions (requires all counts to be 0 cpm).

²Median (minimum, maximum) minutes misclassified (against criterion) on days where misclassification occurred.

Table 3

Associations of sociodemographic characteristics with amount of sedentary time misclassified as non-wear time (Algorithm 1)¹ in 148 accelerometer sub-study participants of the Australian Diabetes, Obesity and Lifestyle study

	Hours misclassified (%)				P
	0	<1	1 to <2	2 to <3	
Age group					0.174
30–39 (n=17)	64.2	7.3	13.8	5.5	9.2
40–49 (n=39)	74.0	8.1	10.1	4.3	3.5
50–59 (n=48)	68.0	4.2	18.1	3.9	5.7
60+(n=44)	64.7	3.8	19.0	8.3	4.2
Gender					0.694
Women (n=86)	68.0	6.2	16.1	4.7	5.0
Men (n=62)	68.5	4.4	15.4	6.6	5.1
Body Mass Index (kg/m ²) ²					0.048
Normal/underweight <25 (n =45)	72.1	8.1	15.3	2.3	2.3
Overweight/obese ≥ 25 (n =102)	66.4	4.3	16.1	7.0	6.2
Employed fulltime					0.076
Yes(n=81)	63.7	6.2	19.1	6.4	4.6
No (n=67)	72.0	4.9	13.0	4.7	5.5
Education					0.283
High school or less (n=70)	70.3	4.7	14.7	6.0	4.3
Post high school (n=78)	66.3	6.1	16.8	5.0	5.7
Household income					0.760
> \$1500 /week (n=52)	68.0	5.6	15.5	4.7	6.2
\$1500/ week or not reported (n=96)	68.3	5.4	15.9	5.9	4.5

¹Table presents crude % of observed days with each level of misclassification, and p-value for association (Generalized Estimating Equations) and n participants (not days)

²N=1 missing data for Body Mass Index

Table 4

Comparison of accelerometer-derived estimates for the US adult population 20 years (NHANES 2003–2004) obtained with different wear-time estimation algorithms

	Algorithm 1 ¹	Algorithm 2 ²
N		
Valid days	19,927	20,508
Participants with 1+ day valid data	3,785	3,807
Participants with 4+ days valid data	3,078	3,156
Mean (Standard Deviation) ³		
Wear time, <i>minutes/day</i>	839 (103)	855 (117)
Sedentary time, <i>minutes/day</i>	473 (122)	494 (139)
Corrected sedentary time, <i>minutes/day</i> ⁴	469 (103)	492 (108)
Sedentary time, % <i>worn time</i>	56.3 (12.5)	57.4 (12.8)

¹ Defines non-wear periods as 60 minutes at zero intensity, including up to two consecutive minutes of maximum intensity 99 counts per minute

² Defines non-wear periods as 60 minutes at zero intensity, including up to a total of two minutes of maximum intensity 49 cpm

³ Mean (Standard Deviation) are weighted (for sampling and non-response) to the US population

⁴ corrected for wear time by the residuals method