Considerations in Processing Accelerometry Data to Explore Physical Activity and Sedentary Time in Older Adults

Considerations in Processing Accelerometry Data to Explore Physical Activity and Sedentary Time in Older Adults.

ABSTRACT

Processing decisions for accelerometry data can have important implications for outcome measures, yet little evidence exists exploring these in older adults. The aim of the current study was to investigate the impact of three potentially important criteria on older adults, physical activity and sedentary time. Participants (n=222: mean age 71.75years (SD=6.58), 57% male) wore ActiGraph GT3X+ for (7 days). Eight data processing combinations from three criteria were explored: low frequency extension (on/off), nonwear time (90/120-min) and intensity cut-points (moderate-to-vigorous physical activity ≥1,041and >2,000 counts/min)). Analyses included Wilcoxon Sign-Rank, paired *t* tests and correlation coefficients (significance, $p<0.05$). Results for low-frequency extension on, 90-min nonwear time and >1,041counts/min showed significantly higher light and moderate-to-vigorous physical activity and lower sedentary time. Cut-points had the greatest impact on physical activity and sedentary time. Processing criteria can significantly impact physical activity and/or sedentary time, potentially leading to data inaccuracies, preventing cross-study comparisons, and influencing the accuracy of population surveillance.

Keywords: accelerometer processing, light physical activity, methodology, moderate-to-vigorous physical activity; sedentary behaviour.

INTRODUCTION

Accelerometry as a device-based measure overcomes many of the challenges that self-reported measurement relies on, such as survey completion and accurate memory recall. This is particularly the case for older adults (≥60 years (UNFPA, 2012)) who may have difficulties with reading/vision, cognition and memory recall (Sallis et al., 2000; Troiano et al., 2008; Copeland & Esliger, 2009; Seymour et al., 2001; Hutto et al., 2013). In addition, accelerometry unlike self-report does not require individuals to differentiate their physical activity (PA) behaviours into differing intensities; for example, walking at a brisk pace (moderate PA) or jogging/running (vigorous PA). Differentiation can be challenging as their perception of intensity may differ with age (Sallis et al., 2000; Troiano et al., 2008; Copeland & Esliger, 2009; Seymour et al., 2001; Hutto et al., 2013;). Consequently, accelerometry has been found to be a valid and reliable tool that can measure raw bodily acceleration across multiple planes. When collecting accelerometry data there are different types of processing options: event based, raw acceleration or count-based. For the purpose of this study we will be concentrating on the count-based approach due to its prevalence in the literature and likely applicability to those wishing to use accelerometry not only for research but also those outside of academia working in the fields of policy and practice. When processing data and implementing a count-based method modern accelerometers allow for four data collection stages: a) collection and processing of raw acceleration data; b) transformation of raw acceleration into a digital representation such as ‘counts’; c) translation of counts into a physiological meaningful indicator, (e.g. intensity); and d) the presentation of data as minutes per hour/day/week for PA behaviours at various levels of intensity (Granat, 2012; Hutto et al., 2013; Migueles et al., 2017). The four aforementioned stages are possible with the use of proprietary software such as that developed by ActiGraph (ActiLife, ActiGraph, Pensacola, FL). However, it should be noted that although count-based measures have been implemented within the current study, many research groups have moved away from this method and are processing their raw accelerometry data with the use of statistical packages such as GGIR (R package) to transform the data collected into PA outputs.

Although accelerometers provide a feasible option for the measurement of PA and sedentary time (ST), it should be noted that they are not without issue and specific recommendations for use have yet to be made for older adults. Challenges can occur during two main phases: (a) data collection; and (b) data processing (Ward et al., 2005; Toftager et al., 2013). As part of the (a) data collection phase, researchers make decisions regarding device selection (cost, memory and battery life), placement, wear time (to ensure reliable estimates of PA), initiation settings (sampling rate, light-emitting diode options, idle sleep mode) and appropriate software packages to manage this phase (Warren et al., 2010). Over the last decade, however, vast improvements have been made with a larger range of devices available, better device specification and improved software packages (faster processing speed, quality and guided initiation processes) (Ward et al., 2005; Toftager et al., 2013). With the occurrence of positive changes regarding the quality of accelerometry use, the focus has now shifted from data collection decisions to challenges regarding decisions during the (b) data processing stage (Evenson et al., 2012; Toftager et al., 2013).

During the (b) data processing stage, it is vital that protocols are put in place to ensure that the accuracy of processed data reflects a participant's reality. This stage is not only reliant on the device functioning and recording of the participant's data, but also it is reliant on researcher decision making and the choices that are made prior to the processing of data (Crouter et al., 2006; Corbett et al., 2017). The first decision relates to the choice of either an epoch-based approach, or an event-based approach (Granat, 2012). In brief, the epoch-based approach consists of determining the activity intensity level for each short period of time (epoch) using cut-points established in trials, making it possible to link accelerometer counts to energy expenditure. Laboratory trials involve a pre-determined activity (e.g. walking, jogging, and running) were participants complete the experimental session by performing the activity in set conditions for set periods of time on a treadmill. Copeland & Esliger (2009) tasked older adults to walk (common activity of older adults) on a treadmill for three, 6-minute conditions at varying speeds (2.4, 3.2 and 4.8km·h^{-1}). Oxygen consumption was determined at rest (seated) and during each of the walking conditions and accelerometers worn (Copeland & Esliger, 2009). Following the testing stage average counts per minute and oxygen consumption were calculated, enabling mean accelerometer and oxygen uptake to be determined for each walking speed (Copeland & Esliger, 2009).

In comparison the event-based approach is conceptually different as it first identifies periods of similar acceleration (event) using various pattern recognition algorithms and/or machine learning techniques (Granat, 2012). This event-based approach is reliant upon the orientation of the device as its sensitive axis and gravity, and the shift in angle of the device. After identifying events, each event is classified into broad categories based on bodily positions (e.g. standing, sitting, lying, walking, running or cycling). However, this is a relatively new approach and the epoch-based approach remains popular, and more common, largely because its typical outcome measures are minutes of PA at varying levels of intensity that can be easily linked to PA recommendations. Although, even with a method that has a long history of use, commonly accepted standards are lacking, and many choices need to be made in each study.

As previously mentioned researchers are presented with a range of criteria to consider including: 1) low frequency extension (LFE) function; 2) nonwear time; and 3) intensity cut-points (Table 1) (Trost et al., 2005; Aguilar-Farias et al., 2014; Gorman et al., 2014; ActiGraph, 2016; Barnett et al., 2016; Aadland et al., 2018). The combination of these criteria in addition to three other criteria which have been previously established for older adults (epoch length, number of valid hours in a day and number of valid days in a week) have the potential for valuable data to be lost or miscalculated if unsuitable criterion decisions are made (Gorman et al., 2010; Evenson et al., 2012; Hutto et al., 2013). Such decisions are particularly relevant for specific population subgroups where legitimate reasons (e.g. body fat, gait issues, age, and sex) may influence the implications of the aforementioned criteria (Corbett et al., 2017; Migueles et al., 2017). As older adults are unlikely to be similar to adults due to a range of factors including health related issues (lower physiological function, muscle atrophy, reduced cardiorespiratory function etc.) and differing lifestyle behaviours (retirement, lack of a structured daily routine, sedentary hobbies etc.), this has the potential to influence their PA and/or ST (McPhee et al., 2016). For that reason, the decision to employ adult processing criteria is questionable and presents a strong argument for an age-specific tailored approach to data processing (Migueles et al., 2017). Furthermore, older adults can be heterogeneous in terms of age and/or health/functional status potentially requiring even more detailed supplementary guidance.

However, to date, previous reviews have only individually discussed specific cut-point thresholds for older adults (Evenson et al., 2012; Aguilar-Farias et al., 2014), epoch length (Gabriel et al., 2010; Ayabe et al., 2013), nonwear time (Choi et al., 2011; Hutto et al., 2013) and the appropriate number of valid days (Sasaki et al., 218); none have discussed, the impact of such data processing criteria decisions in combination (Hutto et al., 2013). In addition, research has shown that although the use of accelerometers within the fields of PA (and ST) and public health is increasing, the methodological processes are poorly reported and it should be noted that no best practice guidelines or consensus exists for any age group but even less so, for older adults regarding data processing (Migueles et al., 2017). Furthermore, with the use of accelerometers increasing across fields (urban planning, urban design, public health, cancer) those implementing accelerometry protocols may not be fully informed on the implications of selecting specific criteria causing them to select and implement criteria which would not be considered appropriate for their research. Therefore, research of this nature is essential in order to consider and explore the field of data processing and to highlight the potential differences that can result dependent on the criteria selected.

The specific research question for the current study is "do changes to accelerometer criteria (LFE, nonwear time and cut-point thresholds) significantly impact the resultant levels of PA and ST (minutes per day) for older adults?" By answering this research question we can contribute to the field of older adult PA and ST research by potentially highlighting the need for a consensus regarding processing decisions. The reporting of such processes and methodologies is imperative in order to align efforts and to ensure the standardisation of studies particularly for demographic sub-groups of the population. As if this does not happen and researchers fail to collaborate and provide transparent information regarding their implemented processing criteria, there is the potential for data inaccuracy, which reduces the ability for cross-study comparisons and this could potentially raise critical questions over population surveillance figures (Strath et al., 2012; Pedisic & Bauman, 2015; Migueles et al., 2017; Aadland et al., 2018).

Aims

The aim of the study was to assess the impact of different accelerometer criteria (LFE, nonwear time and cut-point thresholds) on recorded levels of PA and ST in a sample of healthy free-living older adults. The research questions for the current study aimed to determine the following: if (a) the LFE is applied will that result in significantly higher minutes of PA per day and significantly lower minutes of ST lower per day; (b) When nonwear time is set at 120 minutes, will minutes of PA per day be significantly lower and minutes of ST be significantly higher per day; and (c) When cut-point thresholds are applied that were specifically tested within a sample of older adults will minutes of PA per day be significantly higher and ST be significantly lower per day. We would like to highlight that the aim of the paper is not an exhaustive comparison of all possible criteria but rather an illustrative demonstration the impact of three differing criteria on data processing.

METHODS

For the purposes of the current study, ethical approval was sought from and approved by the Queen's University School of Medicine, Dentistry and Biomedical Sciences and School of Natural and Built Environment ethics committees. In addition, informed written consent was obtained from each participant prior to their participation in the study.

Study design

The current study analysed cross-sectional accelerometer data from a group of older adults (≥60 years) in the United Kingdom who wore ActiGraph GT3X+ devices for a 7-day period. Participants wore the device for seven consecutive days (during waking hours and non-water based activities) and were asked to complete a wear time diary. Participants wore the accelerometer on an elasticated waist belt on their right hip (common placement for this age group) (Migueles et al., 2017). Data were collected February-July 2017.

Sample recruitment

Healthy free-living older adults were recruited from Wave 1 of the Northern Ireland Cohort for the Longitudinal Study of Ageing (NICOLA) (aged 60 years and older); involving 8,500 men and women aged ≥50 years (http://nicola.qub.ac.uk/). Participants were randomly selected from across Northern Ireland and a subsample were invited to participate in the current study (Ellis et al., 2018). Briefly, 71.8% (675/940) were contactable; of those participants, 45.0% (304/675) were recruited, and 83.2% (253/304) of recruits completed the study (Figure 1). The wider study methods have been detailed elsewhere (Ellis et al., 2018).

Accelerometer cleaning and processing

Raw accelerometer activity data were processed using ActiLife 6 software (ActiGraph Inc., Pensacola, FL). Data processing criteria are summarized in Table 2.

Each of the processing criteria were inserted into the combination matrix producing a total of eight difference processing combinations (naming convention: combination-1…combination-8). The accelerometer data was exported to Microsoft Excel (.csv format) for each of the eight processing combinations detailed in Table 2. Within Microsoft Excel, mean minutes of ST/day, mean minutes of light physical activity (LPA)/day and mean minutes of moderate-to-vigorous physical activity (MVPA)/day were extracted. The data file was then transferred to SPSS data analysis (version 23; SPSS Inc, Chicago, IL) for statistical analysis.

Statistical analysis

Following data cleaning and processing, descriptive analyses were performed on the demographic variables (gender, age, ethnicity, relationship status, education and employment situation) of the sample. As the data for MVPA significantly deviated from normal distribution (tested using Shapiro Wilk tests), data were presented as median and interquartile ranges. Normally distributed ST and LPA were presented as mean and *SD* ($p > .05$).

For MVPA, Wilcoxon Sign-Rank tests were performed, and paired *t* tests for ST and LPA to determine if the difference between minutes per day were significantly different when the LFE function was switched on, when the length of nonwear time was changed and when different cut-point thresholds were implemented.

Correlation coefficients were then performed to determine the strength of the relationship between the results for each other the eight combinations. To interpret the Spearman's Rank correlations (MVPA per day) and Pearson's Rank Correlations (LPA per day and ST per day), the following benchmarks were used as reported by Landis & Koch (1977): 0–0.20 = *poor correlation*, 0.21-0.40 = *fair correlation*, 0.41-0.60 = *moderate/acceptable correlation*, 0.61-0.80 = *substantial correlation*, and 0.81-1.0 = *near perfect correlation* (Landis & Koch, 1977). Finally, to determine if any patterns existed within the data, a stacked bar chart was produced alongside the statistical analysis as aforementioned. Significance level was set at $p < .05$.

Accelerometer processing criteria to be tested

In line with the aims of the current study, it was decided by the research team that three processing criteria would be investigated and the rationale for three predetermined criteria would be clearly outlined within the study methodology. Each of the variations tested have been detailed below.

Low-frequency extension

Accelerometers have a tendency to filter out low-frequency acceleration signals as part of the band pass filter stage. By doing so noise, jitter and non-human movement would normally be removed. However, there is now the option to switch on a LFE when processing raw ActiGraph accelerometry data that aims to capture PA at lower intensities e.g. LPA (stretching, light house work, fishing etc.) and/or small steps; enabling the whole activity spectrum to be recorded with greater sensitivity (ActiGraph, 2011; ActiGraph, 2016; Feito et al., 2017). This LFE may also allow for lower frequency movements such as shuffling gait in older adult, which would previously have been removed.

For that reason, this function is thought to be applicable for older adults; however, many cut-point thresholds have been validated prior to the introduction of this software function and studies often fail to mention LFE employment (Evenson et al., 2012; Heesh et al., 2018). Fieto et al., (2017) and Wanner et al., (2013) both concluded that in free-living conditions for adults, the LFE significantly overestimated daily step counts and Wanner et al., (2013) also found it significantly overestimated LPA. However, even though several studies have been performed to determine if the LFE should be implemented when processing raw accelerometry data, none have considered this function in addition to other criterion decisions specifically for an older adult population and in relation to their levels of PA and/or ST (Cain et al., 2013; Wanner et al., 2013; Feito et al., 2017). We hypothesised that when the LFE is switched on minutes of PA per day would be significantly higher and minutes of ST would be significantly lower per day in comparison to when the function is switched off.

Nonwear time

Nonwear and ST are both represented by the absence of acceleration, and expressed as zero counts by device-based PA measurement. By definition, ST relates to the "time spent sitting during commuting, in the workplace and the domestic environment, and during leisure time. Sedentary behaviours such as Television viewing, computer use, or sitting in an automobile typically are in the energy-expenditure range of 1.0 to 1.5 metabolic equivalent of tasks (METs; multiples of the basal metabolic rate)," whereas nonwear time is when a device is not worn (Owen et al., 2010; Hutto et al., 2013). When considering nonwear time (consecutive 'zeros') for older adults in comparison to young/mid-life adults, it is important to consider the possibility that older adults may have differing lifestyles and spend more time sedentary during the waking day. This could be due to older adults accumulating more ST by partaking in relatively sedentary hobbies (reading, listening to music, knitting etc.) or they may have physical health impairments which limits their movement as opposed to noncompliance with the study (Owen et al., 2010; Hutto et al., 2013). Therefore, older adults in comparison to young/mid-life adults may require longer periods of nonwear to reduce the likelihood of misclassification of nonwear versus ST. Consequently, recommendations have been called for to improve the comparability and accuracy of data processing particularly for ST (Cain et al., 2018). A recent review suggested that 60 or 90 min were commonly implemented for older adults (Migueles et al., 2017) with Choi et al., (2011) reporting, for both adults and youth, that a longer time period of 90-minutes was the optimum window in comparison to 60 min. Taking these recommendations into account and considering the work that was already performed it was decided to build upon Choi et al., (2011) work for youth/adults and to test 90-min versus a lengthier period of 120 min for older adults. The rationale for this decision was thought to not only take into consideration the aforementioned sedentary habitual routines and past times of older adults but also with the aim of furthering previous work (Choi et al., 2011; Cain et al., 2018). We hypothesised that when nonwear time is set at 120 minutes, minutes of PA per day will be significantly lower and minutes of ST will be significantly higher per day in comparison with a nonwear time of 90 min.

3) Cut points and threshold classification

Cut points are those threshold values (in counts) that corresponded to a certain energy expenditure that are determined during the cut point calibration study and are used in order to translate and convert the acceleration signal into something which is physiologically more meaningful such as a measure of intensity, that reflects the force applied through the device and in turn into a meaningful output that can easily be understood (minutes by intensity) and compared to PA recommendations.

Previous research has highlighted the need for specific demographic subgroup cut points; having reported that adult cut points may not be appropriate for older adults whose PA behaviours may differ, and for whom, the "energy expenditure" of partaking in a range of activities would be higher than for young/mid-life adults (Ainsworth et al., 2000; Evenson et al., 2012; Corbett et al., 2017; Migueles et al., 2017). If the choice of cut points is unsuitable, this has the potential to significantly impact results and make cross-comparison studies nearly impossible (Freedson et al., 1998; Cain et al., 2013; Corbett et al., 2017). With a wealth of cut points available, many of which have not been validated in specific laboratory trials for the subgroup in question researchers are at risk of employing processing algorithms that have been incorrectly labelled for specific subpopulation groups

For the current analysis, comparisons were made across two distinct sets of cut points used in older adult research to highlight considerations for researchers: (a) those labeled for older adults but trialed in a sample of adults (Davis & Fox, 2007); and (b) those labeled for older adults and trialed in a sample of older adults (Copeland & Esliger, 2009). The first type of cut points is common in PA research as they are ones that have been "labeled" for use in older adult research but are based upon the commonly used Freedson cut-points developed with a sample of adults (males = 24.8 +/- 4.2 years and females = 22.9 +/- 3.8 years) (Freedson et al., 1998; Davis & Fox, 2007). Davis & Fox (2007), reduced their data to bands of 200 counts/min and established a moderate-PA threshold of ≥2,000 counts; this being the closest to the Freedson's counts of 1,952: the boundary between light (<3 METS) and moderate (3-6 METS; (Freedson et al., 1998; Davis & Fox, 2007). Conversely, the second set were developed with a sample of older adults (aged 69.7 +/- 3.5 years) in similar conditions to the Freedson laboratory testing; older adults simultaneously wore an accelerometer and had their oxygen consumption measured using a breathing mask, thus making it possible to link accelerometer count values to energy expenditure (Copeland & Esliger, 2009). Consequently, the threshold for moderate PA for older adults was set at ≥10,41 counts/min (Copeland & Esliger, 2009). We hypothesized that when cut point thresholds are specifically tested within a sample of older adult's minutes of PA per day will be significantly higher and ST will be significantly lower per day in comparison to those labeled for older adults.

Pre-determined processing criteria

Three predetermined criteria were implemented following guidance from previously published literature. (a) epoch length, both cut points thresholds were validated at a 60-s epoch (Freedson et al., 1998; Davis & Fox, 2007; Copeland & Esliger, 2009). (b) and (c) "valid day/week", a period of time required to wear the monitor in order to gauge typical behaviour and to determine habitual daily and/or weekly behavioural patterns (Kocherginsky et al., 2017). A minimum of five monitoring days (≥10 hours per day) was set following guidance from Sasaki et al., (2018).

RESULTS

Demographic characteristics

The majority of participants were: aged between 60 and 70 years old (50%, n=106); male (57%, n=129); white (100%, n=222); married/ living with a partner (68%, n=152); retired (83%, n=185); and had a Diploma/Certificate/Undergraduate, Postgraduate or higher degree (54%, n=119; Table 3).

Physical activity intensity levels

Moderate-to-vigorous physical activity

Results showed that median minutes of MVPA ranged from 17.0 to 61.0 min/day between the eight combinations (Table 4). The lowest median level of MVPA was 17.0 min/day (interquartile range 5.0-34.0), Combination 7 (Davis & Fox, 2007), LFE switched off and 120-min nonwear time; and the highest recorded median level of MVPA per day was 61.0 min/day (interquartile range 33.0-91.5), Combination 2 (Copeland & Esliger, 2009), LFE switched on and 90-min nonwear time (Table 4). Results also showed that the largest range of minutes' of MVPA per day for one combination was 291 min (2-293 min/day; Combination 4); and the smallest range was 131 min (0-131 min/day; Combination 5 and Combination 7).

Light physical activity

Mean minutes of LPA ranged from 189.8 to 250.6 min/day for the eight different combinations (Table 4). The lowest mean level of LPA was 189.8 min/day (*SD* 73.8), Combination 7 (Davis & Fox, 2007), LFE switched off and 120-min nonwear time; and the highest recorded mean level of LPA was 250.6 min/day (*SD* = 70.1), Combination 2 (Copeland & Esliger, 2009), LFE switched on and 90-min nonwear time (Table 4). Results also showed that the largest range of minutes of LPA per day for one combination was 436 min (22-458 min/day; Combination 6 and Combination 8), and the smallest range was 348 min (55-403 min/day; Combination 1).

Sedentary time

Mean ST was found to range from 522.9 to 633.7 min/day for the eight different combinations (Table 4). The lowest mean level of ST was 522.9 (91.8), Combination 2 (Copeland & Esliger, 2009), LFE switched on and 90-min nonwear time; and the highest recorded mean level of ST per day was 633.7 min/day (*SD* = 89.5), Combination 7 (Davis & Fox, 2007), LFE switched off and 120 minutes' nonwear time (Table 4). Results also showed that the largest range of minutes of ST per day for one combination was 575 min (224-799 minutes/day; Combination 4); and the smallest range was 458 minutes (297-755 min/day; Combination 1; 387-845 min/day Combination 5; and 335-793 min/day Combination 6).

Differing processing criteria

Low frequency extension

Wilcoxon Signed Ranks Tests (MVPA) and paired *t* tests (LPA) showed that when the LFE was switched on this resulted in significantly higher median minutes of MVPA and mean minutes of LPA per day in comparison to when the LFE was switched off ($p = .000$) (Table 5). Median differences ranged from approximately 2 to 11 min for MVPA per day and mean differences 25-30 min of LPA per day (Table 4 and 5).

Conversely, paired *t* tests showed that when the low frequency was switched on this resulted in significantly lower mean minutes of ST per day in comparison to when the LFE function was switched off (p = .000) (Table 5). Mean differences ranged from approximately 25-32 min of ST per day (Tables 4 and 5).

Nonwear time

Wilcoxon signed-ranks tests (MVPA) and paired *t* tests (LPA) showed that when 90 min was set for nonwear time significantly higher median minutes per day of MVPA and mean minutes per day of LPA were found in comparison to when 120 minutes was set for nonwear time (Tables 4 and 5). Median differences ranged from approximately 1-3 min per day for MVPA and mean differences ranged from approximately 4 to 6 minutes for LPA (Table 4).

Conversely, paired *t* tests showed that when 90 min was set as nonwear time, this resulted significantly lower mean minutes of ST per day in comparison to when 120 minutes was set for each of the four-combination comparisons (p = .000, Table 5). Mean differences ranged from approximately 17 to 22 minutes of ST per day (Table 4).

Cut points threshold classifications

Wilcoxon signed ranks tests (MVPA) and paired *t* tests (LPA) showed that when Copeland & Esliger (2009) cut point thresholds were used to determine levels of MVPA (median minutes) and LPA (mean minutes) per day results were significantly higher in comparison to results for the Davis & Fox (2007) cut point thresholds (Table 5). Median differences ranged from 31 to 40 min of MVPA per day and mean differences ranged from approximately 26 to 30 min of LPA per day (Table 4). Conversely, paired *t* tests showed that for mean minutes of ST per day results were significantly lower for Copeland & Esliger (2009) cut point thresholds (p = .000) than compared to David & Fox (2007) (Table 5). Mean differences ranged from approximately 61 to 65 minutes of ST per day (Table 4).

Correlation coefficients

When bivariate correlation coefficients were performed for each of the eight data processing combinations for MVPA per day, results showed that each of the correlations were found to be *near perfect* when compared (r = .877-.999; Table 6). The same was found for both LPA per day (r= .822-.996) and ST (r = .891-.992).

Patterns

When the data were reviewed, and minutes of MVPA, LPA and ST per day were presented visually in Figure 2; it is possible to see that cut point thresholds have the greatest impact (Combinations 1-4 vs. Combinations 5-8). When Copeland & Esliger (2009) cut point thresholds (Combinations 1-4) were implemented, Figure 2 shows a pattern of both minutes of MVPA and LPA increasing while the proportion of minutes per day of ST decreases. When David & Fox (2007) cut point thresholds are implemented the reverse pattern can be observed, minutes of MVPA and LPA per day decrease and the proportion of ST increases (Figure 2). Other emerging patterns were in relation to LFE and nonwear time and total recorded minutes per day. When the LFE was switched on, more minutes were recorded per day in comparison when the extension was switched off; when 120 min of nonwear time was set more minutes were recorded per day; and when both the LFE was switched on and nonwear time was set to 120 min, these combinations (4 and 8) resulted in the largest recorded minutes per day.

DISCUSSION

The current study compared the impact of different accelerometer criteria on recorded levels of PA intensity (and ST) in a sample of healthy free-living older adults. To our knowledge this is the first study that has taken multiple accelerometer criteria and applied it to the data collected from a sample of healthy free-living older adults to determine and highlight the impact on results. Research such as this is imperative, as to date, no specific recommendations have been made regarding accelerometry processing and older adults, and in this expanding and rapidly developing field it is important to highlight the potential differences dependent on the criteria selected.

Our research highlighted that significant differences were found for each PA intensity per day (MVPA, LPA and ST) regardless of the criteria that was changed (LFE, nonwear time and cut point thresholds). However, as expected the largest differences in minutes per day for each intensity were found when the cut point threshold classification was changed whilst the LFE and nonwear time were held constant.

Low frequency extension

The option to use a LFE is a recent development for accelerometry and device based PA measurement. LFE has the potential to capture PA at lower intensities such as LPA, steps and/or shuffling gait which may be performed by older adults; its use enables the whole activity spectrum to be recorded with greater sensitivity, and its use has been suggested when implementing PA measurement particularly in a sample of older adults (ActiGraph, 2011; ActiGraph, 2016; Feito et al., 2017). However, as this function is relatively new, many cut point thresholds have not been validated using LFE; prompting research groups to carry out work in this specific area (Cain et al., 2013; Wanner et al., 2013; Fieto et al., 2015; Fieto et al., 2017).

Results from the current study showed that when the LFE was switched on, higher levels of LPA and MVPA and lower levels of ST per day were recorded. These findings were in agreement with previous research (Cain et al., 2013; Wanner et al., 2013; Fieto et al., 2015; Fieto et al., 2017). Cain and colleagues (2013) also compared data from the GT3X with/out LFE with an older generation 7,164 accelerometer and results showed that by using the LFE this made the recorded data (and consequently the results) more comparable with previous studies - in particular, studies that reported the development of cut points thresholds (Cain et al., 2013).

Therefore, despite the fact that the results from the current study were in agreement with previous research, the findings make it difficult to make a specific recommendation regarding the use of the LFE for older adults. We have however, demonstrated the impact that the LFE can have on levels of PA at differing intensities per day for older adults and have supplemented previous research showing that not only do differences exist, but also results indicate, and potentially favour, the use of the LFE for older adults particularly when results will be compared or linked with studies that implemented older accelerometers or when studies implement cut point thresholds that were validated with older accelerometer models (Cain et al., 2013; Fieto et al., 2017). Going forward, more research is required in older adults to confirm these assumptions as the current study cannot state specific criteria recommendations as to date, no gold standard measure exists to compare differing results to. Further research is also required to validate PA with and without the use of LFE and when performing a range of physical activities and when sedentary.

Nonwear time

Considering the current findings alongside previously published research in similar population groups, longer periods of nonwear time would be preferred. Results from the current study showed that near perfect correlations were found for ST per day and differences in mean minutes were minimal (approximately 20 minutes) when 90 min was set in comparison with 120 min. This was similar to Hutto and colleagues (2013) who also demonstrated that there were differences between actual minutes recorded, but their findings support the use of longer periods of nonwear time for older adults. Migueles et al., (2017) also reported that a longer period of 90-min nonwear time was preferred over a shorter 60-min period in order to identify actual wear time in older adults (Migueles et al., 2017). By implementing a protocol allowing for a period of 120 min of nonwear time the likelihood of misclassification and data inaccuracies is reduced (Hutto et al., 2013). The longer periods of time will also have the potential to take account of general ST and the pastimes of older adults that would be considered sedentary but are actually active for instance knitting, reading, painting etc. (Hutto et al., 2013; Choi et al., 2011).

Cut points threshold classifications

This study highlights the differences when using differing cut point thresholds for older adults: (a) established in a sample of younger adults and labeled for older adults (Davis & Fox, 2007); and (b) established in a sample of older adults and labeled for older adults (Copeland & Esliger, 2009). We recognise that this finding was predictable however important to highlight. With the field of accelerometry ever expanding with multi-disciplinary groups implementing accelerometry protocols the possibility of inappropriate criteria being implemented is likely. It is therefore important to acknowledge, that commonly used or labelled criteria may not always be the most appropriate and may cause significantly different results for both PA and ST, resulting in incomparable and inaccurate data which will not reflect a participant's true reality.

Therefore, when selecting cut-points it would be recommended to trace the origin of the cut point back to the original developmental study and to determine what subgroup it was tested in. For their cut points Copeland & Esliger (2009) performed a laboratory-based assessment in a healthy sample of older adults (69.7 +/- 3.5 years) whereas Davis & Fox (2007) based their cut points on the commonly implemented thresholds established by Freedson et al., (1998) who performed laboratory testing in a sample of young adults (males = 24.8 +/- 4.2 years and females = 22.9 +/- 3.8 years).

Copeland & Esliger (2009) showed a strong relationship (*r*= 0.878) with walking speed and accelerometer counts, and accelerometer counts, and oxygen consumption (*r*= 0.60) and following the laboratory-based assessment, cut point thresholds were specifically established for older adults and took account of the differences older adults experience regarding their changing levels of fitness with age (Ainsworth et al., 2000; Copeland & Esliger, 2009). Reports stated that ≥1,041 counts per minute would classify MVPA, which corresponded to a mean $V0_2$ of 13 $ml \cdot kg^{-1} \cdot min^{-1}$ equivalent to 3.7 METs (Copeland & Esliger, 2009). This finding is in line with the Compendium of Physical Activities: when classifying intensity of PA by METs, 3-6 would be considered as moderate-PA (Ainsworth et al., 2000). Copeland & Elisger (2009, page 25) also reported that by using this cut-point for MVPA "there is little chance that a light minute of activity will be inappropriately labeled as MVPA" and the cut-point of ≥1,041 counts per minute would actually be considered a "conservative delineation of MVPA for older adults". Consequently, the use of this cut-point in older adult studies would be recommended.

Recommendations for future research

In line with previous studies that involved samples of older adults we chose to process the accelerometry data in 60-s epochs and extract the acceleration signal from the vertical axis. However, as both criterion have various options that can be selected by researchers during the measurement and analysis period, both criterion need to be considered within the field of device based PA measurement and further research is warranted.

The magnitude of accelerations measured by an accelerometer are recorded and processed at specific time intervals (15-, 30- or 60-s epochs) (Gabriel et al., 2010; Ayabe et al., 2013). Sixty seconds is a standard value used in first generation accelerometers and was selected as a default due to memory and battery capacity constraints rather than choice (Gabriel et al., 2010). Previous research has highlighted the potential for data inaccuracies when measuring at 60 s, whereas the use of shorter epochs may prevent misclassification of activities and be chosen to reflect the understanding of how PA is accumulated in a particular population group (Gabriel et al., 2010). Therefore, if we accumulate/average out acceleration signals over a longer time frame (60 s), we have the potential to dilute the intensity of PA that would have been otherwise reflected in shorter epoch (10, 15 or 30-s epochs). Consequently, there is a requirement for research in line with these technological advancements not only to determine which epoch length is appropriate, but also which epoch should be used with specific sub-groups; for instance, older adults and/or those with disabilities (Migueles et al., 2017).

Axis are also of interest and should be further explored. Accelerometers have the capability of measuring acceleration on 1, 2, or 3 axes; with three axis becoming more common with technological advances. Within the current study only the vertical axis was used due to the methodologies implemented within the original study protocols (Davis & Fox, 2007; Copeland & Esliger, 2009). However, as technology has now advanced to enable not only measurements taken in a single axis but across three planes (triaxial) it is important that research further develops guidelines of axial use across all age groups and subgroups of the population in order to ensure that the accuracy of PA measurement is improved and increased (Howe et al., 2009).

Within this complex field, more research is also required in order to determine the degree of heterogeneity within the subpopulation of older adults, and to design and implement further calibration studies for older adults both laboratory based and in free-living settings. The majority of calibration tests have been performed in laboratory settings as discussed within the introduction and discussion although future research would benefit from understanding free-living PA (Granat, 2012). In addition, in order to expand this field of research more work is required to specifically review LPA in order to expand the full spectrum of activity intensities (ST, light, moderate and vigorous) for older adults. This is in line with current research in the field of LPA and older adults that has shown LPA to be positively associated with well-being, physical health and life satisfaction (those aged 60 years and older); and independent of other PA and ST, LPA to reduce the risk of depressive symptoms for older adults (Bae et al., 2018; Ku et al., 2018). Furthermore, as PA research is moving towards previously mentioned non-proprietary methods more research is required into the processing of raw accelerometry data. However, as software such as R and GGIR cannot be used by all researchers and practitioners as specific knowledge and expertise are required it is important that studies such as this, are performed in order to produce guidance on what is currently available and can be used by the majority of those using accelerometers (i.e. count based measures such as those used by Actigraph).

Strengths and Limitations

A strength of the current study was the implementation of the ActiGraph GT3X+ accelerometer in a large sample of older adults to collect PA and ST data for a period of seven consecutive days. Limitations include the fact PA and ST was only measured in healthy free-living older adults and efforts were not made to review our sample by differences in age or health/functional status as this is an additional complexity that needs to be considered within this age group.

CONCLUSIONS

As highlighted within the current study, the choice of, and the combination of accelerometry processing criteria has the potential to significantly impact the results of a study that aims to objectively measure PA and/or ST. If suitable criteria are not chosen for the targeted population group, this can lead to data inaccuracies and may prevent cross-study and/or cross-country comparisons. It is imperative that research groups present their methodologies in a transparent manner and collaborate with other researchers to ensure standardisation of methods. More research is required in this area before definitive recommendations can be made, specifically studies in both laboratory and free-living settings to determine which criteria are the most accurate and true to an older adult's reality. Until that time, it is important researchers do not implement accelerometry research in older adults with a "one size fits all" data processing approach and prior to the comparison of data across studies, methodological processes should be fully examined (Gorman et al., 2014; Strath et al., 2012).

Abbreviations

IQR: Inter-Quartile Range; LED: Light-Emitting diode; LFE: Low Frequency Extension; LPA; Light Physical Activity; METs: Metabolic Equivalent of Task; MVPA: Moderate-to-

Vigorous-Physical-Activity; n: number; NICOLA: Northern Ireland Cohort for the Longitudinal Study of Ageing; PA: Physical Activity; SD: Standard Deviation; ST: Sedentary Time; UK: United Kingdom.

Conflicts of Interest

The authors declare that they have no competing or conflicting interests including financial interests.

Acknowledgements

Firstly, we would like to acknowledge Professor Ian Young as the founding Principal Investigator of the NICOLA Study from which the current study was sub-sampled. In addition, we are grateful to all the participants of the NICOLA Study, and the whole NICOLA team, which includes nursing staff, research scientists, clerical staff, computer and laboratory technicians, managers and receptionists. The Atlantic Philanthropies, the Economic and Social Research Council, the UKCRC Centre of Excellence for Public Health Northern Ireland, the Centre for Ageing Research and Development in Ireland, the Office of the First Minister and Deputy First Minister, the Health and Social Care Research and Development Division of the Public Health Agency, the Wellcome Trust/Wolfson Foundation and Queen’s University Belfast provide core financial support for NICOLA. The authors alone are responsible for the interpretation of the data and any views or opinions presented are solely those of the authors and do not necessarily represent those of the NICOLA Study team.

REFERENCES

Aadland E, Andersen LB, Anderssen SA, Resaland GK. A comparison of 10 accelerometer nonwear time criteria and logbooks in children. BMC Public Health. 2018;18:323.

ActiGraph. Actigraph, Low Frequency Extension Explained. 2016. Available at: https://actigraph.desk.com/customer/en/portal/articles/2515505-low-frequency-extension-explained.

ActiGraph R&D and Software Departments: ActiLife 5 - User’s Manual. Volume H. Pensacola: ActiGraph; 2011. p. 16.

Aguilar-Farias N, Brown WJ, Peeters GM. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. Journal of science and medicine in sport / Sports Medicine Australia. 2014;17(3):293–9.

Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc. 2000;32(9):498–504.

Ayabe M, Kumahara H, Morimura K, Tanaka H. Epoch length and the physical activity bout analysis: an accelerometry research issue. BMC Res Notes. 2013;6:20.

Bae W, Suh IKY, Ryu J, Heo J. Physical activity levels and well-being in older adults. Psychological Reports. 2018;1–14.

Barnett A, Van Den Hoek D, Barnett D, Cerin E. Measuring moderate-intensity walking in older adults using the ActiGraph accelerometer. BMC Geriatr. 2016;16(1):211.

Cain KL, Bonilla E, Conway TL, Schipperijn J. Defining accelerometer nonwear time to maximize detection of sedentary time in youth. Pediatr Exerc Sci. 2018;30;2;288-95.

Cain KL, Conway TL, Adams MA, Husak LE, Sallis JF. Comparison of older and newer generations of ActiGraph accelerometers with the normal filter and the low frequency extension. Int J Behav Nutr Phys Act. 2013;10(1):51–7.

Cain KL, Sallis JF, Conway TL, et al. Using accelerometers in youth physical activity studies: a review of methods. J Phys Act Health. 2013;10:437–450.

Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc. 2011;43(2):357–364.

Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. J Aging Phys Act. 2009;17:17–30.

Corbett DB, Wanigatunga AA, Valiani V, Hand-berg EM, Buford TW, Brumback B, Casanova R, Janelle CM, Manini TM. Metabolic costs of daily activity in older adults (chores xl) study: Design and methods. Contemporary Clinical Trials Communications. 2017:6;1-8.

Crouter SE, Churilla JR, Bassett DR. Estimating energy expenditure using accelerometers. Eur. J. Appl. Physiol 2006;98,:601–612.

Davis MG, Fox KR. Physical activity patterns assessed by accelerometry in older people. European Journal of Applied Physiology. 2007.

Ellis G, Hunter R, Hino, AAF, Cleland CL, Ferguson S, Murtagh B, Anez RRR, Melo S, Tully M, Kee F, Sengupta S, Reis R. Study protocol: healthy urban living and ageing in place (HULAP): an international, mixed methods study examining the associations between physical activity, built and social environments for older adults the UK and Brazil. BMC Public Health. 2018;18:1135.

Evenson KR, Buchner DM, Morland KB. Objective measurement of physical activity and sedentary behavior among US adults aged 60 years or older. Prev Chronic Dis, 2012;9:E26.

Feito Y, Garner HR, Bassett DR. Evaluation of ActiGraph's Low-Frequency Filter in Laboratory and Free-Living Environments. Med Sci Sports Exerc. 2015;47(1):211–7.

Feito Y, Hornbuckle LM, Reid LA, Crouter SE. Effect of actigraph's low frequency extension for estimating steps and physical activity intensity. PLoS ONE. 2017;12.

Freedson PSE, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. Accelerometer. Med. Sci. Sports Exerc. 1998;30:777–781.

Gabriel KP, McClain JJ, Schmid KK, Storti KL, High RR, Underwood DA, Kuller LH, Kriska AM: Issues in accelerometer methodology: the role of epoch length on estimates of physical activity and relationships with health outcomes in overweight, post-menopausal women. Int J Behav Nutr Phys Act. 2010;7:53-53.

Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. Eur Rev Aging Phys Act. 2014;11:35-49.

Granat MH. Event-based analysis of free-living behaviour. Physiological Measurement. 2012;33(11);1785-800.

Heesh KC, Hill RL, Aguilar-Farias A, van Uffelen JGZ, Pavey T. Validity of objective methods for measuring sedentary behaviour in older adults: a systematic review. International Journal of Behavioural Nutrition and Physical Activity. 2018;15:119.

Howe CA, Staudenmayer JW, Freedson PS. Accelerometer prediction of energy expenditure: vector magnitude versus vertical axis. Med Sci Sports Exerc. 2009;41;12:2199–206.

Hutto B, Howard VJ, Blair SN, Colabianchi N, Vena JE, Rhodes D, Hooker SP. Identifying accelerometer nonwear and wear time in older adults. Int J Behav Nutr Phys Activ. 2013;10: 120-10.

Kocherginsky M, Huisingh-Scheetz M, Dale W, Lauderdale DS, Waite L. Measuring Physical Activity with Hip Accelerometry among U.S. Older Adults: How Many Days Are Enough? PloS one. 2017.

Ku PW, Steptoe A, Liao Y, Sun WJ, Chen LJ. Prospective relationship between objectively measured light physical activity and depressive symptoms in later life. International Journal of Geriatric Psychiatry. 2018.

Kwan M, Woo J, Kwok T. 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. 2004;55:179–182.

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159-174.

McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology. 2016;17:567–580.

Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nystrom C, Mora-Gonzalez J, Lof M, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. Sports Med. 2017.

Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. Exerc Sport Sci Rev. 2010;38(3):105–13.

Pedisic Z, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. Br J Sports Med. 2015;49(4):219–25.

Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Res Q Exerc Sport 2000;71:S1-14.

Sasaki JE, Junior JH, Tribess S, Marocolo Junior M, Stabelini Neto A, Virtuoso Junior JS. Number of days required for reliably estimating physical activity and sedentary behaviour from accelerometer data in older adults. J Sports Sci. 2018;36;14;1572-1577.

Seymour DG, Ball AE, Russell EM, Primrose WR, Garratt AM, Crawford JR. Problems in using health survey questionnaires in older patients with physical disabilities. The reliability and validity of the SF-36 and the effect of cognitive impairment. J Eval Clin Practice. 2001;7(4): 411–418.

Strath SJ, Greenwald MJ, Isaacs R, Hart TL, Lenz EK, Dondzila CJ, Swartz AM. Measured and perceived environmental characteristics are related to accelerometer defined physical activity in older adults. Int J Behav Nutr Phys Act. 2012;9(40).

Toftager M, Kristensen PL, Oliver M, Duncan S, Christiansen LB, Boyle E, Brond JC, Troelsen J: Accelerometer data reduction in adolescents: effects on sample retention and bias. Int J Behav Nutr Phys Act. 2013;10:40.

Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M: Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008;40:181-188.
Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc. 2005;37;531-43.

UNFPA. United Nations Population Fund. Ageing in the Twenty-First Century: A Celebration and A Challenge. 2012. Available at: https://www.unfpa.org/sites/default/files/pub-pdf/Ageing%20report.pdf

Wanner M, Martin BW, Meier F, Probst-Hensch N, Kriemler S. Effects of filter choice in GT3X accelerometer assessments of free-living activity. Med Sci Sports Exerc**.** 2013;45:170–177.

Ward DS, Evenson KR, Vaughn A, Rodgers AB, Troiano RP. Accelerometer use in physical activity: best practices and research recommendations. Med Sci Sports Exerc. 2005;37(11):582–88.

Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, Vanhees L. Assessment of physical activity - a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil. 2010;17:127-139.

Figure 1. Flow Diagram of Participant Recruitment

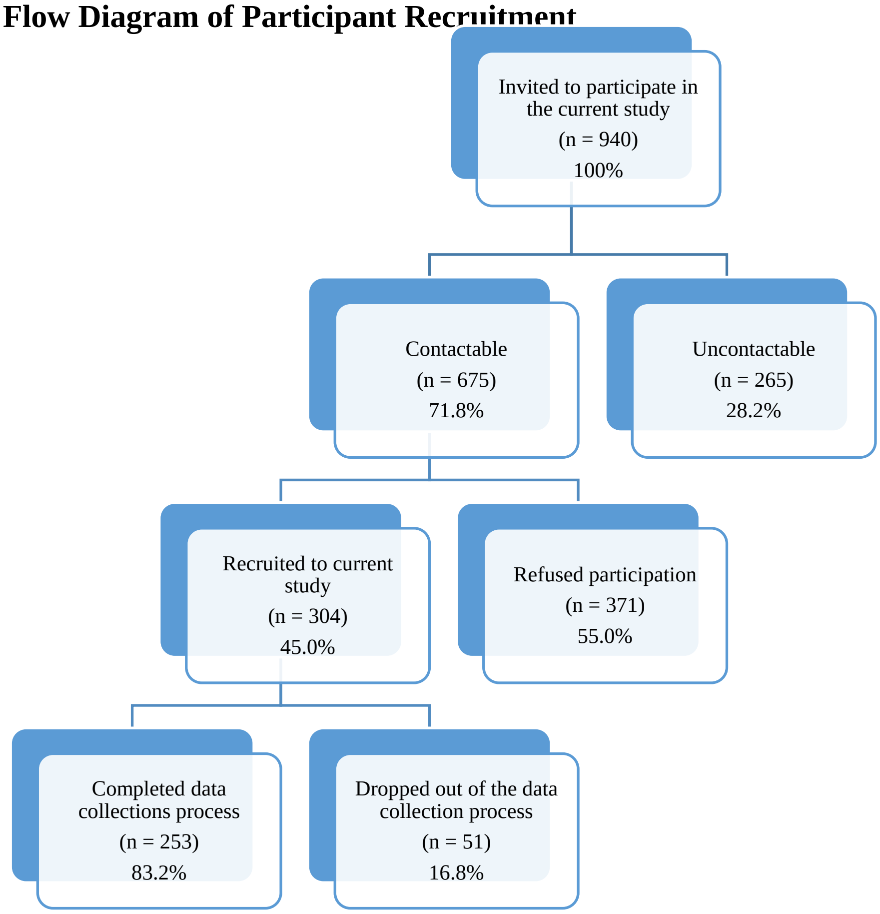

Table 1. Criteria considerations during the data processing stage of an accelerometry study.

Criteria decisions	Definition	Reference
Epoch length	"Accelerometers function by integrating a filtered digitized acceleration signal over a user-specified time interval, commonly referred to as an epoch".	Trost et al., 2005
Low frequency extension function	"The low frequency extension filter, which increased the device's sensitivity to lower intensity activities; thereby, allowing for the measurement of a greater range of physical activity intensities".	Actigraph, 2016
Non-wear time	"Non-wear time is the time during a measurement period where participants do not wear the accelerometer, and should be excluded from further analyses on the assumption that the remaining wear time is sufficiently representative for the whole measurement period".	Aadland et al., 2018
Number of hours in a valid day/Number of days in a valid week	"To monitor activity for a sufficient number of days so that the resulting daily average reflects an individual's usual or habitual level of physical activity".	Trost et al., 2005
Intensity cut-points	"Accelerometer data can be quantified as counts-per-minute with established count cut points and ranges categorizing light, moderate or vigorous PA intensity…Derivation of cut points involves establishing relationships between energy expenditure and accelerometer counts".	Barnett et al., 2016

Table 2. Criteria for accelerometer processing for older adults.

Reference	Epoch (s)	Vector	Combination number	LFE	Nonwear time (min)	Sedentary time	LPA	MVPA	Vigorous physical activity
Copeland & Esliger (2009)*	60	Vertical axis	1	Off	90 minutes	≤99	100-1040	≤1040	NA
			2	On	90 minutes				
			3	Off	120 minutes				
			4	On	120 minutes				
Davis & Fox (2007)*	60	Vertical axis	5	Off	90 minutes	≤199	200-1999	2000-3999	≥4000
			6	On	90 minutes				
			7	Off	120 minutes				
			8	On	120 minutes				

Note. LFE = low-frequency extension; LPA = light physical activity; MVPA = moderate-to-vigorous physical activity; NA = not applicable.

[a]Number of hours in a valid day: 10 hr; and number of valid days in a valid week: 5 days (including one weekend day).

Table 3. Minutes per day of moderate-to-vigorous physical activity and sedentary time by processing variation.

		Overall sample n (%)
Gender	Male	129 (57)
	Female	97 (43)
Age (years)	60-70	106 (50)
	71-80	85 (40)
	81-90	18 (8)
	91 plus	4 (2)
Ethnicity	White	222 (100)
Relationship status	Married or living with a partner	152 (68)
	Single	20 (9)
	Separated	4 (2)
	Divorced	14 (6)
	Widowed	34 (15)
Highest educational attainment	None/Primary school (not complete)/ Primary or equivalent	24 (11)
	High School (GCSE/O-Level/Intermediate/Junior Cert)/High School (A-Level/Leaving Cert)	76 (35)
	Diploma/Certificate/Undergraduate primary degree/Postgraduate/higher degree	119 (54)
Current situation	Retired	185 (83)
	Employed or self-employed	24 (11)
	Permanently disabled or sick, Looking after home or family or other	14 (6)

Table 4. Minutes per day of MVPA and sedentary time by processing variation.

Reference	**Combination number**	**LFE**	**Non-wear time (min)**	**MVPA (median min/day [Interquartile range])**	**LPA (mean min/day [*SD*])**	**Sedentary time (mean min/day [*SD*])**
Copeland & Esliger (2009)	1	Off	90	51.0 (27.0-79.0)	225.4 (66.0)	550.5 (88.4)
	2	On	90	61.0 (33.0-91.5)	250.6 (70.1)	522.9 (91.8)
	3	Off	120	48.0 (25.0-77.0)	219.4 (69.1)	572.6 (93.8)
	4	On	120	59.0 (31.3-91.0)	246.6 (72.0)	540.2 (97.5)
Davis & Fox (2007)	5	Off	90	19.0 (6.0-38.0)	195.9 (71.4)	612.6 (85.6)
	6	On	90	21.0 (7.0-40.0)	224.4 (77.6)	587.8 (89.7)
	7	Off	120	17.0 (5.0-34.0)	189.8 (73.8)	633.7 (89.5)
	8	On	120	20.0 (6.0-38.0)	220.1 (78.7)	605.2 (93.8)

Table 5. Statistical analysis to compare variations in data processing.

Reference	Combination number	LFE	Non-wear time (min)	Moderate-to-vigorous physical activity (min/ day)	Light physical activity (min/ day)	Sedentary time (min/ day)
				Z (% difference)	**Mean difference (% difference)**	**Mean difference (% difference)**
Copeland & Esliger (2009)	1 vs 2	Off vs. On	90	-12.779* (17.86)	-28.128* (10.59)	29.162* (5.14)
	3 vs 4	Off vs. On	120	-13.173* (20.56)	-27.000* (11.67)	32.859* (5.82)
Davis & Fox (2007)	5 vs 6	Off vs. On	90	-12.053* (10.00)	-32.045* (13.56)	26.351* (4.13)
	7 vs 8	Off vs. On	120	-12.360* (16.22)	-31.137* (14.78)	28.889* (4.60)
Copeland & Esliger (2009)	1 vs 3	Off	90 vs. 120	-3.136** (6.06)	1.184** (2.70)	-19.103* (3.94)
	2 vs 4	On	90 vs. 120	-2.983** (3.33)	1.100* (1.61)	-15.402* (3.25)
Davis & Fox (2007)	5 vs 7	Off	90 vs. 120	-2.502*** (11.11)	1.170* (3.16)	-18.906* (3.39)
	6 vs 8	On	90 vs. 120	-2.632*** (4.88)	1.520** (1.93)	-16.105* (2.92)
Copeland & Esliger (2009) vs Davis & Fox (2007)	1 vs 5	Off	90	-12.919* (91.43)	29.811* (14.00)	-62.527* (10.68)
	2 vs 6	On	90	-13.073* (97.56)	26.175* (11.03)	-64.913* (11.69)
	3 vs 7	Off	120	-13.291* (95.38)	29.549* (14.47)	-61.132* (10.13)
	4 vs 8	On	120	-13.319* (98.73)	26.534* (11.36)	-64.949* (11.35)

*p.000; **p <.005; p <.05

Table 6. Bivariate correlation coefficients for each of the eight data processing combinations for moderate-to-vigorous physical activity.

			Combination							
			1	2	3	4	5	6	7	8
Combination	**1**	**Correlation Coefficient**		.994*	.998*	.992*	.911*	.923*	.909*	.927*
	2	**Correlation Coefficient**	.994*		.994*	.999*	.877*	.896*	.879*	.901*
	3	**Correlation Coefficient**	.998*	.994*		.995*	.911*	.926*	.915*	.932*
	4	**Correlation Coefficient**	.992*	.999*	.995*		.877*	.896*	.882*	.904*
	5	**Correlation Coefficient**	.911*	.877*	.911*	.877*		.992*	.999*	.996*
	6	**Correlation Coefficient**	.923*	.896*	.926*	.896*	.992*		.992*	.994*
	7	**Correlation Coefficient**	.909*	.879*	.915*	.882*	.999*	.992*		.997*
	8	**Correlation Coefficient**	.927*	.901*	.932*	.904*	.996*	.994*	.997*	

*p .000

Figure 2. Daily proportion of moderate-to-vigorous physical activity, light physical activity and sedentary time for each data processing combination.