Sanders, GJ, Boddy, LM, Sparks, AS, Curry, WB, Roe, B, Kaehne, A and Fairclough, SJ

Evaluation of wrist and hip sedentary behaviour and moderate-to-vigorous physical activity raw acceleration cutpoints in older adults

Evaluation of wrist and hip sedentary behaviour and moderate-to-vigorous physical activity raw acceleration cutpoints in older adults

George J. Sanders[1], Lynne M. Boddy[2], S. Andy Sparks[3], Whitney B. Curry[4], Brenda Roe[5,6], Axel Kaehne[5], Stuart J. Fairclough[1,7].

Edge Hill University.

Author Note

[1] Physical Activity and Health Research Group, Department of Sport and Physical Activity, Edge Hill University, Ormskirk, UK

[2] Physical Activity Exchange, Research Institute for Sports and Exercise Sciences, Liverpool John Moores University, Liverpool, UK

[3] Sport Nutrition and Performance Research Group, Department of Sport and Physical Activity, Edge Hill University, Ormskirk, UK

[4]Wellbeing and Public Health, Cornwall Council, Truro, UK

[5] Faculty of Health and Social Care, Edge Hill University, UK

[6]Personal Social Services Research Unit, University of Manchester, UK

[7]Department of Physical Education and Sports Sciences, University of Limerick, Ireland

Corresponding author: George J. Sanders, Department of Sport and Physical Activity, Edge Hill University, St Helens Road, Ormskirk, L39 4QP, United Kingdom. Telephone: 01695 657 344 Email: george.sanders2@go.edgehill.ac.uk

Edge Hill University institutional ethical approval number: SPA-REC-2016-343

Word count: 4963 (excluding title page, abstract, references, tables and figures)

Evaluation of wrist and hip sedentary behaviour and moderate-to-vigorous physical activity raw acceleration cutpoints in older adults

Abstract

Wrist-based accelerometers are now common in assessing physical activity (PA) and sedentary behaviour (SB) in population-based studies, but there is a scarcity of raw acceleration cutpoints in older adults. The study aimed to determine and evaluate wrist-based GENEActiv (GA) and hip-based ActiGraph GT3X+ (AG) raw acceleration cutpoints for SB and moderate-to-vigorous PA (MVPA) in older adults ≥60 years of age. A laboratory-based calibration analyses of 34 healthy older adults involved receiver operator characteristic (ROC) curves to determine raw acceleration cutpoints for SB and MVPA. ROC analysis revealed an area under the curve (AUC) of 0.88 for GA SB and MVPA, and 0.90 for AG SB and 0.94 for AG MVPA. Sensitivity optimised SB and specificity optimised MVPA GA cutpoints of 57 mg and 104 mg, and AG cutpoints of 15 mg and 69 mg were also generated, respectively. Cross-validation analysis revealed moderate agreement for GA and AG SB cutpoints, and fair to substantial agreement for GA and AG MVPA cutpoints, respectively. The resultant cutpoints can classify older adults as engaging in SB or not engaging in MVPA but the sensitivity optimised SB cutpoints should be interpreted with a degree of caution due to their modest cross-validation results.

Key Words: *Accelerometry; Activity Monitors; Calibration; Cross-validation; Ageing*

Introduction

There is strong evidence supporting the benefits of reducing sedentary behaviour (SB) and increasing physical activity (PA) in older adults (Devereux-Fitzgerald, Powell, Dewhurst, & French, 2016; Greaney, Lees, Blissmer, Riebe, & Clark, 2016; Kim, Im & Choi, 2016; Lewis, Napolitano, Buman, Williams, & Nigg, 2017). Despite this, older adults (e.g., aged 60+ years; Heo et al., 2017) remain the most sedentary and physically inactive segment of society (Chastin et al., 2017). Older adults spend approximately 80% of their awake time engaged in SB which represents eight to 12 hours/day (Chastin et al., 2017). Furthermore, compliance with recommended PA guidelines is poor in the general population and even worse in older adults (Troiano et al., 2008). Health Survey for England data indicate that only 20% of men and 17% of women aged 65 to 74 years meet recommended PA guidelines (NHS, 2015). This contrasts with 49 % of men and 35 % of women aged 25 to 34 years (NHS, 2015). It is widely acknowledged that increasing age is associated with decreased time spent in PA (Berkemeyer et al., 2016; Harvey, Chastin & Skelton, 2014; Martinez-Gomez et al., 2017; Wullems, Verschueren, Degens, Morse, & Onambélé, 2016) and this trend is apparent in older adults, particularly after retirement age (Martinez-Gomez et al., 2017; Strain et al., 2016). To exemplify this, Buman et al. (2010) reported that when comparing those between 66-69 years old with those aged 80 years and older, engagement in MVPA decreased from 16.2 $min \cdot day^{-1}$ to 10.7 $min \cdot day^{-1}$.

As described in the behavioural epidemiological framework (Sallis, Owen, & Fotheringham, 2000), accurate measurements of SB and PA are needed to detect

potential correlates, identify relationships between such behaviours and associated health outcomes, and evaluate the efficacy of intervention strategies (Lewis et al., 2017). Accelerometry is now commonly adopted for monitoring older adults' SB and PA levels (Mañas et al., 2017; Oguma et al., 2017; Wullems, Verschueren, Degens, Morse, & Onambélé, 2017; Zhu et al., 2017), as it allows for valid, reliable, and accurate assessments of activity intensity, frequency, and duration (Mathie, Coster, Lovell, & Celler, 2004; Prince et al., 2008). Moreover, accelerometers are particularly appropriate for assessing PA in older adults as these devices require no input from the participant over the data collection period, and superior wearer compliance has been demonstrated when compared to younger age groups. For example, in the UK Biobank study of 103,578 adults and older adults, average wear time compliance of a wrist-worn triaxial accelerometers increased by an average of two hours 18 $min \cdot day^{-1}$ (1.6%) for each decade from 40 to 79 years old (Doherty et al., 2017). Accelerometry also eliminates self-report questionnaire bias related to subjective recall of past events which is an ability that can decline with ageing (Barnett, van den Hoek, Barnett, & Cerin, 2016).

The hip has been the conventional attachment site for accelerometers because of its proximity to the centre of mass (Troiano, McClain, Brychta, & Chen, 2014; Van Hees et al., 2011). A limitation of hip-worn accelerometers in studies involving older adults (Copeland & Esliger, 2009; Taylor et al., 2014) is that most commonly used cutpoints applied to data to classify PA intensity have been calibrated for younger adults (Falck, Davis, & Liu-Ambrose, 2016). As the energy expenditure associated with a given metabolic equivalent (MET) activity intensity threshold is lower in older adults compared to younger adults (Hall, Howe, Rana, Martin, & Morey,

2013), using accelerometer cutpoints developed in young adults would likely result in an underestimation of time spent in moderate-to-vigorous intensity PA (MVPA) (Barnett et al., 2016).

Recent accelerometer studies have suggested that the wrist may be a preferable attachment site as it can more accurately capture the arm motions of non-ambulatory based activities such as household chores (Evenson et al., 2015; Landry, Falck, Beets, & Liu-Ambrose, 2015), and is less influenced by atypical gait patterns and walking speed variability, which are both commonly observed in older adults (Ko, Jerome, Simonsick, Studenski, & Ferrucci, 2018). Wrist-worn accelerometers have demonstrated excellent validity against energy expenditure as the criterion measure, and in comparison to hip-worn monitors (Esliger et al., 2011). Furthermore, wrist- and hip-worn accelerometers have demonstrated comparable free-living MVPA classification accuracy (Hargens, et al., 2017). Superior wearer compliance has also been reported for accelerometers worn on the wrist versus the hip in large population-based studies including the National Health and Nutrition Examination Survey (NHANES), Dallas Heart Study, and the UK Biobank study adopting wrist-worn accelerometer protocols (Doherty et al., 2017; Lakoski & Kozlitina, 2014; Troiano et al., 2014). Specifically, wrist-worn data from the 2011 to 2012 cycle of NHANES showed that 70–80% of participants provided at least six days of data with at least 18 hours of wear. This contrasts with 40–70% of participants who provided at least six days of hip-worn accelerometer data with at least 10 hours of wear in the 2003 to 2004 cycle of NHANES (Troiano et al., 2014).

A further development in accelerometer-based SB and PA research is the move toward raw acceleration signal processing. This advance in accelerometer-based PA monitoring, which has traditionally used accelerometer output reduced to dimensionless activity "counts" per user-specified period of time or epoch (Fairclough et al., 2016) is likely to provide greater methodological transparency in post-data collection analytical processes and improve comparability of data between different accelerometer models (Hildebrand, Van Hees, Hansen, & Ekelund, 2014). Devices such as the GENEActiv (Activinsights, Cambs, United Kingdom) and ActiGraph GT3X+ and GT9X (ActiGraph, Pensacola, FL) are capable of collecting and recording raw unfiltered accelerations, which can then be subject to researcher-driven data processing procedures (Welk, McClain, & Ainsworth, 2012). Interunit reliability is acceptable for both brands (Esliger et al., 2011; Santos-Lozano et al., 2012). Although time spent in SB and light, moderate and/or vigorous intensity PA can be quantified from raw acceleration data (Matthew, 2005), currently, no raw acceleration cutpoints for SB and MVPA exist for older adults. Consequently, the aim of this study was to determine laboratory-based wrist-worn GENEActiv and hip-worn ActiGraph GT3X+ raw acceleration cutpoints for SB and MVPA in older adults.

Methods

Study population

A homogenous purposive sample of 34 community dwelling white, British older adults (10 male), aged 60-86 years (mean age = 69.6, SD = 8.0 years) were recruited through leaflets distributed at local fitness centres/gyms and community centres, and

through word of mouth referrals. A sample size of 30 participants was targeted so as to be comparable with recent calibration studies in older adults (Landry et al., 2015; Wullems et al., 2017). Individuals interested in participating in the study were pre-screened for inclusion criteria which set out that participants must be (1) ≥60 years of age, (2) be physically cleared for exercise using the modified Physical Activity Readiness Questionnaire (Modified PAR-Q; Cardinal, Esters, & Cardinal, 1996; Cardinal & Cardinal, 2000), (3) have the ability to walk briskly on a treadmill without assistance, and (4) not be taking any medications that would influence energy expenditure or their ability to perform ambulatory activity. Participants were excluded if they had a medical condition precluding them from exercise, were unable to wear a portable indirect calorimeter during testing, or had limited mobility such that they could not walk on a treadmill independently. Prior to the commencement of the study, institutional ethical approval was received (SPA-REC-2016-343) and all participants provided written informed consent prior to their inclusion.

Anthropometrics

Participants' body mass (Seca mechanical scales, Birmingham, UK) and standing height (Holtain Limited stadiometer, Crymych, UK) were measured in light clothing without shoes. Resting blood pressure was measured upon arrival and immediately prior to commencement of the accelerometer calibration protocol using a Boso Medicus Prestige blood pressure monitor (Boso Bosch + Sohn, Germany).

Study protocol

Participants completed two separate laboratory data collection visits (separated by one week) at the university site. Visit 1 was an initial familiarisation of the protocol structure, equipment, and laboratory. Participants also provided written informed consent and completed anthropometric measures, and a six-minute treadmill walk test (6MWT; ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories, 2002) to establish maximal walking speeds and to familiarise participants with treadmill walking at different speeds. Participants completed a laboratory-based protocol consisting of 16 activities (see Table 1 for protocol description). Older adults typically engage in SB and light intensity PA (LPA; Ku, Fox, Liao, Sun, & Chen, 2016) consisting of activities such as carrying light objects, walking slowly, and household chores (Public Health England, 2017). Thus, to represent these typical activity patterns and to make the distinction between sitting and standing, a range of sedentary, stationary, and LPA activities were included in the study protocol. MVPA activities reflected typical lifestyle behaviours of walking and stepping (e.g., stair climbing). All activities were performed in a standardised order as described in Table 1. Participants were provided with standardised instructions from a predetermined script delivered by the same researcher prior to beginning each activity. The activity protocol only was repeated at visit 2. Visit 1 lasted ~80 minutes whilst visit 2 lasted ~60 minutes as participants did not need to repeat the familiarization components and 6MWT.

[INSERT TABLE 1 ABOUT HERE]

Sedentary, stationary and LPA activities were performed for 3-minutes each, whilst the stepping and walking activities were performed for two minutes. Sedentary, stationary, LPA, and stepping activities were separated by a 1-minute transition period, whilst the treadmill walking activities were separated by 5-second transitions. To ensure consistency across measurement sessions, the first author was trained and led on all aspects of the measurement protocol. The start and end of each activity was timed by the first author with a stopwatch (Fastime, Leicestershire, UK). All instruments were synchronised to the same clock to ensure that criterion measurements for a given period of activity were matched with time-stamped accelerometer data for precisely the same duration of the activity. This allowed for appropriate data comparisons to be made across all recording devices.

Accelerometers

Participants wore GENEActiv (GA) (ActivInsights Ltd., Kimbolton, Cambridgeshire, UK) and ActiGraph (AG) GT3X+ (ActiGraph, Pensacola, FL) triaxial accelerometers on the non-dominant wrist and left hip, respectively. Both monitors were set to collect raw triaxial accelerations at 60 Hz. Participants also wore an activPAL (PAL Technologies Ltd., Glasgow, UK) accelerometer on the left anterior thigh as the criterion measure for SB (Chastin, Culhane & Dall, 2014; Varela Mato, Yates, Stensel, Biddle, & Clemes, 2017). The activPAL monitor has been shown to accurately measure postural allocation when worn on the upper thigh (Kozey-Keadle et al., 2011). It is valid for step counts, time spent sitting, lying, standing, and walking, and is currently regarded as a gold standard for the objective measurement of sedentary/sitting time and patterning of SB (Chastin et al., 2018;

Koster et al., 2016; Kozey-Keadle et al., 2011; Sellers, Dall, Grant, & Stansfield, 2016). The activPAL uses proprietary algorithms to classify activity into periods spent sitting, standing, and stepping. activPAL data were collected at a sampling frequency of 20 Hz. Participants wore the same monitors (matched by serial number) for both visits. A total of six different GA and AG, and four different activPAL monitors were used throughout the study. All devices were used and calibrated as per manufacturers' instructions and initialised approximately 10 minutes before the start of each data collection session.

Immediately after testing, the activity monitors were removed and the data downloaded to a single, secured computer. The GA data were downloaded using GENEActiv PC software version 2.9 (Activinsights, Cambs, United Kingdom) and saved in raw format as .bin files, whilst the AG data were downloaded using ActiLife version 3.13.3 (ActiGraph, Pensacola, FL), and saved in raw format as .gt3x files and converted to time-stamped .csv files to facilitate raw data processing. activPAL data were downloaded using activPAL3 software version 7.2.32 (PAL Technologies Ltd., Glasgow, UK), saved as .datx files and converted to .csv "Event" files for processing. Signal processing of raw GA .bin files and raw AG .csv files was completed using R-package GGIR version 1.5 (https://cran.r-project.org/web/packages/GGIR/) (van Hees et al., 2013b). GGIR facilitates data cleaning and the extraction of user-defined acceleration levels, which can then be set to reflect the intensity thresholds as derived in this study (Rowlands et al., 2016). Concurrent with previous studies (Hildebrand et al., 2014; Fairclough et al., 2016; Menai et al., 2017; Rowlands, Yates, Davies, Khunti, & Edwardson, 2016) and to help promote comparability between studies, the Euclidean Norm Minus One

(ENMO) (van Hees et al., 2013b) was adopted to quantify acceleration relative to gravity (1 mg = 0.00981 ms^{-2}), after which negative values were rounded to zero. The ENMO metric has proven competency and increasing use in the field (Vähä-Ypyä et al., 2015). ENMO has been shown to be comparable to other raw acceleration signal metrics such as MAD and HFEN in predicting PA energy expenditure (Bakrania et al., 2016; Van Hees et al., 2013b) and consequently, has been adopted in large scale studies for the analysis of raw acceleration data (Doherty et al., 2017; Menai et al., 2017). Furthermore, although there are differences in the design of the chosen GA and AG devices that may affect individual axis output, these differences affect the resultant ENMO values in a consistent manner with GA output approximately 10% higher than AG output (Rowlands et al., 2018).

Raw data were further reduced by averaging the ENMO values over 1-second epochs. All resulting values are expressed in milli (10^{-3}) gravity-based acceleration units (mg), where 1g = 9.81 m/s^2. Although the ENMO metric can be sensitive to poor calibration (van Hees, et al., 2013b), GGIR autocalibrates the raw triaxial accelerometer signal in order to reduce such calibration error (van Hees et al., 2014). Where insufficient non-movement periods were available for auto-calibration (due to the relatively short duration of the protocol), back-up calibration coefficients derived from older adult males' and females' free-living data collected with the same accelerometer units were used (Hildebrand et al., 2016; Rowlands et al., 2018). Subsequently, no files were removed from the analytical sample as a result of calibration error. The activPAL "Event" files provided the exact time in seconds that posture changes occurred and each second was classified as sedentary, standing, or stepping. These files were then expanded using an Excel formula to obtain second-

by-second data, with each second subsequently classified as either sedentary or not sedentary. These second-by-second activPAL files were synchronised with the 1-second ENMO values from GA and AG. To exclude any transitional movements, the middle 2-minutes of data from each 3-minute activity were extracted and subsequently utilised for analysis. The full 2-minutes for each of the stepping and walking activities were used. This included the 5-second transition period between each walking speed.

Energy Expenditure

As the criterion measure for MVPA, oxygen consumption (VO_2; $ml \cdot kg \cdot min^{-1}$) was measured with a portable indirect calorimetry system (MetaMax 3B-R2, CORTEX Biophysik GmbH, Leipzig, Germany). The Metamax interface and breathing mask (7600 Series V2, Hans Rudolph, Kansas) were set up and fitted as per the manufacturer instructions. VO_2 was measured using breath-by-breath mode and in order to match time periods across devices, data were stored in second-by-second intervals. These measurements were used to determine energy expenditure (EE), which was then used to classify activity intensity in METs. Resting energy expenditure was measured at visit 2 during the first three lying activities of the protocol. The observed mean resting energy expenditure was 2.89 ml·kg·min-1, which was used to define 1 MET. This value is comparable with previous calibration studies in older adults (Barnett et al., 2016; Evenson et al., 2015; Sergi et al., 2010; Siervo et al., 2014), and is consistent with the expected decrease in RMR associated with ageing (Byrne, Hills, Hunter, Weinsier, & Schutz, 2005; Kwan, Woo, & Kwok,

2004). MVPA was defined as an intensity of 3 METs and above (e.g., ≥8.68 $ml \cdot kg \cdot min^{-1}$) (Shephard, 2011).

Data Analysis

Cutpoint calibration. A randomly counter-balanced sample of 17 participants (12 female, five male) from visit 1, and 17 participants (12 female, five male) from visit 2 provided the calibration data (N = 34). Descriptive statistics for all devices were calculated for each activity in the protocol. The activPAL sedentary events and 3 MET VO_2 values were used as the criterion reference standards for SB and MVPA, respectively. SB and MVPA were each coded as either 0 or 1, where 1 represented the behaviour occurring and 0 represented the behaviour not occurring. Receiver Operating Characteristic (ROC) curve analyses (Jago, Zakeri, Baranowski, & Watson, 2007) were used to determine SB and MVPA cutpoints. Area under the curve (AUC) was calculated for each analysis as a measure of diagnostic accuracy. AUC values of; ≥ 0.90 are considered excellent, 0.80–0.89 good, 0.70–0.79 fair, and < 0.70 poor (Metz, 1978). For each device two different pairs of cutpoints were generated by analysing combinations of sensitivity (Se) and specificity (Sp) on the ROC curves. Our aim was to determine a cutpoint that accurately captured SB and MVPA (Se) whilst limiting misclassification of SB and MVPA (Sp). Two approaches were adopted to achieve this. Firstly, ENMO values that maximised both Se and Sp (Youden index) (Perkins & Schisterman, 2006) were identified as one set of cutpoints (SB_{Youden} and $MVPA_{Youden}$). The Youden index can however result in low positive predicted values (Evenson et al., 2015) and it is recommended that researchers consider the relative importance of Se and Sp (Welk, Laurson,

Eisenmann, & Cureton, 2011), and implications of the selected cutpoints on the biobehavioural impact on the outcome variables (Mackintosh, Fairclough, Stratton, & Ridgers, 2012). Secondly, cutpoints were determined that emphasised Se over Sp for SB cutpoints (SB_{Se}) to minimise the likelihood of classifying SB as PA, with Sp emphasised over Se for MVPA cutpoints ($MVPA_{Sp}$) to reduce the likelihood of misclassifying light PA as MVPA. Both cutpoints reflected recommendations that the lower Se or Sp values should be ≥60% (Lugade, Fortune, Morrow, & Kaufman, 2014). This prioritization approach minimises the risk of individuals being misclassified in the target behaviour and is common in accelerometer calibration (Landry et al., 2015; Mackintosh et al., 2012; Nero, Wallén, Franzén, Ståhle, & Hagströmer, 2015) and fitness standards research (Welk et al., 2011). There was no *a priori* preference for selecting either Youden or Se/Sp-based cutpoints, but instead we wanted to examine the effect of both approaches and make an informed decision as to which may be most suitable once the data were analysed.

Cross-validation. To be consistent with good practice guidelines suggested by Welk (2005), SB_{Youden} and $MVPA_{Youden}$, and SB_{Se} and $MVPA_{Sp}$ cutpoints were cross-validated. Cross-validation analyses were performed with data from the 17 participants (12 female, five male) whose visit 1 data was not used in the calibration analysis, and the 17 participants (12 female, five male) whose visit 2 data was not used in the calibration analysis (N = 34). Two-by-two (2x2) contingency tables were used to check classification agreement. The criterion measure and ENMO data were first categorised into sedentary/not sedentary and MVPA/not MVPA binary codes. Computed Se and Sp, Cohen's kappa coefficients, and percentage agreement between classifications were also assessed. Statistical analyses were performed using

IBM SPSS Statistics, version 24 (IBM, Armonk, NY), with the level of statistical significance set at $p \leq 0.05$.

Results

Participant characteristics

Of the 24 female and 10 male participants, 20 were classed as healthy (five male), seven were classed as overweight (three male), and five were classed as obese (two male) (NHS, 2016). Further descriptive characteristics of the sample are presented in Table 2. The mean (SD) energy expenditure and accelerometer output from GA and AG accelerometers are presented in Table 3 for each of phase of the laboratory protocol.

[INSERT TABLE 2 ABOUT HERE]

[INSERT TABLE 3 ABOUT HERE]

ROC Curve Analysis

ROC curve analysis revealed an AUC for the GA of 0.88 (95% CI: 0.87-0.88; $P < 0.001$) for SB (Figure 1) and 0.88 (95% CI: 0.87-0.88; $P < 0.001$) for MVPA (Figure 2). For the AG the AUC for SB was 0.90 (95% CI: 0.90-0.91; $P < 0.001$) (Figure 3), and 0.94 (95% CI: 0.94-0.95; $P < 0.001$) for MVPA (Figure 4).

[INSERT FIGURES 1, 2, 3, AND 4 ABOUT HERE]

Cutpoint generation

Table 4 displays all the generated cutpoints. The GA cutpoints which maximised both Se and Sp using the Youden Index were SB_{Youden}: 20 mg (Se = 94%, Sp = 72%) and $MVPA_{Youden}$: 32 mg (Se = 88%, Sp = 77%). AG cutpoints were 6 mg (Se = 85%, Sp = 82%) for SB_{Youden} and 19 mg (Se = 86%, Sp = 92%) for $MVPA_{Youden}$. The cutpoints optimising Se and Sp for GA were 57 mg (Se = 99%, Sp = 60%) for SB_{Se} and 104 mg (Se = 60%, Sp = 89%) for $MVPA_{Sp}$. Respective AG cutpoints for SB_{Se} and $MVPA_{Sp}$ were 15 mg (Se = 98%, Sp = 60%) and 69 mg (Se = 60%, Sp = 99%).

Cross-validation

The classification agreement, sensitivity, specificity, and kappa coefficients between calibration and cross-validation data for SB and MVPA cutpoints are shown in Table 4. GA SB_{Youden} (Se = 47%, Sp = 92%) and $MVPA_{Youden}$ (Se = 76%, Sp = 76%) cutpoints demonstrated moderate percentage agreement (73.1–76.2%) and moderate kappa scores (0.42–0.52). AG SB_{Youden} (Se = 62%, Sp = 93%) and $MVPA_{Youden}$ (Se = 86%, Sp = 89%) cutpoints demonstrated high percentage agreement (83.3–87.3%) and moderate to substantial kappa scores (0.59–0.75). Comparatively, lower percentage agreement and kappa scores were observed for both GA SB_{Se} and $MVPA_{Sp}$ (67.2-68.9%, k = 0.36-0.38), and AG SB_{Se} and $MVPA_{Sp}$ (73.2-80.4%, k = 0.46-0.60) cutpoints, respectively.

[INSERT TABLE 4 ABOUT HERE]

Discussion

This is the first study to determine GA wrist- and AG hip-worn raw acceleration cutpoints for SB and MVPA in older adults. ROC curve analyses revealed that wrist GA and hip AG accelerometer raw acceleration cutpoints provide good and excellent discriminations of SB and MVPA, respectively. The SB_{Se} and $MVPA_{Sp}$ cutpoints were comparable to those reported previously in younger adults (Hildebrand et al., 2014; Hildebrand et al., 2016; Menai et al., 2017). The Se values of 99% for GA SB_{Se} and 98% for AG SB_{Se} cutpoints ensure that almost all older adults who are sedentary have ENMO values below the established cutpoints, and therefore have a very low risk of being misclassified as being physically active. The de-emphasis on Sp (e.g., % of older adults correctly identified as not being sedentary) acknowledges the risk that a proportion (up to 40%) of older adults who are physically active may be classified as sedentary. However, SB is an identifiable risk factor affecting physical (e.g., premature mortality, chronic diseases, and all-cause dementia risk) and psychosocial (e.g., self-perceived quality of life, wellbeing, and self-efficacy) determinants of health (Edwards & Loprinzi, 2016; Falck et al., 2016; Lewis et al., 2017) independent of PA (Tremblay et al., 2017). Such misclassification is likely to be beneficial if already physically active older adults were offered further opportunities to take part in interventions to reduce SB and increase PA levels (Chastin et al., 2017). Conversely, the higher Sp values for the GA $MVPA_{Sp}$ cutpoint (89%) and AG $MVPA_{Sp}$ cutpoint (99%) ensures that older adults not engaged in MVPA (e.g., ENMO values below the established cutpoints) are not falsely

classified as being in MVPA and are correctly identified and targeted for PA-promoting interventions (Lyons, Swartz, Lewis, Martinez, & Jennings, 2017). The de-emphasis on Se suggests that up to 40% of older adults who are not in MVPA could have ENMO values above this cutpoint (due to the lower true positive rate, relative to the true negative rate (Sp)), and therefore their behaviour could be misclassified as MVPA (Nero et al., 2015). However, it is more likely to be harmful for an older adult to be wrongly classified as active rather than asking active older adults to take part in additional MVPA. Hence, the goal of the MVPA cutpoint to identify older adults who may have increased health risks due to being below this cutpoint (by favouring Sp over Se) appears to be justified (Nero et al., 2015; Welk, Going, Morrow, & Meredith, 2011).

The SB_{Youden} and $MVPA_{Youden}$ cutpoints yielded greater cross-validation classification accuracy. However, they are significantly lower than existing ENMO adult SB cutpoints for GA wrist-worn (46 mg) and AG hip-worn (47 mg) accelerometers (Hildebrand et al., 2016), and MVPA cutpoints for GA wrist-worn accelerometers in adults (93 mg; Hildebrand et al., 2014) and older adults (100 mg; Menai et al., 2017), and AG hip-worn accelerometers in adults (69.1 mg; Hildebrand et al., 2014). The mean age of the participants of 69.6 years was older than participants in both the Hildebrand et al. (2014) and Hildebrand et al. (2016) studies (34.2 years). Thus, the lower SB and MVPA cutpoints were anticipated given the lower RMR associated in older adults (Byrne et al., 2005; Kwan et al., 2004). Notwithstanding this, the SB_{Youden} and $MVPA_{Youden}$ cutpoints were substantially lower than those previously reported. If these cutpoints were applied in free-living environments it is likely that they would significantly underestimate SB and

unrealistically overestimate MVPA, resulting in participants being falsely classified as being physically active when they are more likely to be sedentary.

Given that acceleration magnitudes are significantly lower for the AG GT3X+ relative to the GA (John, Sasaki, Staudenmayer, Mavilia, & Freedson, 2013; Rowlands et al., 2015; Rowlands et al., 2016), the higher wrist-worn cutpoints relative to the hip-worn cutpoints were consistent with those observed previously (Rowlands et al., 2015; Stiles, Griew & Rowlands, 2013). Our protocol was comparable to previous calibration studies implemented in controlled settings (de Almeida Mendes, da Silva, Ramires Reichert, Martins, & Tomasi, 2017). However, laboratory calibration protocols rely on small deliberate increases in PA intensities and movement patterns within a limited period of time, compared to free-living activities over extended periods (van Hees, Golubic, Ekelund, & Brage, 2013a). Such protocols cannot fully reflect daily SB and PA patterns, and this may limit the accuracy of the SB and MVPA thresholds obtained for wrist- and hip-worn devices when they are applied in free-living environments (Van Hees et al., 2013a). This may have partially explained some of the modest results observed from the cross-validation phase, which repeated the laboratory protocol with the same participants.

One of the main decisions to be made by researchers using either raw acceleration or count-based outcomes is monitor placement location (de Almeida Mendes et al., 2017). After comparing SB and PA estimates from wrist- and hip-worn monitors with energy expenditure, Rosenberger et al. (2013) concluded that SB and MVPA classification accuracy was superior for the hip-worn devices. Our cross-validation

results support this due to the superior percent agreement and kappa scores for the hip-worn AG over the wrist-worn GA in classifying both SB and MVPA. However, our results also demonstrate that wrist-worn accelerometers can provide accurate estimates of SB and MVPA and the subsequent cutpoints performed reasonably well at discriminating both SB and MVPA (Troiano et al., 2014). Indeed, a recent systematic review of raw acceleration calibration studies reported no evidence of meaningful differences in the accuracy of wrist- and hip-worn accelerometers (de Almeida Mendes et al., 2017). Considering the superior wear compliance associated with wrist-worn devices (Doherty et al., 2017), this attachment site may be the most suitable location during free-living protocols.

Several strengths and limitations should be noted when interpreting the results of this study. A main strength was the use of raw acceleration data from two commonly used devices positioned at wrist and hip wear sites. These cutpoints will be of utility to researchers using the raw data capabilities of the GA and AG accelerometers to study SB and PA in older adults. We used best practice analytical procedures to calibrate and cross-validate the cutpoints (Welk, 2005), which were specific to adults aged 60 years and over. We also directly measured resting energy expenditure to allow a sample-specific interpretation of three METs as the MVPA threshold, and used a validated separate criterion measure for SB (Kim, Barry, & Kang, 2015). There were also a number of limitations. The sample may not have been representative of the wider older adult population in respect of their fitness status and motivation to engage in PA, as we recruited a convenience sample of healthy older adults who answered advertisements and showed an interest in the study representing a broad age range. Furthermore, specific older adult populations (e.g., those with

chronic diseases and impaired mobility) may require different SB and PA cutpoints (Landry et al., 2015) that reflect differences in RMR and energy cost during ambulatory PA across this age group (Miller, Strath, Swartz, & Cashin, 2010). Moreover, we incorporated activities that replicate everyday movements and tasks performed by older adults, but recognise that the laboratory setting limits the ecological validity of the resultant data (Hildebrand et al., 2016). The duration of the stepping and walking activities was relatively short and this may have limited the amount of MVPA data available for use in the analysis. Lastly, cross-validation was performed with the same participants using a repeated laboratory protocol rather than with an independent sample using a free-living or a simulated free-living protocol (Welk, 2005). It was felt that the challenges associated with having the participants wear the gas analysis system for an extended period in free-living situations were too great to warrant taking this approach. Consequently, the SB cutpoints which performed modestly in the cross-validation, should be interpreted with caution. Both sets of cutpoints should be further cross-validated in free-living environments (Welk, 2005) with independent samples over whole days using feasible criterion measures such as activPALs and wrist-worn heart rate monitors (Brage et al., 2015).

In conclusion, cutpoints varied dependent upon attachment site, with the wrist-worn GA cutpoints higher than those for the hip-worn AG. The identified GA and AG SB_{Se} and $MVPA_{Sp}$ cutpoints can enable researchers to classify older adults as engaging in SB or not engaging in MVPA, but the SB_{Se} cutpoints should be interpreted with a degree of caution due to their modest cross-validation results. More cross-validation research in independent samples within free-living environments is needed to further test the utility of the SB_{Se} and $MVPA_{Sp}$ cutpoints.

Disclosure statement

There are no conflicts of interest to report.

Funding

This work was funded by Sefton Metropolitan Borough Council and Edge Hill University and received institutional ethical approval from Edge Hill University under SPA-REC-2016-343.

References

ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. (2002). Statement AT: guidelines for the six-minute walking-test. *American Journal of Respiratory and Critical Care Medicine*, *166*, 111-117.

Bakrania, K., Yates, T., Rowlands, A. V., Esliger, D. W., Bunnewell, S., Sanders, J., ... & Edwardson, C. L. (2016). Intensity thresholds on raw acceleration data: euclidean norm minus one (ENMO) and mean amplitude deviation (MAD) approaches. *PloS one*, *11*(10), e0164045.

Banda, J. A., Haydel, K. F., Davila, T., Desai, M., Bryson, S., Haskell, W. L., . . . Robinson, T. N. (2016). Effects of varying epoch lengths, wear time algorithms, and activity cut-points on estimates of child sedentary behavior and physical activity from accelerometer data. *PLOS ONE, 11*(3), e0150534.

Barber, S.E., Forster, A. & Birch, K.M. (2015). Levels and patterns of daily physical activity and sedentary behavior measured objectively in older care home residents in the United Kingdom. *Journal of aging and physical activity*, *23*(1), 133-143.

Barnett, A., van den Hoek, D., Barnett, D. & Cerin, E. (2016). Measuring moderate-intensity walking in older adults using the ActiGraph accelerometer. *BMC geriatrics*, *16*(1), 211-219.

Berkemeyer, K., Wijndaele, K., White, T., Cooper, A.J.M., Luben, R., Westgate, K., Griffin, S.J., Khaw, K.T., Wareham, N.J. & Brage, S. (2016). The descriptive epidemiology of accelerometer-measured physical activity in older adults. *International Journal of Behavioral Nutrition and Physical Activity*, *13*(1), 2-11.

569 Blodgett, J., Theou, O., Kirkland, S., Andreou, P. & Rockwood, K. (2015). The
570 association between sedentary behaviour, moderate–vigorous physical activity and
571 frailty in NHANES cohorts. *Maturitas*, *80*(2), 187-191.

572 Bradbury, K. E., Guo, W., Cairns, B. J., Armstrong, M. E., & Key, T. J. (2017).
573 Association between physical activity and body fat percentage, with adjustment for
574 BMI: a large cross-sectional analysis of UK Biobank. *BMJ Open*, *7*(3), e011843.

575 Brage, S., Westgate, K., Franks, P. W., Stegle, O., Wright, A., Ekelund, U., &
576 Wareham, N. J. (2015). Estimation of free-living energy expenditure by heart rate
577 and movement sensing: a doubly-labelled water study. *PloS one*, *10*(9), e0137206.

578 Buman, M.P., Hekler, E.B., Haskell, W.L., Pruitt, L., Conway, T.L., Cain, K.L.,
579 Sallis, J.F., Saelens, B.E., Frank, L.D., & King, A.C. (2010). Objective light-
580 intensity physical activity associations with rated health in older adults. *American*
581 *Journal of Epidemiology*, *172*(10), 1155-1165.

582 Butte, N.F., Ekelund, U. & Westerterp, K.R. (2012). Assessing physical activity
583 using wearable monitors: measures of physical activity. *Medicine & Science in*
584 *Sports & Exercise*, *44*(1 Suppl 1), S5-S12.

585 Byrne, N.M., Hills, A.P., Hunter, G.R., Weinsier, R.L. & Schutz, Y. (2005).
586 Metabolic equivalent: one size does not fit all. *Journal of Applied physiology*, *99*(3),
587 1112-1119.

588 Cardinal, B.J. & Cardinal, M.K. (2000). Preparticipation physical activity screening
589 within a racially diverse, older adult sample: comparison of the original and revised
590 physical activity readiness questionnaires. *Research quarterly for exercise and*
591 *sport*, *71*(3), 302-307.

592 Cardinal, B.J., Esters, J., & Cardinal, M.K. (1996). Evaluation of the revised
593 physical activity readiness questionnaire in older adults. *Medicine and science in*
594 *sports and exercise*, *28*(4), 468-472.

595 Celis-Morales, C. A., Lyall, D. M., Steell, L., Gray, S. R., Iliodromiti, S., Anderson,
596 J., . . . Gill, J. M. R. (2018). Associations of discretionary screen time with mortality,
597 cardiovascular disease and cancer are attenuated by strength, fitness and physical
598 activity: findings from the UK Biobank study. *BMC Medicine, 16*(1), 77-90.

599 Centers for Disease Control and Prevention. (2015). How Much Physical Activity do
600 Older Adults Need?, Retrieved from
601 https://www.cdc.gov/physicalactivity/basics/older_adults/.

602 Chastin, S. F. M., Dontje, M. L., Skelton, D. A., Čukić, I., Shaw, R. J., Gill, J. M., ...
603 & Dall, P. M. (2018). Systematic comparative validation of self-report measures of
604 sedentary time against an objective measure of postural sitting
605 (activPAL). *International Journal of Behavioral Nutrition and Physical*
606 *Activity*, *15*(1), 21-32.

607 Chastin, S., Gardiner, P.A., Ashe, M.C., Harvey, J.A., Leask, C.F., Balogun, S.,
608 Helbostad, J.L. & Skelton, D.A. (2017). Interventions for reducing sedentary
609 behaviour in community-dwelling older adults. *The Cochrane Library*, 1-13.

610 Chastin, S.F.M., Culhane, B. and Dall, P.M. (2014). Comparison of self-reported
611 measure of sitting time (IPAQ) with objective measurement
612 (activPAL). *Physiological measurement*, *35*(11), 2319-2327.

613 Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.,
614 413). Hillsdale (NJ): Lawrence Erlbaum Associates.

615 Copeland, J.L. & Esliger, D.W. (2009). Accelerometer assessment of physical
616 activity in active, healthy older adults. *Journal of aging and physical activity*, *17*(1),
617 17-30.

618 de Almeida Mendes, M., da Silva, I.C.M., Ramires, V., Reichert, F., Martins, R. &
619 Tomasi, E. (2017). Calibration of raw accelerometer data to measure physical
620 activity: a systematic review. *Gait & posture*, *61*, 98-110.

621 Department of Health. (2011). Physical Activity Guidelines in the UK: Review and
622 Recommendations. Retrieved from
623 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/21374
624 3/dh_128255.pdf.

625 Devereux-Fitzgerald, A., Powell, R., Dewhurst, A. & French, D.P. (2016). The
626 acceptability of physical activity interventions to older adults: a systematic review
627 and meta-synthesis. *Social Science & Medicine*, *158*, 14-23.

628 Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M.H., White,
629 T., van Hees, V.T., Trenell, M.I., Owen, C.G. & Preece, S.J. (2017). Large scale
630 population assessment of physical activity using wrist worn accelerometers: The UK
631 Biobank Study. *PloS one*, *12*(2), p.e0169649.

632 Edwards, M.K. & Loprinzi, P.D. (2016). The association between sedentary
633 behavior and cognitive function among older adults may be attenuated with adequate
634 physical activity. *Journal of Physical Activity and Health*, *14*(1), 52-58.

635 Esliger, D. W., Rowlands, A. V., Hurst, T. L., Catt, M., Murray, P., & Eston, R. G.
636 (2011). Validation of the GENEA Accelerometer. *Medicine and Science in Sports*
637 *and Exercise, 43*(6), 1085-1093.

638 Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008).
639 Calibration of two objective measures of physical activity for children. *Journal of*
640 *Sports Sciences, 26*, 1557–1565.

641 Evenson, K.R., Wen, F., Herring, A.H., Di, C., LaMonte, M.J., Tinker, L.F., Lee,
642 I.M., Rillamas-Sun, E., LaCroix, A.Z. & Buchner, D.M. (2015). Calibrating physical
643 activity intensity for hip-worn accelerometry in women age 60 to 91 years: The
644 Women's Health Initiative OPACH Calibration Study. *Preventive medicine*
645 *reports, 2*, 750-756.

646 Fairclough, S., Noonan, R., Rowlands, A., Van Hees, V., Knowles, Z. & Boddy, L.
647 (2016). Wear compliance and activity in children wearing wrist and hip mounted
648 accelerometers. *Medicine & Science in Sports & Exercise*, *48*(2), 245-253.

649 Falck, R.S., Davis, J.C. & Liu-Ambrose, T. (2016). What is the association between
650 sedentary behaviour and cognitive function? A systematic review. *British Journal of*
651 *Sports Medicine, 51*(10), 800-811.

652 Gorman, E., Hanson, H.M., Yang, P.H., Khan, K.M., Liu-Ambrose, T. & Ashe,
653 M.C. (2014). Accelerometry analysis of physical activity and sedentary behavior in
654 older adults: a systematic review and data analysis. *European Review of Aging and*
655 *Physical Activity*, *11*(1), 35-49.

656 Greaney, M.L., Lees, F.D., Blissmer, B.J., Riebe, D. & Clark, P.G. (2016).
657 Psychosocial factors associated with physical activity in older adults. *Annual Review*
658 *of Gerontology and Geriatrics, 36*(1), 273-91.

659 Hall, K.S., Howe, C.A., Rana, S.R., Martin, C.L. & Morey, M.C. (2013). METs and
660 accelerometry of walking in older adults: standard versus measured energy
661 cost. *Medicine and science in sports and exercise*, *45*(3), 574-582.

662 Hallal, P.C., Andersen, L.B., Bull, F.C., Guthold, R., Haskell, W., Ekelund, U., &
663 Lancet Physical Activity Series Working Group. (2012). Global physical activity
664 levels: surveillance progress, pitfalls, and prospects. *The lancet*, *380*(9838), 247-257.

665 Hargens, T. A., Deyarmin, K. N., Snyder, K. M., Mihalik, A. G., & Sharpe, L. E.
666 (2017). Comparison of wrist-worn and hip-worn activity monitors under free living
667 conditions. *Journal of medical engineering & technology*, *41*(3), 200-207.

668 Harvey, J. a, Chastin S, F. M., & Skelton, D. A. (2014). How sedentary are older
669 people? a systematic review of the amount of sedentary behavior. *Journal of Aging
670 and Physical Activity*, *23*(3), 471–87.

671 Hildebrand, M., Hansen, B. H., van Hees, V. T., & Ekelund, U. (2017). Evaluation
672 of raw acceleration sedentary thresholds in children and adults. *Scandinavian journal
673 of medicine & science in sports*, *27*(12), 1814-1823.

674 Hildebrand, M., Van Hees, V. T., Hansen, B. H., & Ekelund, U. (2014). Age group
675 comparability of raw accelerometer output from wrist- and hip-worn monitors.
676 *Medicine and Science in Sports and Exercise*, *46*(9), 1816-1824.

677 Innes, E. & Straker, L. (1999). Validity of work-related assessments. *Work*, *13*(2),
678 125-152.

679 Jago, R., Zakeri, I., Baranowski, T. & Watson, K. (2007). Decision boundaries and
680 receiver operating characteristic curves: new methods for determining accelerometer
681 cutpoints. *Journal of sports sciences*, *25*(8), 937-944.

682 John, D., Sasaki, J., Staudenmayer, J., Mavilia, M., & Freedson, P. S. (2013).
683 Comparison of raw acceleration from the GENEA and ActiGraph™ GT3X+ activity
684 monitors. *Sensors*, *13*(11), 14754-14763.

685 Kerr, J., Marshall, S. J., Godbole, S., Chen, J., Legge, A., Doherty, A. R., . . . Foster,
686 C. (2013). Using the SenseCam to improve classifications of sedentary behavior in
687 free-living settings. *American Journal of Preventive Medicine, 44*(3), 290-296.

688 Kim, J., Im, J.S. & Choi, Y.H. (2016). Objectively measured sedentary behavior and
689 moderate-to-vigorous physical activity on the health-related quality of life in US
690 adults: The National Health and Nutrition Examination Survey 2003–2006. *Quality*
691 *of Life Research, 26*(5), 1315-1326.

692 Kim, Y., Barry, V. W., & Kang, M. (2015). Validation of the ActiGraph GT3X and
693 activPAL accelerometers for the assessment of sedentary behavior. *Measurement in*
694 *Physical Education and Exercise Science, 19*(3), 125-137.

695 Ko, S.U., Jerome, G.J., Simonsick, E.M., Studenski, S. & Ferrucci, L. (2018).
696 Differential gait patterns by falls history and knee pain status in healthy older adults:
697 results from the Baltimore longitudinal study of aging. *Journal of aging and physical*
698 *activity*, *0*(0), 1-18.

699 Koolhaas, C. M., van Rooij, F. J., Cepeda, M., Tiemeier, H., Franco, O. H., &
700 Schoufour, J. D. (2018). Physical activity derived from questionnaires and wrist-
701 worn accelerometers: comparability and the role of demographic, lifestyle, and
702 health factors among a population-based sample of older adults. *Clin Epidemiol, 10*,
703 1-16.

704 Koster, A., Shiroma, E. J., Caserotti, P., Matthews, C. E., Chen, K. Y., Glynn, N. W.,
705 & Harris, T. B. (2016). Comparison of sedentary estimates between activPAL and
706 hip-and wrist-worn ActiGraph. *Medicine and science in sports and exercise*, *48*(8),
707 1514-1522.

Kozey-Keadle, S., Libertine, A., Lyden, K., Staudenmayer, J., & Freedson, P. S. (2011). Validation of wearable monitors for assessing sedentary behavior. *Medicine & Science in Sports & Exercise*, *43*(8), 1561-1567.

Ku, P. W., Steptoe, A., Liao, Y., Sun, W. J., & Chen, L. J. (2018). Prospective relationship between objectively measured light physical activity and depressive symptoms in later life. *International journal of geriatric psychiatry*, *33*(1), 58-65.

Kwan, M., Woo, J. & Kwok, T. (2004). The standard oxygen consumption value equivalent to one metabolic equivalent (3.5 ml/min/kg) is not appropriate for elderly people. *International journal of food sciences and nutrition*, *55*(3), 179-182.

Landry, G.J., Falck, R.S., Beets, M.W. & Liu-Ambrose, T. (2015). Measuring physical activity in older adults: calibrating cutpoints for the MotionWatch 8©. *Frontiers in aging neuroscience*, *7*, 166-173.

Lewis, B.A., Napolitano, M.A., Buman, M.P., Williams, D.M. & Nigg, C.R. (2017). Future directions in physical activity intervention research: expanding our focus to sedentary behaviors, technology, and dissemination. *Journal of behavioral medicine*, *40*(1), 112-126.

Lohne-Seiler, H., Hansen, B.H., Kolle, E. & Anderssen, S.A. (2014). Accelerometer-determined physical activity and self-reported health in a population of older adults (65–85 years): a cross-sectional study. BMC Public Health, 14(1), 284-293.

Lugade, V., Fortune, E., Morrow, M. & Kaufman, K. (2014). Validity of using tri-axial accelerometers to measure human movement—Part I: Posture and movement detection. *Medical Engineering and Physics*, *36*(2), 169-176.

730 Lyons, E. J., Swartz, M. C., Lewis, Z. H., Martinez, E., & Jennings, K. (2017).
731 Feasibility and acceptability of a wearable technology physical activity intervention
732 with telephone counseling for mid-aged and older adults: a randomized controlled
733 pilot trial. *JMIR mHealth and uHealth*, *5*(3), 28-63.

734 Mackintosh, K. A., Fairclough, S. J., Stratton, G., & Ridgers, N. D. (2012). A
735 calibration protocol for population-specific accelerometer cut-points in children.
736 *PLOS ONE, 7*(5), e36919.

737 Mañas, A., del Pozo-Cruz, B., Guadalupe-Grau, A., Marín-Puyalto, J., Alfaro-Acha,
738 A., Rodríguez-Mañas, L., García-García, F.J. & Ara, I. (2017). Reallocating
739 accelerometer-assessed sedentary time to light or moderate-to vigorous-intensity
740 physical activity reduces frailty levels in older adults: An isotemporal substitution
741 approach in the TSHA study. *Journal of the American Medical Directors*
742 *Association,* doi: 10.1016/j.jamda.2017.11.003

743 Martinez-Gomez, D., Bandinelli, S., Del-Panta, V., Patel, K. V., Guralnik, J. M., &
744 Ferrucci, L. (2017). Three-year changes in physical activity and decline in physical
745 performance over 9 years of follow-up in older adults: The Invecchiare in Chianti
746 Study. *Journal of the American Geriatrics Society*, *65*(6), 1176-1182.

747 Matchar, D,B., & Orlando, L,A. (2007). The relationship between test and outcome.
748 In: Price CP, editor. *Evidence-Based Laboratory Medicine; Principles, Practice and*
749 *Outcomes*, 2nd Edition. Washington DC, USA: AACC Press, 53-66.

750 Mathie, M.J., Coster, A.C., Lovell, N.H. & Celler, B.G. (2004). Accelerometry:
751 providing an integrated, practical method for long-term, ambulatory monitoring of
752 human movement. *Physiological measurement*, *25*(2), R1-R20.

753 Matthew, C.E. (2005). Calibration of accelerometer output for adults. *Medicine &*
754 *Science in Sports & Exercise, 37*(11 Suppl), S512-522.

755 Menai, M., Van Hees, V.T., Elbaz, A., Kivimaki, M., Singh-Manoux, A. & Sabia, S.
756 (2017). Accelerometer assessed moderate-to-vigorous physical activity and
757 successful ageing: results from the Whitehall II study. *Scientific reports, 7*, 45772-
758 45780.

759 Metz, C.E. (1978) Basic principles of ROC analysis. In *Seminars in nuclear*
760 *medicine* (Vol. 8, No. 4, 283-298). Elsevier.

761 Miller, N.E., Strath, S.J., Swartz, A.M. & Cashin, S.E. (2010). Estimating absolute
762 and relative physical activity intensity across age via accelerometry in
763 adults. *Journal of aging and physical activity*, *18*(2), 158-170.

764 Murphy, S.L. (2009). Review of physical activity measurement using accelerometers
765 in older adults: considerations for research design and conduct. *Preventive*
766 *medicine*, *48*(2), 108-114.

767 National Health Service. (2015). *Physical activity guidelines for older adults*.
768 Retrieved from http://www.nhs.uk/Livewell/fitness/Pages/physical-activity-
769 guidelines-for-older-adults.aspx.

770 National Health Service. (2016). *What is the body mass index (BMI)?*. Retrieved
771 from https://www.nhs.uk/common-health-questions/lifestyle/what-is-the-body-mass-
772 index-bmi/.

773 Nero, H., Wallén, M.B., Franzén, E., Ståhle, A. & Hagströmer, M. (2015).
774 Accelerometer cutpoints for physical activity assessment of older adults with
775 Parkinson's disease. *PloS one, 10*(9), 1-11.

Oguma, Y., Osawa, Y., Takayama, M., Abe, Y., Tanaka, S., Lee, I.M. & Arai, Y. (2017). Validation of questionnaire-assessed physical activity in comparison with objective measures using accelerometers and physical performance measures among community-dwelling adults aged≥ 85 years in Tokyo, Japan. *Journal of Physical Activity and Health*, *14*(4), 245-252.

Perkins, N. J., & Schisterman, E. F. (2006). The inconsistency of "optimal" cutpoints obtained two criteria based on the receiver operating characteristic curve. *American Journal of Epidemiology*, *163*(7), 670-675.

Phillips, L. R., Parfitt, G., & Rowlands, A. V. (2013). Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *Journal of science and medicine in sport*, *16*(2), 124-128.

Prince, S.A., Adamo, K.B., Hamel, M.E., Hardt, J., Gorber, S.C. & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 56-79.

Public Health England. (2017). *Health Report 2017: Sefton.* Retrieved from http://fingertipsreports.phe.org.uk/health-profiles/2017/e08000014.pdf.

Rosenberger, M.E., Haskell, W.L., Albinali, F., Mota, S., Nawyn, J. & Intille, S. (2013). Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Medicine and science in sports and exercise*, *45*(5), 964-975.

Rowlands, A.V., Fraysse, F., … Catt, M. (2015). Comparison of measured acceleration output from accelerometry-based activity monitors. *Medicine and science in sports and exercise*, *47*, 201-210.

799 Rowlands, A. V., Mirkes, E. M., Yates, T., Clemes, S., Davies, M., Khunti, K., &
800 Edwardson, C. L. (2018). Accelerometer-assessed physical activity in epidemiology:
801 are monitors equivalent?. *Medicine & Science in Sports & Exercise*, *50*(2), 257-265.

802 Rowlands, A. V., Yates, T., Davies, M., Khunti, K. & Edwardson, C. L. (2016). Raw
803 accelerometer data analysis with GGIR R-package: Does accelerometer brand
804 matter? *Medicine and science in sports and exercise, 48*, 1935–1941.

805 Sallis, J. F., Owen, N., & Fotheringham, M. J. (2000). Behavioral epidemiology: a
806 systematic framework to classify phases of research on health promotion and disease
807 prevention. *Annals of Behavioral Medicine*, *22*(4), 294-298.

808 Santos-Lozano, A., Torres-Luque, G., Marín, P.J., Ruiz, J.R., Lucia, A. &
809 Garatachea, N. (2012). Intermonitor variability of GT3X
810 accelerometer. *International journal of sports medicine*, *33*(12), 994-999.

811 Schaefer, C. A., Nigg, C. R., Hill, J. O., Brink, L. A., & Browning, R. C. (2014).
812 Establishing and evaluating wrist cutpoints for the GENEActiv accelerometer in
813 youth. *Medicine and Science in Sports and Exercise, 46*(4), 826-833.

814 Sellers, C., Dall, P., Grant, M., & Stansfield, B. (2016). Validity and reliability of the
815 activPAL3 for measuring posture and stepping in adults and young people. *Gait &*
816 *Posture*, *43*, 42-47.

817 Sergi, G., Coin, A., Sarti, S., Perissinotto, E., Peloso, M., Mulone, S., Trolese, M.,
818 Inelmen, E.M., Enzi, G. & Manzato, E. (2010). Resting VO2, maximal VO2 and
819 metabolic equivalents in free-living healthy elderly women. *Clinical nutrition*, *29*(1),
820 84-88.

Shaw, R.J., Čukić, I., Deary, I.J., Gale, C.R., Chastin, S.F., Dall, P.M., Skelton, D.A., & Der, G. (2017). Relationships between socioeconomic position and objectively measured sedentary behaviour in older adults in three prospective cohorts. *BMJ Open, 7*(6), 1-11.

Shephard, R. (2011). Compendium of physical activities: A Second update of codes and MET values. *Yearbook Of Sports Medicine*, *2012*, 126-127.

Siervo, M., Bertoli, S., Battezzati, A., Wells, J.C., Lara, J., Ferraris, C. & Tagliabue, A. (2014). Accuracy of predictive equations for the measurement of resting energy expenditure in older subjects. *Clinical nutrition*, *33*(4), 613-619.

Sparling, P.B., Howard, B.J., Dunstan, D.W. & Owen, N. (2015). Recommendations for physical activity in older adults. *BMJ: British Medical Journal (Online*), *350,* 1-5.

Stiles, V. H., Griew, P. J., & Rowlands, A. V. (2013). Use of accelerometry to classify activity beneficial to bone in premenopausal women. *Medicine and science in sports and exercise*, *45*(12), 2353-2361.

Strain, T., Fitzsimons, C., Foster, C., Mutrie, N., Townsend, N., & Kelly, P. (2016). Age-related comparisons by sex in the domains of aerobic physical activity for adults in Scotland. *Preventive medicine reports*, *3*, 90-97.

Szklo, M. & Nieto, F.J. (2007). *Stratification and adjustment: Multivariate analysis in epidemiology*, in Epidemiology: Beyond the Basics. Gaithersburg, MD, Aspen Publishers, 227-295.

Taylor, L.M., Klenk, J., Maney, A.J., Kerse, N., MacDonald, B.M. & Maddison, R. (2014). Validation of a body-worn accelerometer to measure activity patterns in octogenarians. *Archives of physical medicine and rehabilitation*, *95*(5), 930-934.

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., … Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) - Terminology consensus project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1), 75-92.

Treuth, M.S., Schmitz, K., Catellier, D.J., McMurray, R.G., Murray, D.M., Almeida, M.J., Going, S., Norman, J.E. & Pate, R. (2004). Defining accelerometer thresholds for activity intensities in adolescent girls. *Medicine and science in sports and exercise*, *36*(7), 1259-1266.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, *40*(1), 181–8.

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British Journal of Sports Medicine, 48*(13), 1019-1023.

Vähä-Ypyä, H., Vasankari, T., Husu, P., Mänttäri, A., Vuorimaa, T., Suni, J., & Sievänen, H. (2015). Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based mean amplitude deviation (MAD). *PLoS One*, *10*(8), e0134813.

van Hees, V. T., Fang, Z., Langford, J., Assah, F., Mohammad, A., da Silva, I. C., . . . Brage, S. (2014). Autocalibration of accelerometer data for free-living physical

activity assessment using local gravity and temperature: an evaluation on four continents. *Journal of Applied Physiology, 117*(7), 738-744.

van Hees, V. T., Golubic, R., Ekelund, U., & Brage, S. (2013a). Impact of study design on development and evaluation of an activity-type classifier. *Journal of Applied Physiology, 114*(8), 1042-1051.

van Hees, V. T., Gorzelniak, L., Dean Leon, E. C., Eder, M., Pias, M., Taherian, S., . . . Brage, S. (2013b). Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLoS One, 8*(4), e61691.

van Hees, V. T., Renstrom, F., Wright, A., Gradmark, A., Catt, M., Chen, K. Y., . . . Franks, P. W. (2011). Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *Plos One, 6*, e22922.

Varela Mato, V., Yates, T., Stensel, D., Biddle, S. and Clemes, S.A. (2017). Concurrent validity of Actigraph-determined sedentary time against the Activpal under free-living conditions in a sample of bus drivers. *Measurement in Physical Education and Exercise Science*, *21*(4), 212-222.

Welk, G. J. (2005). Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Medicine Science Sports Exercise, 37*(11), S501-S511.

Welk, G. J., Eisenmann, J. C., Schaben, J., Trost, S. G., & Dale, D. (2007). Calibration of the biotrainer pro activity monitor in children. *Pediatr Exerc Sci, 19*(2), 145-158.

887 Welk, G. J., Going, S. B., Morrow, J. R., & Meredith, M. D. (2011). Development of
888 new criterion-referenced fitness standards in the FITNESSGRAM® program:
889 rationale and conceptual overview. *American journal of preventive medicine*, *41*(4),
890 S63-S67.

891 Welk, G. J., Laurson, K. R., Eisenmann, J. C., & Cureton, K. J. (2011). Development
892 of youth aerobic-capacity standards using receiver operating characteristic curves.
893 *American Journal of Preventive Medicine*, *41*(4, Supplement 2), S111-S116.

894 Welk, G.J., McClain, J. & Ainsworth, B.E. (2012). Protocols for evaluating
895 equivalency of accelerometry-based activity monitors. *Medicine and science in*
896 *sports and exercise*, *44*(1 Suppl 1), S39-49.

897 Wullems, J.A., Verschueren, S.M., Degens, H., Morse, C.I. & Onambélé, G.L.
898 (2016). A review of the assessment and prevalence of sedentarism in older adults, its
899 physiology/health impact and non-exercise mobility counter-
900 measures. *Biogerontology*, *17*(3), 547-565.

901 Wullems, J.A., Verschueren, S.M., Degens, H., Morse, C.I. & Onambélé, G.L.
902 (2017). Performance of thigh-mounted triaxial accelerometer algorithms in objective
903 quantification of sedentary behaviour and physical activity in older adults. *PloS*
904 *one*, *12*(11), e0188215.

905 Zhu, W., Wadley, V.G., Howard, V.J., Hutto, B., Blair, S.N. & Hooker, S.P. (2017).
906 Objectively measured physical activity and cognitive function in older
907 adults. *Medicine and science in sports and exercise*, *49*(1), 47-53.