Accelerometer assessed physical activity in epidemiology: Are monitors equivalent?

Short title: Equivalence of wrist-worn accelerometers

AV Rowlands*,[1,2,3] EM Mirkes,[4] T Yates,[1,2] S Clemes,[5] M Davies,[1,2] K Khunti,[1,2,6] CL Edwardson[1,2]

*Corresponding author

1. Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, UK
2. NIHR Leicester Biomedical Research Centre, UK
3. Alliance for Research in Exercise, Nutrition and Activity (ARENA), Sansom Institute for Health Research, Division of Health Sciences, University of South Australia, Adelaide, Australia
4. Department of Mathematics, University of Leicester, UK
5. School of Sport, Exercise and Health Sciences, University of Loughborough, UK
6. NIHR Collaboration for Leadership in Applied Health Research and Care East Midlands, Leicester General Hospital, UK

Corresponding author: Alex Rowlands, Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, LE5 4PW, UK. alex.rowlands@le.ac.uk. Tel: +44 116 258 8595

Abstract

Purpose: Accelerometers are increasingly being used to assess physical activity in large-scale surveys. Establishing whether key physical activity outcomes can be considered equivalent between three widely-used accelerometer brands would be a significant step towards capitalising on the increasing availability of accelerometry data for epidemiological research. **Methods:** Twenty participants wore a GENEActiv, Axivity AX3 and ActiGraph GT9X on their non-dominant wrist and were observed for two-hours in a simulated living space. Participants undertook a series of seated and upright light/active behaviours at their own pace. All accelerometer data were processed identically using open-source software (GGIR) to generate physical activity outcomes (including average dynamic acceleration (ACC) and time within intensity cut-points). Data were analysed using pairwise 95% equivalence tests (±10% equivalence zone), intra-class correlation coefficients (ICC) and limits of agreement. **Results:** The GENEActiv and Axivity could be considered equivalent for ACC (ICC=0.95, 95% confidence interval (CI) 0.87 to 0.98), but ACC measured by the ActiGraph was approximately 10% lower (ICC: GENEActiv/ActiGraph 0.86, 95% CI 0.56 to 0.95; Axivity/ActiGraph 0.82, 95% CI 0.50 to 0.94). For time spent within intensity cut-points, all three accelerometers could be considered equivalent to each other for over 85% of outcomes (ICC$\geq$0.69, lower 95% CI$\geq$0.36), with the GENEActiv and Axivity equivalent for 100% of outcomes (ICC$\geq$0.95, lower 95% CI$\geq$0.86). **Conclusions:** GENEActiv and Axivity data processed in GGIR are largely equivalent. If comparing GENEActiv or Axivity to the ActiGraph, time spent within intensity cut-points has good agreement. These findings can be used to inform selection of appropriate outcomes if comparing outputs from these accelerometer brands.

Key words: ActiGraph, Axivity, GENEActiv, GGIR, GT9X, Link

Introduction

Pooling data from multiple surveys has facilitated robust and generalizable estimates of risk factors (e.g. smoking, overweight) for cardiovascular events, cardiovascular disease and mortality that have informed clinical and public health practice (1, 2). Physical inactivity is also an established risk factor for chronic disease (3) but, until recently, physical activity measurement in epidemiological and surveillance studies has relied on self-report. This is imprecise which has complicated comparison or aggregation of data across populations.

Over the past few years, it has become feasible to move to large-scale objective measurement of physical activity with wrist-worn accelerometers worn 24 hours a day, seven days a week. As the latest generation of accelerometers measure acceleration in SI units, there is great potential for aggregation of measures of physical activity into very large multinational databases. Data harmonisation would facilitate a step-change in our ability to: a) compare prevalence or levels of activity/inactivity across populations, b) quantify dose-response associations between activity and health and c) to identify the factors that impact on these associations. A key advantage would be the ability to address these questions in very large samples across countries and/or populations for a wide range of health outcomes.

There are three brands of accelerometers providing acceleration data in SI units being used in large surveys: the Axivity, ActiGraph, and GENEActiv. For example, UK Biobank, a large-scale prospective epidemiological resource containing baseline phenotypic and genotypic data on 500,000 participants, has recently used the Axivity wrist-worn accelerometer in over 100,000 participants (4) and the Breakthrough Generation Study (5) has used the Axivity on 4,800 women to date. The US National Health and Nutrition Examination Survey (NHANES) used the ActiGraph wrist-worn accelerometer in cycles 2011-2012 and 2013-2014 (approximately 9,000-10,000 participants examined per cycle). The Pelotas Birth Cohort (6), the Melbourne Child Health Checkpoint (7), the Cork Children's Lifestyle Study (8), and the British Whitehall II study (9) used the GENEActiv in approximately 10,000, 4,000, 1,000 and 3,750 participants, respectively.

Before outcomes from studies using different brands of accelerometer can be pooled into multinational databases, it is necessary to demonstrate comparability of data outputs between brands.

Pooling accelerometer data from these studies is more viable now than ever before. Earlier accelerometers processed data into counts using proprietary algorithms (10); this complicated the interpretation and comparison of data from studies using different devices (11). However, evidence suggests that, despite the outcome being non-proprietary accelerations, data may not be equivalent between brands (12, 13, 14, 15, 16).

To facilitate transparent processing of these raw data the generation and use of open-source resources is encouraged (16, 18). GGIR is an open-access package in R [http:/cran.r-project.org] that can be used to process and analyse raw accelerations from the GENEActiv and the Actigraph using identical methods (17, 18, 19, 20). As it is open-source and an efficient method for processing and analysing raw data to obtain the key outcomes required for characterising habitual physical activity it has been used widely to analyse GENEActiv and ActiGraph data (e.g. 6, 9, 20, 21, 22, 23, 24, 25, 26). We have shown that the outcomes from GENEActiv and ActiGraph processed through GGIR are broadly comparable, although comparisons at lower magnitudes of acceleration may be problematic (20). More focussed assessment of specific sedentary and light activities is needed to examine this further.

The Axivity and GENEActiv measured similar magnitudes of accelerations when tested in a mechanical shaker (27), but there are no data comparing the Axivity and ActiGraph or comparing physical activity outcomes from the Axivity to the GENEActiv or ActiGraph during actual wear. This paper will introduce a function that converts raw Axivity files to a format that facilitates identical processing and analysis of Axivity files in GGIR. It is important to consider the accelerometer, processing and analysis together when establishing whether outcomes can be considered equivalent or not, as each of these steps can impact on the final outcome variable (16). Given the widespread use of a) these accelerometers to assess physical activity and b) GGIR for processing and analysis of the data, establishing which outcomes can be considered equivalent between

accelerometers would be a significant step towards capitalising on the increasing availability of accelerometry data in epidemiological research.

We aimed to establish whether the Axivity, GENEActiv and ActiGraph GT9X result in equivalent physical activity outcomes when data are processed and analysed identically with GGIR. We considered the magnitude of acceleration, time spent above and below published intensity cut-points and the distribution of time across an incremental acceleration range, both for specific activities and for the duration of a simulated free-living situation in a laboratory.

Methods

A convenience sample of 20 adult participants was recruited from Loughborough University and University of Leicester (staff and students) via email and word of mouth. All participants provided written informed consent, and the study was approved by the Ethics Committee of Loughborough University. Data were collected between January and April 2016.

Height and body mass were measured to the nearest 0.5 cm and 0.1 kg, respectively. Each participant was fitted with an Axivity AX3 (Axivity Ltd, Newcastle, UK), GENEActiv (ActivInsights Ltd, Cambridgeshire, UK) and ActiGraph GT9X Link (ActiGraph LLC, Pensacola, FL, USA) on their non-dominant wrist. This was part of a larger study employing further activity monitors; to reduce the need for multiple wrist-straps the Axivity was taped to the GENEActiv and the ActiGraph was worn immediately proximal to the GENEActiv and Axivity. In our previous studies comparing output from the GENEActiv and the ActiGraph, differences in output were consistent whether the monitors were taped together (12, 28) or worn adjacent on the wrist (20).

Protocol

The study took place in a laboratory mocked up as a living space with items of furniture and lasted approximately two hours with participants tested in groups of two or three. Participants were asked to undertake a series of seated activities (watching television, using the computer, eating and reading) in any manner, at their own pace and in any order they chose. Minimal instructions were given with participants simply asked to ensure they undertook each activity at least once and for a minimum of 10 min. The aim was to mimic free-living postures/behaviours as closely as possible. A researcher observed the participants continuously and recorded their activity and posture (seated or standing) minute-by-minute. Participants then performed six upright light and active behaviours in a randomised order for five minutes each: standing still, standing up to work on a computer, dusting, sweeping, washing pots and walking.

Accelerometers

The Axivity AX3, GENEActiv and ActiGraph GT9X Link (from herein: Axivity, GENEActiv and ActiGraph) are triaxial accelerometry-based activity monitors with a dynamic range of +/- 8 *g*, where *g* is equal to the Earth's gravitational pull. All accelerometers were set to capture and store accelerations at their maximum sampling frequency of 100 Hz. The "idle sleep mode" in the ActiGraph software (Actilife v. 6.13.0) was disabled. Axivity data were downloaded using OmGui open-source software (OmGui Version 1.0.0.28, Open Movement, Newcastle University, UK) and saved in raw format as .cwa files. GENEActiv data were downloaded using GENEActiv PC software version 2.2 and saved in raw format as .bin files. ActiGraph data were downloaded using ActiLife v. 6.13.0, saved in raw format as .gt3x files and converted to .csv format for data processing.

Data processing and outcome measures

Raw .cwa Axivity files were converted to .bin files using our new function, 'AccelerometerCWA2BINConverter', to enable analysis in GGIR, thus identical processing and analysis as GENEActiv and ActiGraph files. This function includes resampling of the data to a standard frequency as specified in the header of the cwa file; this is necessary because the Axivity sample frequency is unreliable

and varies over time. This function is available at: https://github.com/Mirkes/AccelerometerCWA2BINConverter.

All accelerometer files were analysed with R-package GGIR version 1.4 in R (http://cran.r-project.org) (18, 19). This included auto-calibration using local gravity as a reference (19), detection of sustained abnormally high values, calculation of the average magnitude of dynamic acceleration (i.e. resultant vector magnitude, corrected for gravity and expressed as Euclidean Norm Minus One (ENMO) in milli-gravitational units (m*g*) averaged over 1-second epochs) and generation of participant-specific csv files with accelerometer output in 1 s epochs. Where insufficient non-movement periods were available for auto-calibration we used back-up calibration coefficients derived from free-living data collected with the same accelerometer unit.

A number of pre-specified outcomes were assessed: average acceleration (m*g*); % time accumulated within cut-points for sedentary, light and moderate-to-vigorous physical activity (MVPA); distribution of time across acceleration levels in 40 m*g* resolution (0-40 m*g*, 40-80 m*g*... >200 m*g*). Mean acceleration was calculated for each activity separately, seated activities only, upright activities only and over the total time period (i.e. including all activities and transitions, approximately two hours). Cut-point and distribution of time outcomes were calculated for seated activities only, upright activities only and over the total time period.

Time spent sedentary and time in light activity were calculated using cut-points of 30, 40 and 50 m*g* to enable evaluation of the equivalency of a range of sedentary cut-points (14, 15). The accuracy of these cut-points was further calculated for seated activity (i.e. % seated time classified as sedentary) and for upright activity (i.e. % upright time classified as not sedentary). MVPA was calculated using an acceleration cut-point of 100 m*g* (13). All outcomes were calculated for all three devices.

Data analysis

Descriptive statistics (mean (SD)) were calculated for all outcomes. The level of agreement between outputs from the three brands of accelerometer was determined pairwise using intraclass correlation coefficients (ICC, single measures, absolute agreement) with 95% confidence intervals (CI) and limits of agreement (LoA) (29). We used pairwise 95% equivalence tests to determine whether the 95% CI for the mean of one accelerometer fell within a proposed equivalence zone of the second accelerometer (30). We selected ±10% of the mean as our proposed equivalence zone as in previous studies comparing activity monitors (31, 32). Equivalence results are presented with the reference accelerometer selected according to the following hierarchy: GENEActiv, Axivity, ActiGraph. However, as no accelerometer can be considered the gold standard this was arbitrary, consequently all equivalence analyses were repeated with the alternate accelerometer in each pairing selected as the reference to test whether this affected the conclusions.

Based on our previous work comparing GGIR physical activity outputs from the GENEActiv and the ActiGraph (20), we anticipated the standard deviation of the differences between the log transformed outputs from the two accelerometers would be less than 0.05 m*g* and the ratio between the mean outputs from the two accelerometers would be within 1 ± 0.05. Log transforming the data enables hypotheses about ratios to be analysed in terms of differences. Given this effect size, using Minitab (v17), we determined that a sample size of 12 was required to provide 90% power (alpha = 0.05) to conclude that the difference between physical activity outcomes from a pair of accelerometers was within 10% of the mean when this was in fact true.

Descriptive statistics, ICCs and LoA were conducted in IBM SPSS Statistics v22.0. Equivalency testing and power analyses were carried out in Minitab (v17). Alpha was set at 0.05.

Results

Twenty participants (13 females, 7 males; age (mean (SD)): 23.2 (5.9) y; body mass index: 25.2 (3.6) $kg.m^{-2}$) took part. The Axivity accelerometers were unavailable for one testing session (three participants), one GENEActiv file and two ActiGraph files did not process, and one participant did not complete the seated

activities. Therefore the sample included 17 Axivity, 19 GENEActiv, and 18 ActiGraph files for the overall time period and upright activities and 16, 18 and 18, respectively, for seated activities. Pairwise N's were 16 for GENEActiv/Axivity, 17 for GENEActiv/ActiGraph and 15 for Axivity/ActiGraph, exceeding the sample size of 12 required to achieve 90% power. Participant characteristics were similar for included and excluded files. Running the analyses with listwise deletion did not change the results (N=14), so pairwise analyses were retained to maximise sample sizes.

TABLE 1 HERE

The total testing period lasted approximately 2 h and included 1 h 20 min of seated activities, 24 min of upright activities and 16 min of transition between activities. Descriptive statistics are presented in Table 1 and agreement statistics in Table 2.

*Acceleration (*m*g)*

Agreement between pairs of accelerometers was largely good, with ICCs of 0.82-0.95 for the total period, 0.73-0.85 for all seated activities combined (except the Axivity/ActiGraph pairing, ICC = 0.59) and 0.75-0.97 all upright activities combined, Table 2. Results for specific activities suggested that the poorest agreement between pairs of accelerometers was obtained when using the computer (sitting or standing) and reading. Although the mean biases between pairs of accelerometers tended to be low, some of the 95% LoA were relatively large, particularly for the Axivity/ActiGraph pairing.

Overall, the highest agreement was between the GENEActiv and Axivity devices (ICC = 0.95, 95% CI: 0.87, 0.98; mean bias = -0.1 m*g*, 95% limits = ±6.8 m*g*), which could be considered equivalent (i.e. the 95% CI for the mean of the Axivity fell within ±10% of the mean of the GENEActiv) for the total period (Figure 1a), upright activities combined (Figure 1a) and four of the ten specific activities (not for seated activities, standing still or standing computer, Figures 1b and 1c). In contrast, the ActiGraph could not be considered equivalent to either the GENEActiv or the Axivity at all (Figures 1b and 1c).

TABLE 2 HERE

FIGURE 1 HERE

Sedentary, light intensity and MVPA cut-points

Agreement between accelerometers was good for classification of sedentary time (ICC ≥0.84; mean bias ≤±1.5 percentage points) for all accelerometer pairings, irrespective of cut-point, and all could be considered equivalent (Table 3, Figure 1d). The 95% LoA were narrowest for the 50 m*g* cut-point (Table 3). All accelerometers could also be considered equivalent for classification of light intensity activity (Figure 1d), with pairings weakest for the 30 m*g* cut-point (ICC ≥0.69). As for classification as sedentary, the highest ICC's (≥0.83), lowest mean biases (≤±0.3 percentage points) and narrowest 95% limits of agreement were found for the 50 m*g* cut-point. Although ICCs for classification of MVPA were high (≥0.84) LoA were large (±2.3 – 4.1 percentage points) relative to the means (approx. 13%) and only the GENEActiv/Axivity pairing could be considered equivalent (Figure 1d).

Although, all accelerometer pairings could be considered equivalent for six out of seven outcomes, the highest agreement was found for the GENEActiv and Axivity device, irrespective of cut-point (ICC >0.95, mean bias ≤±0.4 percentage points, 95% limits <±5 percentage points).

TABLE 3 HERE

Accuracy of sedentary cut-points

Accuracy of classification of seated activities as sedentary was high for all cut-points (>87%), Table 1, and equivalent between accelerometers (Table 4, Figure 1e). However, the upright activities were misclassified approximately 60% of the time with the 30 m*g* cut-point and only equivalent for the GENEActiv/Axivity pairing. When applying the 40 m*g* or 50 m*g* cut-point, the upright activities were still misclassified approximately 50% of the time, but all accelerometers could be considered to have equivalent accuracy.

Again, the highest agreement was found for the GENEActiv and Axivity device, irrespective of cut-point (ICC >0.83, mean bias ≤±1.3 percentage points, 95% limits ≤±5.5 percentage points).

TABLE 4 HERE

Incremental acceleration ranges

Agreement was high for all accelerometer pairings for % time in 40 m*g* increments (ICC ≥0.71, mean bias <±1.6 percentage points). As with average m*g* and time spent in cut-point categories, the strongest agreement was for the GENEActiv and Axivity pairing (≥0.93, mean bias ≥±0.25 percentage points), which could be considered equivalent for all categories, except 160-200 m*g,* which was borderline equivalent (Table 5, Figure 1f). The GENEActiv/ActiGraph and Axivity/ActiGraph pairings could be considered equivalent in the 0-40, 40-80 and 80-120 m*g* categories (Figure 1f).

The highest agreement was found for the GENEActiv and Axivity device, irrespective of 40 m*g* range (ICC >0.93, mean bias ≤±0.25 percentage points, 95% limits ≤±4.0 percentage points).

TABLE 5 HERE

Re-running the equivalence analyses with the alternate accelerometer selected as the reference did not impact on whether accelerometer brands were considered equivalent or not, except two cases which were previously borderline, but could be considered equivalent when the alternate accelerometer was the reference. These were: 1) accuracy of the sedentary cut-point, 30 mg, for classification as upright by the ActiGraph and the Axivity (Ratio: 0.93 (95% CI: 0.90, 1)); 2) time spent between 160 and 200 mg by the Axivity and the GENEActiv (Ratio: 0.96, (95% CI: 0.90, 1.03)), Figures 1e-f, Tables 2-5.

Discussion

The Axivity, GENEActiv and ActiGraph wrist-worn raw acceleration accelerometers are widely used to assess physical activity in large-scale surveys (16); consequently the generation of equivalent physical activity outcomes from these tools would aid epidemiological comparisons. We used an open-source software package (GGIR) to identically process and analyse data from the three accelerometer brands to establish the degree of equivalence and agreement across specific activities, types of activities and the entire semi-structured pseudo free-living period. To the best of our knowledge, this is the first time physical activity outcomes from the Axivity accelerometer have been compared to those from the GENEActiv or the ActiGraph.

The GENEActiv and Axivity accelerometers had excellent equivalence and agreement across the majority of outcome measures including acceleration (m*g*). However, acceleration was around 11% lower in the ActiGraph data. Despite this, time spent in sedentary and light intensity could be considered equivalent between all three accelerometers, irrespective of cut-point employed. The GENEActiv and Axivity could also be considered equivalent for MVPA; agreement with the ActiGraph was also high for MVPA, although not within the proposed 10% equivalence zone. The higher agreement between accelerometer brands evident when considering variables derived from acceleration, i.e. time accumulated in cut-point categories or acceleration ranges, rather than the acceleration itself, is consistent with our previous research comparing the GENEActiv and ActiGraph GT3X+ (20). The high correspondence between accelerometer brands for intensity categories is important as quantities of time accumulated within sedentary, light and moderate-to-vigorous intensity cut-points are arguably the most commonly cited physical activity outcome measures.

The observation protocol we employed enabled us to comprehensively evaluate the accuracy of three sedentary cut-points (14, 15). Sedentary/light cut-points of 40-50 m*g* were the most accurate at classifying sedentary and upright time and had higher agreement between accelerometers than 30 m*g*. However, all cut-points, irrespective of accelerometer brand were poor at classifying upright time. The inability to differentiate between postures using magnitude of acceleration alone has been previously reported (14, 15) with use of further features from the acceleration signal recommended for classification of posture (e.g. 33, 34). When

using cut-points or the distribution of time across acceleration ranges, it is perhaps best to think of classifying a spectrum of inactive to active time rather than referring to sedentary time, which infers posture (35).

The equivalency results suggest that equivalence is worse at low accelerations (seated activities), for time spent at accelerations greater than 120 m*g,* and for time spent in MVPA, most notably for the ActiGraph. However, as the proposed equivalence zone is ±10% of the output magnitude, this is most likely a function of the low numbers involved, i.e. accelerations <20 m*g* and ≈4% of time spent at accelerations >120 m*g*. For example, when comparing accelerations during seated activities, the distribution of time across 40 m*g* increments in acceleration and time spent in MVPA between accelerometer brands, despite sometimes not reaching equivalence, fairly high ICC's, low mean bias and relatively narrow limits of agreement were evident.

It is possible that the closer agreement between the GENEActiv and Axivity than for either accelerometer with the ActiGraph is due to the taping together of the GENEActiv and Axivity accelerometers while the ActiGraph was worn proximal to the GENEActiv and Axivity. While this may have contributed, the 11% higher acceleration from the GENEActiv relative to the ActiGraph (across the total period) in the current study is not dissimilar to the 13-16% higher acceleration observed in our earlier studies where the GENEActiv and ActiGraph GT3X+ were taped together at the wrist (28) and hip (12). These consistent differences we, and others, have observed may relate to technical differences between the brands (12, 20, 36); specifically it appears there is some onboard processing of the raw acceleration signal of the ActiGraph device, but details of this are proprietary (36, 37).

The most recent version of the ActiGraph, the GT9X Link, was used in this study. How this compares to the previous version, the GT3X+, which is the accelerometer that has been deployed in large surveys, e.g. NHANES, is important. The accelerometer sensor in the GT9X is the same as the sensor in GT3X+ so good comparability between the two would be anticipated; perhaps more importantly there are differences in the

design of the devices that may affect orientation when worn. However, this should affect individual axis output rather than resultant metrics, such as the ENMO average acceleration metric. Recently, Montoye and colleagues (38) presented a comparison of the GT9X and GT3X+ ActiGraphs. They reported that raw acceleration data were highly correlated between models. Further, our results comparing the GT9X to the GENEActiv are very similar to those from our previous comparison of the GT3X+ to the GENEActiv (20).

Strengths of this study include the large range of sedentary and light activities incorporated in a simulated free-living protocol designed to encourage a range of natural self-paced behaviours common in normal daily life. This elicited an average acceleration across the total period similar to the daily average acceleration observed in free-living individuals (6), strengthening the ecological validity of the results. The observation facilitated the evaluation of specific activities and the accuracy and comparability of sedentary/light thresholds. The evaluation of the agreement of key physical activity outcomes between accelerometer brands will facilitate selection of the most appropriate outcomes to use when comparing studies which have used different accelerometer brands. Critically, the GGIR accelerometer processing package used in this study (17, 18, 19) can be easily applied to large datasets and is available open-source, as is the function we created to convert Axivity raw files to a format that can be analysed in GGIR. Limitations of the study include: the small sample size, although the study was powered appropriately; the self-selected homogenous young fit cohort; and the relatively small amount of time spent in MVPA, given the focus of our protocol design on sedentary and light activities. Further, the short duration of the study precluded the examination of sleep and the range of behaviours across a 24 h day.

In conclusion, this study suggests that key physical activity outcomes from GGIR (acceleration and time spent in intensity cut-points) can be considered equivalent between the Axivity and GENEActiv accelerometers. Comparability with the ActiGraph is reduced, with acceleration and MVPA approximately 9-11% lower; however time spent sedentary and time in light intensity activity can be considered equivalent. It should be noted that these results are generalisable only to studies using the same wear location (non-dominant wrist) and processing the data in GGIR. To ensure the comparability of the accelerometer brands was tested, and not

confounded by differences in wear location, all accelerometer brands were worn on the non-dominant wrist. This is common practise in most studies (e.g. 6, 9, 21, 22, 23, 24), but a notable exception is UK Biobank where the Axivity was worn on the dominant wrist (4). The next step is to examine the agreement between accelerometer brands as they are commonly used, i.e. in a true free-living environment with the GENEActiv and ActiGraph again on the non-dominant wrist, but the Axivity on the dominant wrist, as in UK Biobank.

Acknowledgements and funding

The authors would like to thank Hannah Goodes, Carly Kingdon and Megan Waters for their assistance with data collection and the participants for their involvement with this study. The research was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at University Hospitals of Leicester and Loughborough University, the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care – East Midlands (NIHR CLAHRC – EM) and the Leicester Clinical Trials Unit. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. No external sources of funding were accessed.

Conflict of Interest

The results of the present study do not constitute endorsement by ACSM. We declare that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

References

1. Chen Y, Copeland WK, Vedanthan R et al. Association between body mass index and cardiovascular disease mortality in East Asians and South Asians: pooled analysis of prospective data from the Asia Cohort Consortium. BMJ 2013; 347:f5446.

2. Mons U, Müezzinler A, Gellert C et al. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. BMJ 2015; 350:h1551. doi: https://doi.org/10.1136/bmj.h1551

3. Lee I-M, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major noncommunicable diseases worldwide: An analysis of burden of disease and life expectancy. Lancet 2012; 380: 218–229.

4. Doherty A, Jackson D, Hammerla N et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PLoS Med 2017. doi.org/10.1371/journal.pone.0169649.

5. Swerdlow AJ, Jones ME, Schoemaker MJ et al. The Breakthrough Generations Study: design of a long-term UK cohort study to investigate breast cancer aetiology. Br J Cancer 2011; 105: 911-917.

6. da Silva ICM, van Hees VT, Ramires VV et al. Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. Int J Epidemiol. 2014; 43:1959-1968.

7. Wake M, Clifford S, York E et al. Introducing Growing Up in Australia's Child Health CheckPointA physical and biomarkers module for the Longitudinal Study of Australian Children. Family Matters 2014; 94: 15-23.

8. Li X, Kearney PM, Keane E et al. Levels and sociodemographic correlates of accelerometer-based physical activity in Irish children: a cross-sectional study. J Epidemiol Community Health 2017, doi:10.1136/jech-2016-207691

9. Menai M, van Hees VT, Elbaz A, Kivimaki M, Singh-Manoux A, Sabia S. Accelerometer assessed moderate-to-vigorous physical activity and successful ageing: results from the Whitehall II study. Sci Rep 2017; 8:45772. doi:10.1038/srep45772.

10. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Brit J Sports Med 2014; 48:1019-1023.

11. Wijndaele K, Westgate K, Stephens SK et al. Utilization and harmonization of adult accelerometry data: Review and expert consensus. Med Sci Sports Exerc. 2015; 47: 2129–39.

12. Rowlands AV, Fraysse F, Catt M et al. Comparison of measured acceleration output from accelerometry-based activity monitors. Med Sci Sports Exerc 2015; 47:201-210.

13. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. Age-Group Comparability of Raw Accelerometer Output from Wrist- and Hip-Worn Monitors. Med Sci Sport Exerc. 2014; 46: 1816-1824.

14. Hildebrand M, Hansen BH, van Hees VT, Ekelund U. Evaluation of raw acceleration sedentary thresholds in children and adults. Scand J Med Sci Sports 2016; doi: 10.1111/sms.12795.

15. Bakrania K, Yates T, Rowlands AV et al. Developing and validating intensity-based thresholds on raw accelerometer data for discriminating between sedentary behaviours and light-intensity physical activities: a MAD approach. PLoS One. 2016; 11(10): e0164045. doi:10.1371/journal.pone.0164045.

16. van Hees VT, Thaler-Kall K, Wolf KH et al. A. Challenges and Opportunities for Harmonizing Research Methodology: Raw Accelerometry. Methods Inf Med. 2016; 7;55:525-532.

17. van Hees VT, Renstrőm F, Wright A et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. PLoS ONE. 2011; 6(7): e22922. doi: 10.1371/journal.pone.0022922.

18. van Hees VT, Gorzelniak L, Dean León 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. doi: 10.1371/journal.pone.0061691.

19. van Hees VT, Fang Z, Langford J et al. Auto-calibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. J Appl Physiol. 2014; 117(7): 738-744.

20. 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; 48:1935-1941.

21. Ramires VV, Wehrmeister F C, Böhm AW et al. Physical activity levels objectively measured among older adults: a population-based study in a Southern city of Brazil. Int J Behav Nutr Phys Act 2017; 14:13. doi.org/10.1186/s12966-017-0465-3

22. Bell JA, Hamer M, van Hees V et al. Healthy obesity and objective physical activity. Am J Clin Nutr 2015; 102:268-275.

23. Horta BL, Schaan BD, Bielemann RM et al. Objectively measured physical activity and sedentary-time are associated with arterial stiffness in Brazilian young adults. Atherosclerosis 2015; 243:148–54.

24. Fairclough SJ, Noonan R, Rowlands AV, van Hees V, Knowles Z, Boddy LM. Wear compliance and activity in children wearing wrist and hip mounted accelerometers. Med Sci Sports Exerc 2016; 48, 243-253.

25. Rowlands AV, Cliff DP, Fairclough SJ et al. Moving forward with backwards compatibility: Translating wrist accelerometer data. Med Sci Sports Exerc 2016; 48: 2142-2149.

26. Scott JJ, Rowlands AV, Cliff D, Morgan PJ, Plotnikoff RC, Lubans DR. Testing the validity and feasibility of the GENEActiv accelerometer in free-living adolescents. J Sci Med Sport. 2017; doi.org/10.1016/j.jsams.2017.04.017

27. Ladha C, Ladha K, Jackson D, Olivier P. Table validation of openmovement Ax3 accelerometer. In: Proceedings of the 3rd International Conference of Ambulatory Monitoring Physical Activity and Movement; 2013 June 17-19: Amherst, MA (USA). University of Amherst; 2013: p.69–70.

28. Stiles VH, Griew PJ, Rowlands AV. Use of accelerometry to classify activity beneficial to bone in premenopausal women. Med Sci Sports Exerc 2013; 45:2353-2361.

29. Bland JM, Altman GA. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1(8476):307-310.

30. Wellek S. Testing Statistical Hypotheses of Equivalence. Boca Raton (FL): Chapman & Hall/CRC; 2003, xv, 284 p.

31. Lee J-M, Kim Y, Welk GJ. Validity of consumer-based physical activity monitors. Med Sci Sports Exerc 2014; 46: 1840-1848.

32. Kim Y, Welk GJ. Criterion validity of competing accelerometry-based activity monitoring devices Med Sci Sports Exerc 2015; 47: 2456-2463.

33. Rowlands AV, Olds TS, Hillsdon M et al. Assessing sedentary behaviour with the GENEActiv: Introducing the Sedentary Sphere. Med Sci Sports Exerc 2014; 46:1235-1247.

34. Sasaki JE, Hickey AM, Staudenmeyer JW, Kent JA, Freedson PS. Performance of activity classification algorithms in free-living older adults. Med Sci Sports Exerc 2016; 48:941-950.

35. Sedentary Behaviour Research Network. 2012. Standardized use of the terms “sedentary” and “sedentary behaviours”. Appl Physiol Nutr Metab 2012; 37:540–542.

36. John D, Sasaki J, Staudenmayer J, Mavilia M, Freedson PS. Comparison of raw acceleration from the GENEA and ActiGraphTM GT3X+ activity monitors. Sensors. 2013; 13: 14754-63.

37. Brønd JC, Arvidsson D. Sampling frequency affects the processing of ActiGraph raw acceleration data to activity counts. J Appl Physiol 2015; doi: 10.1152/japplphysiol.00628.2015.

38. Montoye AM, Nelson M, Bock J et al. Comparability of raw and count-based data from the ActiGraph GT9X Link and GT3X+ accelerometers. In: Proceedings of the 5th International Conference of Ambulatory Monitoring Physical Activity and Movement; 2017 June 21-23: Bethesda, MD (USA). National Institutes of Health, Natcher Conference Center; 2017: Abstract: 3-61.

Table 1. Physical activity outcomes (mean, SD) by monitor (GENEActiv, Axivity, ActiGraph)

Note: all activities approx 2 h, seated only 1 h 20, upright only 24 min. Note total time includes all observed activities and transitions
All and upright activities: GENEActiv N = 19, Axivity N = 17, ActiGraph N = 18
Seated activities: GENEActiv N = 18, Axivity N = 16, ActiGraph N = 18

		Mean			Standard deviation		
		GENEActiv	Axivity	ActiGraph	GENEActiv	Axivity	ActiGraph
		Acceleration (m*g*)					
Total period		42.1	40.9	37.8	10.3	11.5	11.2
Seated only		13.1	13.8	12.7	4.3	6.5	5.1
Upright only		87.3	83.8	78.9	25.5	24.7	24
Individual activities							
Seated	Reading	10.6	11.5	9.8	4.7	7.4	5.6
	Eating	20.2	21.6	20.1	7.5	8.7	9.8
	Computer	11.4	12.1	12.3	4.6	7.1	7.4
	TV	11.1	10.7	9.5	6.5	8.6	6.5
Upright	Standing still	16.6	13.2	18.7	20.6	10.6	21.4
	Walking	160	151.4	139.8	52.7	61.7	44.3
	Sweeping	187.7	176.7	175.5	92.7	84.3	81
	Washing pots	60.2	61.4	53.9	18.6	20.4	19.5
	Dusting	78.6	75	70.9	40.6	37.8	37.4
	Standing computer	20.7	21.6	14.4	8.5	7.8	8.4
		% Time in cut-point categories					
Total time	Cut-point (mg)						
Sedentary	<30	67.6	67.5	69.7	7.4	7.7	6.4
	<40	72.2	72.4	74.5	6.6	6.4	5.7
	<50	75.7	76.2	77.8	5.9	5.7	5.3
Light	30-100	18.8	19.5	18.4	5.3	5.8	2.9
	40-100	14.2	14.6	13.7	4.4	4.3	2.1
	50-100	10.7	10.8	10.3	3.6	3.4	1.6
MVPA	>100	13.6	13.0	11.9	3.5	4.0	4.4
		% Accuracy of sedentary cut-points for % time seated only and upright					
	Cut-point (mg)						
Seated only	<30	89.2	87.6	89.2	6.2	7.2	7.0
	<40	91.8	91.0	92.3	5.6	6.0	6.1
	<50	93.6	93.1	94.1	5.1	5.5	5.4
Upright	>30	37.8	39.1	40.3	4.9	5.3	6.2
	>40	44.6	45.7	46.7	5.2	5.8	6.6
	>50	49.8	51.0	51.7	5.6	6.8	7.2
		% time in 40 mg incremental ranges					
Total time	Range (mg)						
	0-40	72.2	72.4	74.5	6.6	6.4	5.7
	40-80	10.8	11.2	10.3	3.3	3.3	1.8
	80-120	6.4	6.2	6.0	2.3	1.9	1.0
	120-160	4.2	4.2	3.8	1.2	1.2	1.2
	160-200	2.5	2.5	2.3	0.9	1.1	1.2
	>200	4.0	3.5	3.1	2.7	2.4	2.5

Table 2. Agreement between pairs of monitors (GENEActiv/ActiGraph, GENEActiv/Axivity, Axivity/ActiGraph) for average dynamic acceleration.

		Intra-class correlation (ICC)[a]			Agreement (Bland and Altman)[b]						Equivalency[c]		
		ICC (95% CI)			mean bias			95% limits of agreement			Can be considered equivalent (see Figure 1)		
	Monitor 1	GEN	GEN	Ax	GEN	GEN	Ax	GEN	GEN	Ax	GEN	GEN	Ax
	Monitor 2	ActiG	Ax	ActiG	ActiG	Ax	ActiG	ActiG	Ax	ActiG	ActiG	Ax	ActiG
Total period		0.86 (0.56, 0.95)	0.95 (0.87, 0.98)	0.82 (0.50, 0.94)	3.3	-0.1	3.6	9.6	6.8	12.0	×	✓	×
Seated only		0.73 (0.40, 0.89)	0.85 (0.62, 0.95)	0.59 (0.15, 0.84)	0.6	-0.6	1.3	7.0	6.1	10.2	×	×	×
Upright		0.87 (0.40, 0.96)	0.97 (0.91, 0.99)	0.75 (0.28, 0.92)	9.0	-0.9	10.2	19.6	11.6	26.6	×	✓	×
Seated activities	Reading	0.67 (0.31, 0.87)	0.79 (0.48, 0.92)	0.37 (-0.15, 0.73)	1.1	-0.8	1.7	4.3	4.1	7.4	×	×	×
	Eating	0.82 (0.56, 0.93)	0.89 (0.70, 0.96)	0.64 (0.22, 0.86)	5.3	3.8	8.1	5.3	3.8	8.1	×	×	×
	Computer	0.60 (0.17, 0.84)	0.83 (0.58, 0.94)	0.49 (-0.04, 0.80)	-0.1	-0.4	0.5	5.6	3.7	7.6	×	×	×
	TV	0.79 (0.51, 0.92)	0.90 (0.73, 0.97)	0.79 (0.50, 0.92)	1.4	0.2	1.6	4.1	3.7	4.9	×	×	×
Upright activities	Standing still	0.95 (0.88, 0.98)	0.90 (0.75, 0.96)	0.68 (0.27, 0.88)	-0.9	-1.0	-1.1	13.2	9.1	19.1	×	×	×
	Walking	0.87 (0.56, 0.96)	0.98 (0.95, 0.99)	0.83 (0.52, 0.94)	13.9	-3.6	16.0	39.5	19.3	54.8	×	✓	×
	Sweeping	0.95 (0.58, 0.99)	0.99 (0.97, 0.99)	0.94 (0.74, 0.98)	20.1	4.6	16.9	36.2	24.4	46.2	×	✓	×
	Washing pots	0.85 (0.52, 0.94)	0.92 (0.79, 0.97)	0.80 (0.45, 0.93)	6.2	-0.5	6.5	20.5	14.7	21.9	×	✓	×

	Intra-class correlation (ICC)[a]			Agreement (Bland and Altman)[b]						Equivalency[c]		
	ICC (95% CI)			mean bias			95% limits of agreement			Can be considered equivalent (see Figure 1)		
Monitor 1	GEN	GEN	Ax	GEN	GEN	Ax	GEN	GEN	Ax	GEN	GEN	Ax
Monitor 2	ActiG	Ax	ActiG	ActiG	Ax	ActiG	ActiG	Ax	ActiG	ActiG	Ax	ActiG
Dusting	0.94 (0.68, 0.98)	0.99 (0.98, 0.99)	0.91 (0.57, 0.98)	9.4	0.2	10.1	22.0	9.4	23.3	×	✓	×
Standing computer	0.70 (0.02, 0.91)	0.60 (0.18, 0.84)	0.37 (-0.11,0.74)	5.4	-2.1	8.0	9.4	13.4	14.0	×	×	×

ActiG = ActiGraph GT9X; Ax = Axivity; GEN = GENEActiv,

[a]Single measure, absolute agreement. All $p < 0.05$ apart from Axivity/ActiGraph ENMO for the specific activities reading (seated), computer (seated) and standing computer (upright)

[b]Bland and Altman: Bias calculated as Monitor 1 - Monitor 2

[c]95% confidence interval (CI) for mean of second monitor falls into the proposed equivalence zone (i.e. +/-10% of the mean) of the mean of the first monitor.

Table 3. Agreement between pairs of monitors (GENEActiv/ActiGraph, GENEActiv/Axivity, Axivity/ActiGraph) for % time in cut-point categories.

		Intra-class correlation (ICC)[a] ICC (95% CI)			Agreement (Bland and Altman)[b] mean bias			95% limits of agreement			Equivalency[c] Can be considered equivalent (see Figure 1)		
	Monitor 1 Monitor 2	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG
Sedentary	<30	0.84 (0.63, 0.94)	0.95 (0.87, 0.98)	0.87 (0.64, 0.96)	-0.9	0.3	-1.5	6.3	4.9	6.1	✓	✓	✓
	<40	0.89 (0.69, 0.96)	0.96 (0.88, 0.99)	0.90 (0.61, 0.97)	-1.2	0.0	-1.5	4.3	4.0	4.0	✓	✓	✓
	<50	0.90 (0.70, 0.97)	0.96 (0.89, 0.99)	0.92 (0.68, 0.98)	-1.1	-0.2	-1.3	3.6	3.4	3.4	✓	✓	✓
Light	30-100	0.70 (0.36, 0.88)	0.95 (0.86, 0.98)	0.69 (0.29, 0.88)	-0.5	-0.4	0.0	4.2	3.8	5.5	✓	✓	✓
	40-100	0.86 (0.66, 0.95)	0.96 (0.90, 0.99)	0.79 (0.47, 0.92)	-0.2	-0.2	0.1	2.1	2.5	3.1	✓	✓	✓
	50-100	0.89 (0.73, 0.96)	0.97 (0.92, 0.99)	0.83 (0.56, 0.94)	-0.3	0.0	-0.2	1.4	1.8	2.1	✓	✓	✓
MVPA	>100	0.85 (0.47, 0.95)	0.96 (0.88, 0.98)	0.84 (0.49, 0.95)	1.4	0.1	1.5	3.4	2.3	4.1	×	✓	×

ActiG = ActiGraph GT9X; Ax = Axivity; GEN = GENEActiv,

[a]Single measure, absolute agreement. All $p < 0.05$ apart from Axivity/ActiGraph ENMO for the specific activities reading (seated), computer (seated) and standing computer (upright)

[b]Bland and Altman: Bias calculated as Monitor 1 - Monitor 2

[c]95% confidence interval (CI) for mean of second monitor falls into the proposed equivalence zone (i.e. +/-10% of the mean) of the mean of the first monitor.

Table 4. Agreement between pairs of monitors (GENEActiv/ActiGraph, GENEActiv/Axivity, Axivity/ActiGraph) for % accuracy of sedentary cut-points.

		Intra-class correlation (ICC)[a] ICC (95% CI)			Agreement (Bland and Altman)[b] mean bias			95% limits of agreement			Equivalency[c] Can be considered equivalent (see Figure 1)		
	Monitor 1 Monitor 2	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG
Seated activities only	<30	0.88 (0.70, 0.96)	0.92 (0.77, 0.97)	0.80 (0.50, 0.93)	-0.4	1.3	-1.6	6.6	4.9	8.9	✓	✓	✓
	<40	0.95 (0.86, 0.98)	0.96 (0.89, 0.99)	0.90 (0.70, 0.97)	-0.8	0.7	-1.4	3.6	3.0	4.9	✓	✓	✓
	<50	0.97 (0.89, 0.99)	0.99 (0.96, 0.99)	0.95 (0.78, 0.98)	-0.7	0.4	-1.1	2.3	1.7	3.0	✓	✓	✓
Upright activities only	>30	0.65 (0.25, 0.86)	0.83 (0.58, 0.94)	0.76 (0.19, 0.93)	-2.4	-0.3	-2.7	8.3	5.5	6.0	×	✓	⊗
	>40	0.78 (0.45, 0.92)	0.90 (0.73, 0.96)	0.79 (0.29, 0.94)	-1.9	0.0	-2.6	6.6	4.7	6.1	✓	✓	✓
	>50	0.82 (0.55, 0.93)	0.92 (0.78, 0.97)	0.83 (0.49, 0.94)	-1.7	0.0	-2.2	6.6	4.8	6.8	✓	✓	✓

ActiG = ActiGraph GT9X; Ax = Axivity; GEN = GENEActiv,

[a]Single measure, absolute agreement. All $p < 0.05$ apart from Axivity/ActiGraph ENMO for the specific activities reading (seated), computer (seated) and standing computer (upright)

[b]Bland and Altman: Bias calculated as Monitor 1 - Monitor 2

[c]95% confidence interval (CI) for mean of second monitor falls into the proposed equivalence zone (i.e. +/-10% of the mean) of the mean of the first monitor.

⊗ Could be considered equivalent when the alternate accelerometer (monitor 2) was selected as the reference.

Note: All and upright activities: GENEActiv N = 19, Axivity N = 17, ActiGraph N = 18; Seated activities: GENEActiv N = 18, Axivity N = 16, ActiGraph N = 18

Table 5. Agreement between pairs of monitors (GENEActiv/ActiGraph, GENEActiv/ActiGraph, Axivity/ActiGraph) for distribution of time across 40 m*g* increments of acceleration

	Intra-class correlation (ICC)[a] ICC (95% CI)			Agreement (Bland and Altman)[b] mean bias			95% limits of agreement			Equivalency[c] Can be considered equivalent (see Figure 1)		
Monitor 1 Monitor 2	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG	GEN ActiG	GEN Ax	Ax ActiG
0-40	0.89 (0.69, 0.96)	0.96 (0.88, 0.99)	0.90 (0.61, 0.97)	-1.17	0.03	-1.54	4.3	4.0	4.0	✓	✓	✓
40-80	0.82 (0.57, 0.93)	0.95 (0.86, 0.98)	0.78 (0.47, 0.92)	-0.10	-0.25	0.33	2.1	2.2	2.7	✓	✓	✓
80-120	0.90 (0.74, 0.96)	0.94 (0.83, 0.98)	0.78 (0.47, 0.92)	-0.02	0.23	-0.19	0.8	1.5	1.3	✓	✓	✓
120-160	0.77 (0.48, 0.91)	0.94 (0.83, 0.98)	0.80 (0.50, 0.93)	0.29	0.06	0.29	1.4	0.8	1.4	×	✓	×
160-200	0.79 (0.52, 0.92)	0.93 (0.82, 0.98)	0.71 (0.35, 0.89)	0.17	-0.10	0.33	1.4	0.8	1.8	×	ⓧ	×
>200	0.89 (0.51, 0.97)	0.97 (0.93, 0.99)	0.81 (0.46, 0.93)	0.84	0.02	0.78	1.9	1.1	2.6	×	✓	×

ActiG = ActiGraph GT9X; Ax = Axivity; GEN = GENEActiv,

[a]Single measure, absolute agreement. All $p < 0.05$ apart from Axivity/ActiGraph ENMO for the specific activities reading (seated), computer (seated) and standing computer (upright)

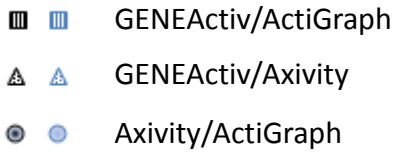
[b]Bland and Altman: Bias calculated as Monitor 1 - Monitor 2

[c]95% confidence interval (CI) for mean of second monitor falls into the proposed equivalence zone (i.e. +/-10% of the mean) of the mean of the first monitor.

ⓧ Could be considered equivalent when the alternate accelerometer (monitor 2) was selected as the reference.

Note: All and upright activities: GENEActiv N = 19, Axivity N = 17, ActiGraph N = 18; Seated activities: GENEActiv N = 18, Axivity N = 16, ActiGraph N = 1

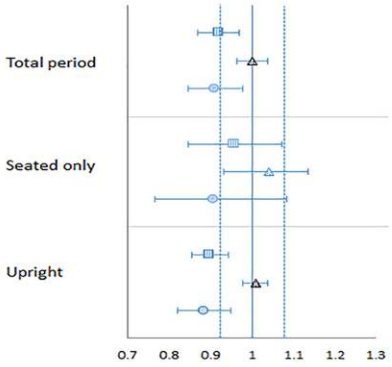
Figure 1 Equivalence between pairs of accelerometer brands for a) acceleration, b) acceleration during individual seated activities, c) acceleration during individual upright activities, d) % time spent within cut-point categories, e) sedentary cut-point accuracy, and f) % time spent in incremental 40 m*g* ranges.



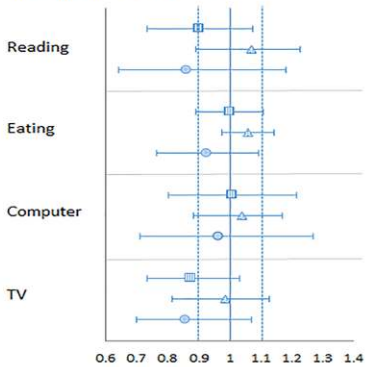
Error bars represent the 95% confidence interval of the ratio. Equivalence = 1 (solid line), proposed equivalence zone (i.e. +/-10% of the mean) represented by dashed vertical lines.

Black marker denotes the 95% confidence interval (CI) for the mean of the second monitor falls into the proposed equivalence zone of the mean of the first monitor, blue marker denotes it falls outside the proposed equivalence zone.
*Could be considered equivalent when the alternate accelerometer was selected as the reference (see text for details)

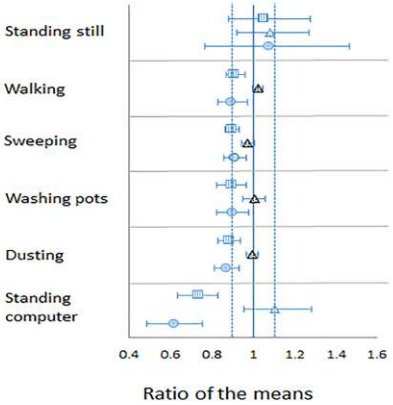
a) Equivalence of acceleration (mg)



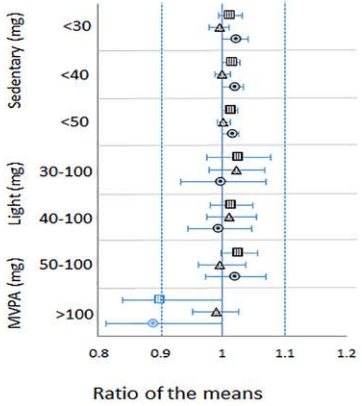
b) Equivalence of acceleration (individual seated activities)



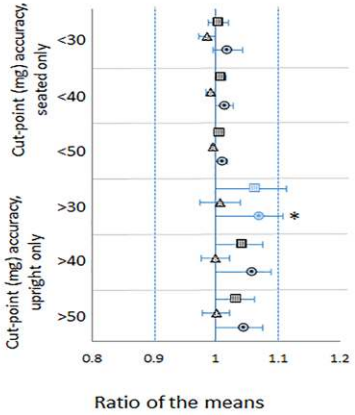
c) Equivalence of acceleration (individual upright activities)



d) Equivalence of % time spent within cut-point categories (total period)



e) Equivalence of sedentary cut-point accuracy



f) Equivalence of % time spent in incremental 40 mg ranges (total period)