This is the published version of a paper published in Scandinavian Journal of Medicine and Science in Sports.

Citation for the original published paper (version of record):

https://doi.org/10.1111/sms.14541

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:gih:diva-7967
ORIGINAL ARTICLE

Fundament for a methodological standard to process hip accelerometer data to a measure of physical activity intensity in middle-aged individuals

Daniel Arvidsson1 | J. Fridolfsson1 | E. Ekblom-Bak2 | Ö. Ekblom2 | G. Bergström3,4 | M. Börjesson5,6

1Center for Health and Performance, Department of Food and Nutrition, and Sport Science, Faculty of Education, University of Gothenburg, Gothenburg, Sweden
2Department of Physical Activity and Health, Swedish School of Sport and Health Sciences, Stockholm, Sweden
3Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
4Center for Health and Performance, Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
5Sahlgrenska University Hospital, Region Västra Götaland, Gothenburg, Sweden
6Sahlgrenska University Hospital, Region Västra Götaland, Gothenburg, Sweden

Abstract

Background: There is a lack of a methodological standard to process accelerometer data to measures of physical activity, which impairs data quality and comparability. This study investigated the effect of different combinations of settings of multiple processing components, on the measure of physical activity and the association with measures of cardiometabolic health in an unselected population of middle-aged individuals.

Methods: Free-living hip accelerometer data, aerobic fitness, body mass index, HDL:total cholesterol ratio, blood glucose, and systolic blood pressure were achieved from 4391 participants 50–64 years old included in The Swedish CArdioPulmonary bioImage Study (SCAPIS) baseline measurement (cross-sectional). Lab data were also included for calibration of accelerometers to provide comparable measure of physical activity intensity and time spent in different intensity categories, as well as to enhance understanding. The accelerometer data processing components were hardware recalibration, frequency filtering, number of accelerometer axes, epoch length, wear time criterium, time composition (min/24 h vs. % of wear time). Partial least regression and ordinary least regression were used for the association analyses.

Results: The setting of frequency filter had the strongest effect on the physical activity intensity measure and time distribution in different intensity categories followed by epoch length and number of accelerometer axes. Wear time criterium and recalibration of accelerometer data were less important. The setting of frequency filter and epoch length also showed consistent important effect on the associations with the different measures of cardiometabolic health, while the effect of recalibration, number of accelerometer axes, wear time criterium and expression of time composition was less consistent and less important. There was a large range in explained variance of the measures of cardiometabolic health depending...
1 | INTRODUCTION

The use of wearable devices to assess physical activity (PA) has expanded rapidly to replace self-report methods for more detailed and accurate data. Accelerometers are common wearable devices, worn at the hip, wrist or thigh, recording acceleration along three perpendicular axes.\(^1\)\(^2\) The raw data are commonly processed in multiple steps to a continuous metrics representing PA intensity (e.g., counts, milligravity (mg)). This metrics can be calibrated to estimate energy expenditure and time spent in PA intensity categories, for example, sedentary (SED), light PA (LPA), moderate PA (MPA), vigorous PA (VPA), very vigorous PA (VVPA).

However, data collected device-based does not automatically improve accuracy, as multiple processing steps require multiple decisions, introducing measurement errors.\(^2\) For example, previous research shows that the number of accelerometer axes,\(^3\)–\(^7\) frequency filtering,\(^8\)\(^9\) epoch length,\(^5\)\(^10\)–\(^13\) and wear time\(^14\)\(^15\) are processing steps (or components) where different settings (e.g., narrow filter vs. wider filter, short vs. long epoch length) affect the PA measure. Unfortunately, there is no uniform standard to process accelerometer data, which hampers comparability.\(^1\)\(^2\)

A uniform standard needs to be developed based on the simultaneous evaluation of the effect of multiple processing components and settings on the PA metrics, and ideally also toward a reference measure to determine criterion validity. In most cases, there is no suitable free-living reference measure. Alternatively, the importance of the processing components and settings would be determined from prediction of health, that is, predictive validity. Still, their importance depends on the specific measure of health, reflecting certain aspects of PA, for example, volume (or energy expenditure) or intensity. Therefore, predictive validity should be determined with different measures of health. Previous research typically evaluated a single processing component, with or without association to measures of health.

The aim of this study was to investigate the simultaneous effect of multiple accelerometer data processing components and settings on PA and the association with measures of cardiometabolic health in a large sample of middle-aged individuals. The goal is to provide a fundament for a uniform standard for such a population.

2 | METHODS

2.1 | Study overview

Free-living hip accelerometer data and measures of cardiometabolic health from the Swedish CArdioPulmonary bioImage Study (SCAPIS) baseline measurement (cross-sectional) together with lab data from the Measuring Energy expenditure and Diary intake at different Activity Levels (MEDAL) study were used to evaluate the effect on PA (Section 1) and associations with measures of cardiometabolic health (Section 2). Hip accelerometer data were processed to a continuous measure of PA intensity, which is commonly used in research.\(^1\) The hip placement is close to the center of mass and provides the best representation of the full body acceleration.\(^16\) Further, simple linear algorithms predict energy expenditure from hip acceleration data with the same accuracy as more complex algorithms.\(^17\) This study took its starting point from the ActiGraph accelerometer and the processing steps affecting its metrics counts, as they are most commonly used in research.\(^1\)

2.2 | Study samples

The SCAPIS is a multicenter cohort including 30 154 randomly selected men and women aged 50–64 years from six regions in Sweden.\(^16\) Baseline recruitment and data collection was performed in 2013–2018. Baseline data from 6265 individuals in the Gothenburg region were used in the present study. SCAPIS was approved by the Ethical
review board in Umeå (2021-228-31M) and the present study by the Regional ethical board in Gothenburg (638-16). All participants gave written informed consent. Complete data were achieved from 4391 participants. The Gothenburg region was compared to the other regions in SCAPIS. Although statistical significance was reached in many of the comparisons, the differences were minor in most of the cases. The MEDAL study aims for methodological improvements of accelerometry. Data were collected from 48 adults 18–40 years in a controlled lab setting and under free-living conditions in 2021–2022. Participants were recruited through announcements at the University of Gothenburg and at local sport clubs. Only lab data were used in the present study. The MEDAL study was approved by the Swedish Ethical Review Authority (2019-05316) and the participants gave written informed consent.

### 2.3 Accelerometer data collected in SCAPIS

SCAPIS participants were instructed to wear the ActiGraph GT3X+, wGT3X+, or wGT3X-BT (3%, 15%, 82%) accelerometer (ActiGraph) in an elastic belt over the right hip during seven consecutive days and to take it off during sleep and during water activities. The accelerometers were initiated to collect data at a sampling rate of 30 Hz (the preset sample rate) with an acceleration amplitude of ±6 g.

### TABLE 1 Characteristics of the Gothenburg region compared to the other five regions in SCAPIS.

<table>
<thead>
<tr>
<th>SCAPIS total N = 26264 variable</th>
<th>Study sample N = 5623</th>
<th>Region N = 3690</th>
<th>Region N = 4887</th>
<th>Region N = 5000</th>
<th>Region N = 4742</th>
<th>Region N = 2322</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean (SD)</td>
<td>57.5 (4.3)</td>
<td>57.4 (4.3)</td>
<td>57.5 (4.4)</td>
<td>57.7 (4.4)*</td>
<td>57.4 (4.4)*</td>
<td>57.5 (4.3)</td>
</tr>
<tr>
<td>Sex, % males</td>
<td>52.3</td>
<td>50.0*</td>
<td>50.0*</td>
<td>51.3</td>
<td>51.1</td>
<td>51.1</td>
</tr>
<tr>
<td>BMI, kg·m⁻², mean (SD)</td>
<td>26.7 (4.3)</td>
<td>26.7 (4.3)</td>
<td>26.9 (4.4)</td>
<td>27.1 (4.6)***</td>
<td>27.1 (4.6)***</td>
<td>27.1 (4.6)***</td>
</tr>
<tr>
<td>HDL:Chol ratio, mean (SD)</td>
<td>0.31 (0.10)</td>
<td>0.32 (0.10)**</td>
<td>0.31 (0.09)</td>
<td>0.26 (0.07)***</td>
<td>0.30 (0.09)**</td>
<td>0.30 (0.09)**</td>
</tr>
<tr>
<td>Glucose, mmol·L⁻¹, mean (SD)</td>
<td>5.7 (1.0)</td>
<td>5.7 (1.0)***</td>
<td>5.8 (1.2)</td>
<td>5.9 (1.1)***</td>
<td>5.5 (1.0)***</td>
<td>5.5 (1.0)***</td>
</tr>
<tr>
<td>SBP, mmHg, mean (SD)</td>
<td>122 (17)</td>
<td>127 (17)***</td>
<td>133 (17)***</td>
<td>125 (16)***</td>
<td>126 (16)***</td>
<td>126 (16)***</td>
</tr>
<tr>
<td>Wear time, min·day⁻¹, mean (SD)</td>
<td>872 (101)</td>
<td>877 (98)</td>
<td>846 (74)</td>
<td>984 (164)</td>
<td>893 (97)***</td>
<td>893 (97)***</td>
</tr>
<tr>
<td>Accepted days, days, mean (SD)</td>
<td>6.5 (1.1)</td>
<td>6.3 (1.2)***</td>
<td>6.3 (1.0)***</td>
<td>6.6 (0.9)***</td>
<td>6.6 (0.9)***</td>
<td>6.6 (0.9)***</td>
</tr>
<tr>
<td>SED, min·day⁻¹, mean (SD)</td>
<td>469 (111)</td>
<td>465 (113)</td>
<td>456 (96)***</td>
<td>568 (158)***</td>
<td>476 (108)</td>
<td>476 (108)</td>
</tr>
<tr>
<td>LPA, min·day⁻¹, mean (SD)</td>
<td>348 (88)</td>
<td>349 (88)</td>
<td>336 (82)***</td>
<td>361 (87)***</td>
<td>357 (85)***</td>
<td>357 (85)***</td>
</tr>
<tr>
<td>MPA, min·day⁻¹, mean (SD)</td>
<td>50 (26)</td>
<td>56 (29)***</td>
<td>49 (26)</td>
<td>50 (26)</td>
<td>54 (28)***</td>
<td>54 (28)***</td>
</tr>
<tr>
<td>VPA + VVPA, min·day⁻¹, mean (SD)</td>
<td>4 (9)</td>
<td>5 (9)</td>
<td>4 (8)</td>
<td>4 (8)</td>
<td>4 (9)</td>
<td>4 (9)</td>
</tr>
</tbody>
</table>

Note: Processing component settings: No recalibration, triaxial, 0.29–1.63 Hz filter, 60 s epochs, 10 h wear time criterion, 24 h time composition (original settings in SCAPIS). Participants with complete data of all the variables compared were included (aerobic fitness was not compared), of the original total sample of N = 30 154. Original crude intensity categories used in SCAPIS (Sasaki et al 2011). ANOVA with Bonferroni post hoc test for continuous variables and chi-square for categorical variable. * p < 0.05, ** p < 0.01, *** p < 0.001 (comparison to study sample presented).

Abbreviations: Chol, total cholesterol; HDL, high-density lipoprotein; LPA, light physical activity; MPA, moderate physical activity; SBP, systolic blood pressure; VPA, vigorous physical activity; VVPA, very vigorous physical activity.

### 2.4 Accelerometer and oxygen uptake data collected in MEDAL

The participants arrived to the lab in the morning in a fasting state not having performed any strenuous PA. The protocol consisted of 20 min resting in supine position, and 4 min in each of the activities sit, stand, stand with manual work, walk 4 and 6 km/h, and run 9, 12 and 15 km/h or until voluntary exhaustion. The participants were resting in the supine position 5 min before start of measurement while the equipment was prepared. Acceleration data were collected at the right hip with the Axivity AX3 accelerometer (Axivity Ltd., Newcastle upon Tyne) at a sampling rate of 100 Hz and with an acceleration amplitude of ±8 g. The raw data were resampled at 30 Hz and truncated at 6 g to correspond to SCAPIS data.

Oxygen uptake data were collected with the Oxycon Pro (Jaeger, BD Corporation). Resting oxygen consumption was calculated as the lowest 2-min mean oxygen consumption during the last 10 min of resting. Data collected from the last 2 min of each activity were used to calculate MET values by the quotient of total oxygen uptake and resting oxygen uptake. The MET values were used to calibrate processed accelerometer data. This calibration to a common scale was necessary in order to directly compare time composition and association with the measures of cardiometabolic health between processing components and settings, as the PA intensity measure is generated on different scales. The explained
variance ($R^2$) of METs from the calibration was determined for the combination of processing component settings and used as an indication of criterion validity, that is, the ability to assess METs. This is relevant only for the first three processing steps presented below, that is, hardware recalibration, frequency filtering, and number of acceleration axes used. The combination of settings of these three processing steps reached an explained variance of METs between 88%–94%.

A spectrum of 36 PA intensity categories of 0.5 METs each was calibrated up to 18 METs using linear regression. Crude intensity categories were also calibrated. They were defined as SED (<1.5 METs), LPA (1.5-< 3.0 METs), MPA (3.0-< 6.0 METs), VPA (6.0-< 9.0 METs), and VVPA (>9.0 METs). Still, the crude intensity categories may hide important variations in PA and associations with measures of health. For example, MPA is a broad category that includes activities corresponding to walking slow (3–4 km/h) up to walking fast (6–7 km/h), where the associations to health at the lower intensities may be markedly different compared to the upper intensities.

### 2.5 Accelerometer data processing

Figure 1 presents the processing steps from raw data to PA intensity and time spent in MET calibrated PA intensity categories applied to lab and free-living data. The raw data from each of the three axes (vertical, antero-posterior, medio-lateral) are bidirectional with positive and negative values.

The first step was to include hardware recalibration or not. Modern accelerometers are microelectromechanical system (MEMS) sensors with onboard firmware reducing loss of calibration. The assumption would be that recalibration is not required, although there are variations between units across activities and intensities. An unpublished observation indicated that the calibration could be lost. Inter-monitor variability contributes to random error which reduces the strength of association.

For some processing methods, hardware recalibration is recommended, for example, Euclidean Norm Minus One (ENMO). Hardware recalibration was performed by an autocalibration method. Stationary periods were identified and the discrepancy between the vector magnitude (VM) of these periods and the gravitational acceleration (1g) was minimized. VM was calculated as:

$$VM = \sqrt{(vert^2 + ant - post^2 + med - lat^2)}.$$ 

Robust calibration requires stationary periods from each of the three acceleration axes for several days, which was determined from on average 6–7 days free-living data in SCAPIS. Sufficient number of stationary periods from all three axes was achieved from all participants except one to perform the calibration with high quality.

The second step was to process accelerometer data to represent PA intensity. This was done either by applying a frequency filter or by the ENMO method. In this step, accelerometer data are turned into positive values. Four different filters were applied with an upper cut-off (low pass filter) at 1.63, 4 or 10 Hz, or without cut-off. As free-living data were collected with 30 Hz sampling rate, a movement frequency of maximum 15 Hz can be captured according to the Nyquist theorem. All filters included a lower cut-off at 0.29 Hz to filter out gravity and low-frequency noise (high pass filter). The 0.29–1.63 Hz filter is applied when generating ActiGraph counts. After filtering, negative values were turned into positive by taking the absolute value. A dead-band at 68 mg was applied to attenuate acceleration signals not representing movement. In addition, the 0.29–1.63 Hz filter included truncation to 2.13 g to be consistent with previous studies. ENMO was included as it represents a method without frequency filtering. However, it also applies an alternative way of turning accelerometer data into positive values.

The third step was to produce the final PA intensity measure (mg) was to use either vertical or VM data. Older accelerometers detected vertical acceleration only. Vertical acceleration is the major signal during ambulatory movement, but horizontal acceleration contributes to total acceleration during running. ENMO was originally developed from triaxial data. Therefore, vertical data was not used with this method.

The fourth step was to aggregate the mg data into epochs by taking the mean mg of each epoch. Commonly used epoch lengths of 1, 3, 10 and 60 s were analyzed. In the processing of accelerometer data, the information content is normally reduced from a sampling rate of 30–100 Hz into epochs of 1–60 s. Shorter epochs capture variation in PA intensity better than longer epochs.

The fifth step was to identify valid days by different wear time criteria. The settings used were 8, 10,12 or 14 h of wear time which are common in previous research. The optimal wear time criterion is a compromise between the inclusion of sufficient amount of time and enough participants. Increasing the wear time criterion would reduce the number of valid days and the sample size. Non-wear-time was defined as at least 60 min of consecutive zeros with allowance of up to 2 min between 0 mg and the acceleration value corresponding to 1.5 METs.
The *sixth and final step* was to apply the MET calibrations developed from lab data to calculate the time composition in different intensity categories. The time composition may be expressed by 24 h (i.e., min/day) or by wear time (i.e., % of wear time). Either option may be biased. Some participants may achieve more time in, for example, MPA and VPA due to more wear time when the time composition is expressed by min/day, while other participants with less wear time may seem more physically active when the time composition is expressed as % of wear time as they wore the accelerometer mostly when doing PA.

### 2.6 Cardiometabolic health

Five traditional measures of cardiometabolic health were used: aerobic fitness, body mass index (BMI), high density lipoprotein: total cholesterol ratio (HDL:Chol), blood glucose, and systolic blood pressure (SBP). Aerobic fitness (ml·kg⁻¹·min⁻¹) was estimated from the Ekblom-Bak two-point submaximal exercise test using validated algorithms covering the age range in SCAPIS. BMI was calculated as body weight/body height² (kg·m⁻²). A 100 mL venous blood sample was collected after an overnight fast and used for analysis of HDL (mmol·L⁻¹), total cholesterol (mmol·L⁻¹) and plasma glucose (mmol·L⁻¹). SBP (mmHg) was determined after 5 min at rest in supine position with an automatic device (Omron M10-IT, Omron Health Care Co).

### 2.7 Statistics

Vertical data are not used with ENMO, reducing the number of combinations of processing component settings from 640 to 576. Predictive validity was determined from the association between the PA intensity spectrum generated for the 576 combinations and the five measures of cardiometabolic health. As the 36 PA intensity spectrum categories can be collinear (in SCAPIS, correlation range 0.02–0.99, 6.8% \( r > 0.70 \), 1.7% \( r > 0.90 \)), partial least square regression (PLS) was used. The PLS combines the spectrum categories into one or more latent variables to maximize the covariance with the measure of health. The number of latent variables in each PLS model was determined from cross-validation by Monte Carlo resampling with 1000 repetitions. A backward selection procedure with a cut-off of a quarter of a standard deviation was used to ensure that a model with more latent variables was significantly better than using fewer latent variables. The optimal model would not include too many latent variables increasing the risk of being overfitted. Statistical significance of the PLS model was determined by permutation tests with 10⁴ repetitions. The strength of the PLS model for each of the 576 combinations is expressed by the explained variance (R²) of the measure of cardiometabolic health. The models were standardized for sex. The explained variance was thereafter used as outcome and the processing steps (components) as predictors in ordinary least square (OLS) regression for each measure of health. The six processing steps (Figure 1) were treated as categorical variables and their respective settings were transferred to dummy variables. The regression analyses were executed by excluding the reference setting for each processing component. The reference settings were no recalibration, vertical data, 0.29–1.63 Hz filter, 60-second epoch length, 8 h wear time, and 24 h time composition (min/day), respectively. They were selected based on their use in previous research. The result can be interpreted as the size and direction (positive or negative) of the effect.
on the explained variance by changing from the reference setting to another setting. The standardized regression coefficients were used to make the effects comparable between processing components. The explained variances and the regression coefficients were used as indicators of predictive validity.

Finally, selectivity ratio plots were used to visualize the effect of the processing component settings on the pattern of the association between the PA intensity spectrum categories and the measures of cardiometabolic health. The selectivity ratio represents the contribution of each of the 36 intensity spectrum categories to the association with the measure of health. 95% confidence interval was calculated for each intensity spectrum variable by bootstrapping with 10^4 repetitions. All data processing and statistics were performed in MATLAB 2020a (MathWorks). The MATLAB function ‘plsregress’ is available in the Statistics and Machine Learning Toolbox.

3 | RESULTS

3.1 | Section 1: Physical activity (Figure 2, montage of multiple figures)

There was no systematic difference in METs between recalibrated and not recalibrated data, although some random variation in the difference occurred, more in accelerometer data processed with the 0.29–1.63 Hz filter compared to the 0.29–10 Hz filter (Hardware calibration and frequency filtering). The mean (sd) calibration error before recalibration was 15.4 (33.7) mg and 3.9 (1.6) mg after recalibration.

The volume of acceleration frequencies during treadmill activity increased up to the peak at 3 Hz during the highest running speed, with additional signals frequencies up to 15 Hz (frequency filtering). The frequencies up to 3 Hz correspond to step frequency. The 0.29–1.63 Hz filter started to attenuate the acceleration signals from its upper cut-point at 1.63 Hz and with higher magnitude as the signal frequency increased. Consequently, an important part of the acceleration information from most ambulatory movement is eliminated. Similar results were observed during free-living in SCAPIS. Due to the different conditions of free-living compared to the strict treadmill protocol, the acceleration signal differs in frequency distribution. The volume of acceleration signal frequency is lower during free-living and peaks at 2 Hz. Further, the peak at 3 Hz corresponding to running is considerably smaller than from treadmill. ENMO also included acceleration signals of low frequencies as no frequency filter is applied. MET calibration for time distribution in intensity categories resulted in a distinct deviation of the 0.29–1.63 Hz filtered data from the other frequency filters.

FIGURE 2 Effect of the different processing on the physical activity intensity measure. The montage visualizes the different components and settings in separate figures.
ENMO contributed to the lowest amount of time distributed in PA of moderate and higher intensity.

Changing from vertical to VM data increased the PA intensity output during treadmill walking and running, and required MET calibration on a different scale before investigating time distribution during free-living in SCAPIS (Axes and frequency filtering). Although VM data increased the time being physically active during free-living, the size of the difference was small compared to expanding the frequency filter.

Each reduction of the epoch length from 60 to 1 s increased time spent in SED, MPA, VPA and VVPA and decreased time in LPA (epoch length). Increasing the wear time criterium from 8 h up to 14 h successively reduced the number of valid days in SCAPIS, with the sharpest decline from 12 to 14 h (wear time). However, the wear time criterium had only minor effect on time distribution.

Finally, the size of the difference in time distribution of MET calibrated data among processing settings was compared between processing components (all together). The largest difference between settings occurred for frequency filtering followed by epoch length and number of acceleration axes. The difference in time distribution between wear time settings and hardware recalibration settings was small.

3.2 | Section 2: Association with cardiometabolic health (Figure 3, montage of two figures)

Hardware recalibration and using VM data had generally minor positive effect on the explained variances (left figure). Specifically, recalibration with ENMO processing increased the explained variance between 0.02 (0.02) (glucose) and 0.80 (0.78) (aerobic fitness) mean (sd) % units compared to without recalibration. This is larger compared to the effect of recalibration with frequency filtering, where the mean (sd) change in explained variance varied between 0.0001–0.11 (0.005–0.42) % units. For glucose, there was a large negative effect of using VM data.

Expanding the frequency filter from the reference setting of 0.29–1.63 Hz had positive effect on the explained variance of all measures of cardiometabolic health, with large effect for aerobic fitness, HDL:Chol and SBP (left figure). However, a wider filter than 0.29–4 Hz did not contribute to more explained variance and only the 0.29–4 Hz filter had positive effect for BMI. ENMO contributed little to improvement of the explained variance.

Reducing the epoch length had positive effect on explained variance of all measure of cardiometabolic health, but not to the same degree as expanding the frequency filter (left figure). An exception was for BMI, where epoch length was the main processing component with 1-second epoch as the strongest setting.

Changing wear time criterium had only small effect on the explained variance of the measures of cardiometabolic health, but for some of them there was a negative effect by increasing wear time (left figure). Expressing time composition by wear time (%) reduced the explained variance compared to expressing by 24 h, although the effect varied between the measures of cardiometabolic health.

The right figure presents the variation of the size of explained variance as well as the processing settings among the lowest respective highest explained variances. The specific combination of the processing settings had large effect on the size of explained variance. Further, the size of explained variance differed between the measures of cardiometabolic health, with the highest explained variance for aerobic fitness and lowest for SBP.

The prevalence of specific processing settings was investigated among the 20 combinations of settings contributing to the lowest respective highest explained variances (right figure). The recalibration setting was prevalent more often among the highest explained variances compared to the lowest explained variances, except for glucose. The opposite was observed for the VM setting, which was more prevalent among the highest explained variances for aerobic fitness only. While the 0.29–1.63 Hz filter and ENMO settings were more prevalent among the lowest explained variances, the 0.29–4 Hz or wider filter settings were more prevalent among the highest explained variances with the exception for BMI where the 0.29–1.63 Hz filter was the most prevalent setting. However, the later only occurred together with an epoch length of 1 s. The 1-second and 3-second epoch lengths were the most prevalent epoch settings among the highest explained variances. Yet, the 1-second and 3-second epochs were also the most prevalent epoch settings among the lowest explained variances of HDL:Chol and glucose but in combination with the 0.29–1.63 Hz filter. For wear time, there was no clear prevalence pattern, while the 24 h time composition (min/day) was the most prevalent time composition setting among the highest explained variances.

Finally, the effect of the processing setting on the pattern of the association was most evident for the frequency filtering component and aerobic fitness. This association pattern is presented here as an example (Figure 4) (all association patterns are presented in supplement Figure S1–S3). Using the narrow 0.29–1.63 Hz filter moves the explained variance to higher PA intensities compared to the other filters, which corresponds to the upward move of the time distribution presented in Section 1.
4 | DISCUSSION

The decisions on processing of hip accelerometer data had large effect on the PA intensity measure as well as on the association with measures of cardiometabolic health. The largest consistent effect was achieved when widening the frequency filter from 0.29–1.63 to 0.29–4 Hz, confirming the higher predictive validity of this processing setting. Widening the frequency filter further did not add more relevant information. In addition, reducing the epoch length from 60 s to less than 10 s contributed to important changes in the time distribution of PA and improvements in explained variance, confirming also the higher predictive validity of using shorter epochs. Although VM data would capture more information on PA than vertical, it does not necessarily improve the association with cardiometabolic health. Wear time and time composition may have some effect on the PA and association with
cardiometabolic health, while minor effect would be expected from hardware recalibration.

The 0.29–1.63 Hz frequency filter, applied in the generation of the ActiGraph counts, is set too narrow to capture variation in movement intensities. Normal walking pace at 4 km·h⁻¹ correspond to a step frequency of 1.7 Hz, fast walking at 6 km·h⁻¹ to 2.4 Hz, jogging at 8 km·h⁻¹ to 2.6 Hz, and running at 12 km·h⁻¹ to 2.8 Hz. Hence, a large proportion of the acceleration signal is eliminated already for walking. Widening the filter up to 4 Hz would capture most relevant acceleration signals generated in SCAPIS, as shown with the signal frequency distribution with a peak at 2 Hz and a smaller peak at 3 Hz that indicates running (Figure 2, Frequency filtering). Still, capturing acceleration signals with frequencies higher than 4 Hz may be important in other populations, such as children or young adults including athletes.

The narrow 0.29–1.63 Hz filter impairs discrimination of acceleration signals generated at different PA intensities and increase the risk of misclassification of MET calibrated data. The misclassification can be in both directions, that is, both underestimation and overestimation. However, as more time is accumulated in MPA (minutes-hours) compared to VPA-VVPA (seconds-minutes), more time is misclassified to higher PA intensities associated with higher frequencies (MPA → VPA-VVPA) than the opposite (VPA-VVPA → MPA). The consequence is the paradoxal result of more time being physically active using the 0.29–1.63 Hz filter compared to wider filters in this and other studies. Previous research has shown that VPA may be required to increase HDL. As for aerobic fitness, the results herein show the importance of using a wider frequency filter to capture higher PA intensities associated with more HDL (Figure 3).

ENMO processes accelerometer data without frequency filter, which was evident from the frequency spectrum where low frequency noise was included (Figure 2, Frequency filtering). ENMO turns acceleration data into positive values by first subtracting the gravity component and thereafter zeroing all remaining negative values. Unfortunately, the consequence is a differential measurement error, as accelerometer data from individuals and activities generating larger acceleration amplitudes are attenuated more than from individuals and activities relying on higher step frequency with smaller acceleration amplitudes. This effect was indicated when comparing children (smaller amplitude) and adults (larger amplitude) in previous studies, where children generated higher values with ENMO compared to methods where all negative values are turned to positive. ENMO will increase the random error and reduce the explained variance with measures of health as shown in the present study (Figure 3). ENMO also exists as a high pass filtered version (0.2–15 Hz). These additions to ENMO had small effects on the explained variance of PA energy expenditure from the doubly labeled water method, and may not affect the predictive validity of cardiometabolic health. However, this remains to be evaluated.

The epoch length was the second most important processing component for the explained variance of the measures of cardiometabolic health. As in previous studies, with each reduction in epoch length the time in SED, MPA and VPA increased while LPA decreased. Although, the effect on the strength of the associations was relatively small compared to frequency filtering, except for BMI. In this case the epoch length was the most important processing mode and 1-second epochs contributed to the strongest associations. Shorter epochs seem to capture variation in PA related to energy expenditure and total body fat better than expanding the frequency filter in SCAPIS. However, previous studies have shown stronger associations with measures of health with longer epochs, which contrasts the results in the present study. In addition to different samples investigated, these studies processed the accelerometer data with the narrow ActiGraph filter. A wide epoch length together with a narrow frequency filter is a very restrictive combination of settings. This combination may separate individuals highly active in continuous PA (e.g., walking, running) even further from other individuals and the association with health appears even stronger, which may explain the contrasting results.

Vertical acceleration is the dominating acceleration axis during all ambulatory movement, while horizontal acceleration adds information during running. Using VM data compared to vertical data contributed to more PA being captured in this study (Figure 2, Axes and frequency filtering) as well as in other studies. Vertical acceleration adds information during running. Using VM data compared to vertical data contributed to more PA being captured in this study. In a study performed on the pilot sample to SCAPIS, VM data reduced the strength of the association between PA and the risk of metabolic syndrome compared to vertical data. The signal frequency pattern in the SCAPIS sample with the major peak at 2 Hz indicates that most of the ambulatory movement is performed as walking, while the smaller peak at 3 Hz indicates a low proportion of running (Figure 2, Frequency filtering). Hence, vertical acceleration captures most relevant PA in SCAPIS, but VM data may add relevant information for the explained variance of aerobic fitness (due to running), as this was the only measure of cardiometabolic health where a positive effect was observed using VM data.
(Figure 3). Hence, it seems that irrelevant noise may be captured with VM data.

Increasing the wear time criterium up to 14 h reduced the number of valid days and the explained variance of aerobic fitness, BMI, HDL:Chol, and SBP but not for glucose, although the effects were small (Figure 3). In a previous study, the slope of the association between MVPA and body fat percentage decreased when increasing wear time criterium.14 Expressing the time composition by 24 h (i.e., min-day\(^{-1}\)) or by wear time (i.e., %) varies largely between studies. In the present study, expressing by wear time reduced the explained variance of all measures of cardiometabolic health compared to expressing by 24 h, but the effect was only minor for some of the measures. Hence, these two processing components may be of some importance. A wear time of 10 h and expressing the PA by 24 h may be useful decisions for SCAPIS and similar populations, when the accelerometer is taken off during sleep.

Finally, it seems that recalibration of the .gt3x files from the ActiGraph accelerometer is not required, as there was only a minor effect on the PA intensity measure and the explained variance of cardiometabolic health (Figure 2, Hardware calibration and frequency filtering, and Figure 3). Previous studies have confirmed high intraclass correlation of the ActiGraph GT3X+ accelerometer used in SCAPIS.21,22 Still, it may be important to check the data after firmware updates. An interesting observation was the distinct pattern of larger individual variation of the difference between recalibrated and not recalibrated data for the 0.29–1.63 Hz filter compared to the wider 0.29–10 Hz filter (Figure 2, Hardware calibration and frequency filtering). Hence, there may be additional random error with a narrow frequency filter due to the calibration status of the accelerometers. Recalibration is recommended for ENMO.24 Frequency filtering would attenuate the effect of recalibration.24 ENMO does not use frequency filtering but subtracts the gravity component as part of the acceleration data processing. This likely explains why there was a larger effect of recalibration on the association with the measures of cardiometabolic health for ENMO compared when including frequency filtering, even if the effect was small. Consequently, the contribution of hardware recalibration may depend on what metrics used. The ActiGraph .gt3x files used in this study demonstrated a calibration offset of 15.4 mg and a standard deviation of 33.7 mg. With the application of a bandpass filter of 0.29–4 Hz, recalibration would be of no importance. However, we cannot draw conclusions about the importance of recalibration of other accelerometer brands.

Based on the results from this study, we propose standard settings for the six processing components investigated to promote high data quality and comparability between studies (Figure 5). This standard is primarily applicable to hip accelerometer data collected with the ActiGraph GT3X accelerometer during daytime in middle-aged populations, to generate the continuous measure of PA intensity.9,19,26 Similar investigations are encouraged in other populations and with other measures of health. In order to further develop a methodological standard to process hip accelerometer data, other accelerometer processing and metrics would be compared. Still, this need to be performed in a stepwise and logical manner to not mixture the complexity of different combinations of processing which would otherwise interfere with interpretation and understanding.

4.1 Strength and weaknesses

This is the first study to investigate the effect of multiple accelerometer data processing components on PA and the association with measures of cardiometabolic health in a large, unselected cohort. Statistics were used to consider linearity between the large number of PA intensity categories investigated. The comprehensive analyses of the associations with the different measures of cardiometabolic health allowed determination of the predictive validity of the settings included. Still, there are other processing components and accelerometer metrics not included in this study. For example, the settings of non-wear time have been investigated in previous research. As the non-wear time setting affects the sedentary part of the PA intensity spectrum and this part has only weak association with the measures of cardiometabolic health in SCAPIS when performing PLS regression, it was expected to be of little relevance in this study. Other metrics used in research to include for comparison to could be, for example, MAD, Activity Index, and MIMS. Further, the results cannot be directly applicable to other accelerometer brands than the ActiGraph accelerometer. Previous research has shown that different accelerometer brands do not always generate equal output when applying the same processing even if the differences can be considered small.26,35

The results are only applicable to the five measures of cardiometabolic health investigated. They cannot directly be applied to other populations with different PA patterns, although the fundamental principles are still valid. Finally, this study determined the predictive validity of the different combinations of accelerometer data processing settings and not the criterion validity. In the latter, a free-living criterion would be required and the doubly labeled water method would provide a criterion measure, although it is energy expenditure
and not a measure of movement. Information about the criterion validity is of importance when the PA behavior or change in PA behavior is of interest per se, for example, to determine the volume of PA to increase energy expenditure to lose body fat. The MET calibrations performed in this study could be used to assess free-living energy expenditure. The explained variance of METs was generally very high. Still, free-living criterion validation with the doubly labeled water method would be required, which was not available for this study. Information about both criterion and predictive validity would be optimal to draw conclusions about the importance of the different accelerometer data processing components and settings.

5 | CONCLUSIONS

Accelerometry is more objective than self-report when it comes to data collection. However, several decisions on accelerometer data processing must be made and have large effect on the time distribution of PA and the association with measures of health. The results of the present study provide the fundament for a methodological standard of settings when processing hip accelerometer data collected during daytime to a continuous measure of PA intensity in middle-aged populations. The implementation of this standard would improve quality and comparability in epidemiological and clinical research. Future studies would target other populations, measures of health, but also other PA metrics.

AUTHOR CONTRIBUTIONS
DA conceptualized the study. JF performed the processing of data, and DA and JF performed the analysis and visualization. EEB, ÖE, GB, and MB were responsible for data collection. DA drafted the manuscript and JF, EEB, ÖE, GB, and MB made substantial contribution to revision. All authors have approved the final version of the manuscript and are accountable for their contribution to and the quality of the work.

FUNDING INFORMATION
The main funding body of SCAPIS is the Swedish Heart-Lung Foundation, but is also supported by the Knut and Alice Wallenberg Foundation, the Swedish Research Council and ANNOVA (Sweden’s Innovation agency, the University of Gothenburg and Sahlgrenska University Hospital, Karolinska Institutet and Stockholm County Council, Linköping University and University Hospital, Lund University and Skåne University Hospital, Umeå University and University Hospital, Uppsala University and University Hospital. Further, the accelerometer data processing was enabled by resources in project SNIC 2021/22–530 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX, partially funded by the Swedish Research Council through grant agreement no. 2018–05973.
REFERENCES


DATA AVAILABILITY STATEMENT

SCAPIS data and materials are accessible by formal application to the SCAPIS administration and presenting ethical approval (scapis.org). MEDAL data and materials are accessible by reasonable request to the corresponding author (daniel.arvidsson@gu.se).

PERSPECTIVES

A methodological standard to process accelerometer data is important for quality and comparability between studies. This cross-sectional study investigated multiple components of processing hip accelerometer data to a physical activity intensity metrics in middle-aged individuals. Frequency filtering and epoch length had large effect on the distribution of physical activity intensity and association with different measures of cardiometabolic health, while hardware recalibration, number of acceleration axes, wear time criterium, and time composition measure (minutes vs. %) were less important. The proposed standard to process hip accelerometer data therefore includes specific settings for frequency filtering and epoch length to capture the association between physical activity intensity and cardiometabolic health in middle-aged individuals. The standard is more flexible to the settings of the other processing components investigated due to their minor influence.

ORCID

Daniel Arvidsson © https://orcid.org/0000-0003-2365-187X

J. Fridolfsson © https://orcid.org/0000-0002-7003-4025

**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Arvidsson D, Fridolfsson J., Ekblom-Bak E., Ekblom Ö., Bergström G., Börjesson M.. Fundament for a methodological standard to process hip accelerometer data to a measure of physical activity intensity in middle-aged individuals. Scand J Med Sci Sports. 2023;00:1-13. doi:10.1111/sms.14541