Impact of ActiGraph Sampling Rate on Free-Living Physical Activity Measurement in Youth

Kimberly A. Clevenger[1], Jan C. Brønd[2], Kelly A. Mackintosh[3], Karin A. Pfeiffer[4], Alexander H.K. Montoye[5], Melitta A. McNarry[3]

[1]Behavioral Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute; [2]University of Southern Denmark, Odense, Denmark; [3]Applied Sports, Technology, Exercise and Medicine Research Centre, Swansea University, Bay Campus, Swansea, SA1 8EN, Wales, UK; [4]Department of Kinesiology, Michigan State University, East Lansing, MI, USA; [5]Integrative Physiology and Health Science Department, Alma College, Alma, MI, USA

Corresponding author:

Kimberly A. Clevenger, PhD, MPH

Address: 9609 Medical Center Dr, Rockville, MD 20850

Phone: 2192529365

Email: kimberly.clevenger@nih.gov

Disclosures

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention, the National Cancer Institute, or the National Institutes of Health. No copyrighted materials were used in this research.

Impact of ActiGraph Sampling Rate on Free-Living Physical Activity Measurement in Youth

Abstract-250 words

ActiGraph sampling frequencies of more than 30 Hz may result in overestimation of activity counts in both children and adults, but research on free-living individuals has not included the range of sampling frequencies used by researchers. **Objective:** We compared count- and raw-acceleration-based metrics from free-living children and adolescents across a range of sampling frequencies. **Approach:** Participants (n=445; 10-15 y) wore an ActiGraph accelerometer for at least one 10-h day. Vector Magnitude counts, Mean Amplitude Deviation, Monitor-Independent Movement Summary units, and activity intensity classified using six methods (four cut-points, two-regression model, and artificial neural network) were compared between 30 Hz and 60, 80, 90, and 100 Hz sampling frequencies using mean absolute differences, correlations, and equivalence testing. **Main results:** All outcomes were statistically equivalent, and correlation coefficients were ≥0.970. Absolute differences were largest for the 30 vs. 80 and 30 vs. 100 Hz count comparisons. For comparisons of 30 with 60, 80, 90, or 100 Hz, mean (and maximum) absolute differences in minutes of moderate-to-vigorous physical activity per day ranged from 0.1 to 0.3 (0.4 to 1.5), 0.3 to 1.3 (1.6 to 8.6), 0.1 to 0.3 (1.1 to 2.5), and 0.3 to 2.5 (1.6 to 14.3) across the six classification methods. **Significance:** Acceleration-based outcomes are comparable across the full range of sampling rates and therefore recommended for future research. If using counts, we recommend a multiple of 30 Hz because using a 100 Hz sampling rate resulted in large maximum individual differences and epoch-level differences, and increasing differences with activity level.

Keywords: accelerometer, methodology, signal processing, sampling frequency, youth, pediatric

Introduction

When accelerometers were first used to characterize physical activity (PA) intensity, data were sampled at a low frequency and stored using proprietary measures, such as activity counts, to save storage space and battery life (Sasaki *et al.*, 2016). However, devices can now store raw acceleration data sampled at high frequencies for long periods of time. The most commonly used research-grade accelerometer brand (Wijndaele *et al.*, 2015), ActiGraph, has allowed users to specify the sampling rate and download raw acceleration data since the release of the GT3X in 2010 (John and Freedson, 2012). Since this capability became available, researchers have used a variety of sampling rates (Migueles *et al.*, 2017), even though most methods for analyzing these data were developed using a sampling rate of 10 or 30 Hz (e.g., Freedson *et al.*, 1998).

ActiGraph sampling rate has been shown to impact the conversion of ActiGraph raw acceleration data to activity counts in both adults (Brønd and Arvidsson, 2016; Clevenger *et al.*, 2021) and children (Clevenger *et al.*, 2019). While this raises concerns regarding the potential impact on surveillance, assessment of interventions, and comparability across studies, two of the three prior studies in this area (Clevenger *et al.*, 2019; Brønd and Arvidsson, 2016) were limited to lab-based or semi-structured activities. Specifically, a lab-based study estimated that 0.3 to 2.6% of epochs classifying PA intensity could be impacted by using a 100 instead of 30 Hz sampling rate (Clevenger *et al.*, 2019). This is congruent with the only free-living study, which employed a small sample of 20 adults, whereby approximately 3.8 to 5.4% of epochs were classified differently when comparing 100 and 30 Hz sampling rates, translating to differences of 3.6 to 5.4 min·day^{-1} of moderate-to-vigorous intensity PA (MVPA; Clevenger *et al.*, 2021). A more comprehensive understanding of the impact of ActiGraph sampling rate on free-living PA measurement using a larger sample size is therefore needed.

A major limitation of prior research (Clevenger *et al.*, 2021; Clevenger *et al.*, 2019) is the comparison of only 100 to 30 Hz sampling rates. Brønd et al. (2016) used a mechanical set-up

to demonstrate that multiples of 30 Hz (i.e., 60 and 90 Hz) were minimally affected by chosen sampling rate. This phenomenon should be replicated in free-living individuals as this would allow the capture of data at a higher sampling rate (90 Hz) without affecting count data. Additionally, given that United States surveillance efforts, such as the National Health and Nutrition Examination Survey (NHANES), use a sampling rate of 80 Hz (Troiano *et al.*, 2014), it is important to identify the impact of this specific sampling rate on outcomes. For 2003-2004 and 2005-2006 NHANES cycles, only count data were released, which may be impacted by sampling rate (Brønd and Arvidsson, 2016). Data from 2011-2012 and 2013-2014 NHANES cycles included a relatively new metric, Monitor-Independent Movement Summary (MIMS) units (John *et al.*, 2019), which is meant to be sampling rate agnostic as raw acceleration data are first up-sampled to 100 Hz, then filtered and integrated. However, whether MIMS is comparable across a range of sampling rates remains to be investigated.

In addition to MIMS, researchers are increasingly using other acceleration-based metrics, such as Mean Amplitude Deviation (MAD; Vähä-Ypyä *et al.*, 2015) due to the potential for improved comparability across device brands (Wijndaele *et al.*, 2015). Two prior studies have reported that raw acceleration metrics are less impacted by ActiGraph sampling rate than activity counts, but these were limited to small sample sizes (n=20-29) and/or lab-based investigations (Clevenger *et al.*, 2021; Clevenger *et al.*, 2019). Additionally, researchers have primarily compared accelerometer metrics across sampling rates or compared PA intensity classified using cut-point approaches (Clevenger *et al.*, 2021; Clevenger *et al.*, 2019; Brønd and Arvidsson, 2016). Little research has studied the impact of sampling rate when using other analytic techniques, such as machine learning (Clevenger *et al.*, 2019; Small *et al.*, 2021) or two-regression models (e.g., Crouter *et al.*, 2012), which may better predict activity intensity due to greater use of the rich information available from accelerometers. Thus, further investigation into the impact of ActiGraph sampling rate on raw acceleration-based metrics and the use of more advanced analytic techniques is warranted.

In this paper, we sought to compare count- and raw-acceleration-based metrics and PA intensity, as estimated using cut-point and more advanced techniques, between ActiGraph data sampled at 30 Hz compared to data sampled at 60, 80, 90, or 100 Hz. We hypothesized that outcomes would be statistically equivalent across sampling rates but that the smallest differences would be evident for the raw acceleration-based metrics and for the comparison between 30 Hz and multiples of 30 Hz (60 and 90 Hz).

Methods

Data collection and processing

This is a secondary data analysis of PA data from 10-15 year-old (mean ± standard deviation: 13.0 ± 1.1 y) youth in Wales (United Kingdom) with an average body mass index of 20.8 ± 4.1 $kg \cdot m^{-2}$. Prior to data collection, the protocol was approved by the Institutional Review Board. For each child who provided assent, one parent/guardian provided written informed consent. Youth (N=488; 260 boys, 224 girls, 4 missing values) wore an ActiGraph wGT3X-BT (ActiGraph LLC, Pensacola, FL) on an elastic belt over their right hip for seven consecutive days, except while sleeping or participating in water-based activities (i.e., showering, swimming).

Monitors were initialized to collect raw acceleration data at a sampling rate of 100 Hz and data were downloaded as .gt3x files using ActiLife software (version 6.13.4; firmware 1.2.0). In accord with previous research (Clevenger *et al.*, 2021; Clevenger *et al.*, 2019; Brønd *et al.*, 2017), the 100 Hz data were resampled to 30, 60, 80, and 90 Hz using Java software (Oracle Corp., Redwood Shores, CA) and the *resample* function available in MATLAB (MathWorks Inc., Natwick, MA) which uses a polyphase antialiasing filter. These sampling rates were chosen because of their frequent use in the literature (particularly for 30 and 100 Hz; Migueles *et al.*, 2017), use in NHANES (80 Hz; Troiano *et al.*, 2014), and Brønd et al.'s (2016) prior supposition that counts are unaffected by sampling rates that are multiples of 30 (i.e., 60 and 90 Hz). While

ActiLife users can select sampling rates of 40, 50, and 70 Hz, these were not included in the present study due to limited use in prior research (Migueles *et al.*, 2017). Thus, each participant had five .gt3x files for comparison.

For each of the five .gt3x files, counts·s^{-1} and raw acceleration were exported from ActiLife and loaded into R Studio (version 1.2.1335) using the 'AGread' package (version 1.1.1). Wear-time was classified using the 'accelerometry' package (version 3.1.2; Van Domelen and Pittard, 2014), with continuous strings of 20 minutes of zero counts in the vertical axis considered as non-wear (the most commonly used wear-time criteria in this age group; Migueles *et al.*, 2017). Participants were required to have at least one day of 10-h of valid wear data to be included in the subsequent analysis, a criterion which allowed the retention of the largest sample size (Migueles *et al.*, 2017). We note that the present study does not aim to produce unbiased estimates of PA intensity, but rather, to compare estimates across sampling rates. Only times classified as wear from all five sampling rates were included in subsequent analysis.

Three accelerometer metrics were calculated at a 5-s epoch for direct comparison; this epoch length was selected to capture the intermittent and transient nature of children's activity. Vector magnitude (VM; counts·5-s^{-1}) was calculated as the square root of the sum of the squared counts in each of the three axes. Mean amplitude deviation (MAD; m*g*) represents the typical distance between the square root of the sum of the squared values of the raw acceleration signals from each axis and the mean value for a given time period (Aittasalo *et al.*, 2015). MIMS were generated using a multi-step process of interpolation to 100 Hz, extrapolation to account for differences in device dynamic range, bandpass filtering, rectification and integration, and summation across the three axes (John *et al.*, 2019). MAD was calculated using the 'acc' package (version 1.3.3) while MIMS were calculated using the 'MIMSunit' package (version 0.10.0).

In addition, six classification methods were used to calculate the percentage of wear-time in sedentary time, light PA, and MVPA. These methods were chosen to include a variety of epoch lengths (1, 5, 10, 15, 60 s), model types (cut-points, two-regression, artificial neural network), and inputs (MAD, mean and variance of acceleration, VM counts). Of note, no methods for classifying activity intensity from MIMS are currently available. The classification methods included:

1) Hangii et al. (2013) cut-points for VM counts·s^{-1} which classified activity intensity as sedentary time (<3 counts·s^{-1}), light PA, (3-56 counts·s^{-1}), or MVPA (≥56 counts·s^{-1});

2) Aittasalo et al. (2015) cut-points for MAD (per 5-s) which classified intensity as sedentary time (<26.9 m*g*), light PA (26.9-331.9 m*g*), or MVPA (≥332 m*g*);

3) Crouter et al. (2012) two-regression model which predicts METs from VM counts·10-s^{-1} and the coefficient of variation in counts·10-s^{-1} which was then classified as sedentary time (≤1.5 METs), light PA (>1.5 to <3 METs), or MVPA (≥3 METs). The 'TwoRegression' package (version 0.1.2) was used to calculate the coefficient of variation (Hibbing, 2018).;

4) Montoye et al. (2019) artificial neural network using mean and variance in counts·15-s^{-1} to predict METs which were then classified as sedentary time (≤1.5 METs), light PA (>1.5 to <3 METs), or MVPA (≥3 METs). We note that this model was developed using data from a protocol involving a large amount of physical activity and therefore may not be intended to classify lower intensities.;

5) Romanzini et al. (2014) cut-points for VM counts·15-s^{-1} which classified intensity as sedentary time (≤180 counts·15-s^{-1}) light PA (181-756 counts·15-s^{-1}), or MVPA (≥757 counts·15-s^{-1}).;

6) Brønd et al. (2019) cut-points for VM counts·min^{-1} which classified intensity as sedentary time (<115 counts·min^{-1}), light PA (115-2160 counts·min^{-1}), or MVPA (≥2161 counts·min^{-1}).

Statistical analyses

We compared 30 vs. 60, 30 vs. 80, 30 vs. 90, and 30 vs. 100 Hz sampling rates. Other comparisons were not conducted to control familywise (type 1) error and because they have little practical relevance as the standard ActiGraph sampling rate is 30 Hz. It is therefore of greater interest to compare these other sampling rates (60-100 Hz) to the 30 Hz sampling rate, which by itself is used in approximately 70% of studies employing ActiGraph monitors (Migueles *et al.*, 2017). MAD, VM, and MIMS were compared at the epoch level (every 5-s) through calculation of mean absolute difference, mean absolute percent difference, and Pearson's correlation coefficient (*r*). Average MAD and VM, and summed MIMS (John *et al.*, 2019) over the entire wear period, and percentage of time spent in each activity intensity according to the six classification methods were calculated for each participant. Mean absolute differences, percentage differences, and Pearson's correlation coefficients were calculated. The 'blandr' package (version 0.5.1) was used to generate Bland Altman plots (Bland and Altman, 1986).

Two, one-sided tests of equivalence (TOST) were used to compare mean MAD and VM, total MIMS, and percentage of wear-time in each activity intensity using the 'TOSTER' package (version 0.4.0; Lakens, 2017). Specifically, 90% confidence intervals around the mean difference for each variable were constructed. If the confidence interval did not overlap or exceed the equivalence bounds, then the monitors were considered equivalent ($p<0.05$). Equivalence bounds were set as 5% of the mean value for each variable except MVPA, for which the equivalence bounds were modified 0.5 percentage points, as using the 5% of the mean criterion resulted in extremely narrow bounds that have little practical meaning (Clevenger *et al.*, 2021; Clevenger *et al.*, 2020). Equivalence bounds are reported in Supplementary Table 1. Power analysis (using the 'TOSTER' package) for the most stringent comparison (MVPA) indicates that our sample size was adequate to detect statistical equivalence.

Results

Of the total sample (N=488), 445 participants had sufficient data to be retained for the present analysis. Participants had approximately 86.5 ± 36.8 h of wear-time across 5.3 ± 2.0 days. Overall means for each accelerometer metric at a 5-s epoch (MAD, VM, MIMS) and for percentage of time spent in sedentary time, light PA, or MVPA according to the six different analysis methods are found in Table 1. Overall, participants spent from 1.7% (Aittasalo *et al.*, 2015) to 46.7% (Montoye *et al.*, 2019) of their time in MVPA.

Table 1. Accelerometer metrics and percentage of wear-time spent in each activity intensity by ActiGraph sampling frequency

	30 Hz	60 Hz	80 Hz	90 Hz	100 Hz
Metric					
VM (counts·5s^{-1})	70.4 ± 22.6	70.4 ± 22.6	71.4 ± 23.0	70.6 ± 22.7	72.4 ± 23.4
MAD (m*g*)	36.6 ± 12.9	37.2 ± 13.0	37.3 ± 13.0	37.3± 13.0	37.3 ± 13.0
MIMS	24,515.5± 11,115.9	24,520.5 ± 11,118.8	24,518.7 ± 11,118.1	24,517.0 ± 11,117.4	24,509.6 ± 11,114.0
Hangii VM counts·s^{-1} cut-points					
Sedentary (%)	73.5 ± 6.3	73.4 ± 6.3	73.4 ± 6.4	73.4 ± 6.3	73.3 ± 6.4
Light (%)	16.9 ± 3.7	16.9 ± 3.7	16.8 ± 3.7	16.9 ± 3.7	16.7 ± 3.6
MVPA (%)	9.6 ± 3.5	9.6 ± 3.5	9.8 ± 3.5	9.7 ± 3.5	9.9 ± 3.6
Aittasalo MAD·5-s^{-1} cut-points					
Sedentary (%)	76.4 ± 6.3	76.2 ± 6.4	76.2 ± 6.4	76.2 ± 6.4	76.2 ± 6.4
Light (%)	21.9 ± 5.5	22.0 ± 5.6	22.0 ± 5.6	22.1 ± 5.6	22.0 ± 5.6
MVPA (%)	1.7 ± 1.2	1.8 ± 1.2	1.8 ± 1.2	1.8 ± 1.2	1.8 ± 1.2
Crouter VM two-regression model					
Sedentary (%)	70.5 ± 7.1	70.5 ± 7.1	70.4 ± 7.1	70.5 ± 7.1	70.3 ± 7.1
Light (%)	15.9 ± 3.4	15.9 ± 3.4	15.8 ± 3.4	15.9 ± 3.4	15.6 ± 3.3
MVPA (%)	13.6 ± 4.6	13.6 ± 4.6	13.8 ± 4.6	13.6 ± 4.6	14.0 ± 4.7
Montoye artificial neural network					
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Light (%)	53.4 ± 9.1	53.4 ± 9.1	53.3 ± 9.1	53.4 ± 9.1	53.3 ± 9.1
MVPA (%)	46.6 ± 9.1	46.6 ± 9.1	46.7 ± 9.1	46.6 ± 9.1	46.7 ± 9.1
Romanzini VM counts·15-s^{-1} cut-points					
Sedentary (%)	71.8 ± 6.9	71.7 ± 6.9	71.7 ± 7.0	71.7 ± 6.9	71.6 ± 7.0
Light (%)	18.4 ± 4.0	18.4 ± 4.0	18.3 ± 3.9	18.4 ± 4.0	18.2 ± 3.9
MVPA (%)	9.8 ± 4.1	9.8 ± 4.1	10.0 ± 4.2	9.8 ± 4.1	10.2 ± 4.2
Brønd VM counts·min^{-1} cut-points					
Sedentary (%)	44.3 ± 9.5	44.3 ± 9.5	44.2 ± 9.4	44.3 ± 9.5	44.2 ± 9.5
Light (%)	42.0 ± 6.7	42.0 ± 6.7	41.9 ± 6.7	42.0 ± 6.7	41.8 ± 6.7
MVPA (%)	13.7 ± 5.4	13.7 ± 5.4	13.9 ± 5.4	13.7 ± 5.4	14.0 ± 5.5

Data are shown as mean ± standard deviation; VM: Vector Magnitude; MAD: Mean Amplitude Deviation; MIMS: Monitor-Independent Movement Summary; MVPA: Moderate-to-Vigorous Physical Activity. Note: the Montoye artificial neural network was developed using a protocol involving a high proportion of physical activity, resulting in the low estimates of sedentary time.

Mean absolute difference and percentage difference, correlation coefficients, and results from the equivalence tests are found in Table 2 (30 vs. 60 Hz), Table 3, (30 vs. 80 Hz), Table 4 (30 vs. 90 Hz), and Table 5 (30 vs. 100 Hz). For all tested sampling frequencies (60, 80, 90, and 100 Hz), mean metric values and percentage of time spent in each activity intensity were considered statistically equivalent to 30 Hz (p<0.001). While all correlations were classified as high (r≥0.970), the largest mean absolute differences and percentage differences were found for the comparison of 30 vs. 80 and 30 vs. 100 Hz. For example, percentage difference in VM was 0.1, 1.4, 0.3, and 2.8% for the comparison of 60, 80, 90, and 100 Hz, respectively, with 30 Hz.

Table 2. Mean absolute differences, correlations, and observed confidence intervals for the equivalence tests for accelerometer metrics and percentage of time spent in each activity intensity using 30 vs. 60 Hz sampling frequency

	Mean Absolute Difference ± SD	Mean Absolute Percentage Difference ± SD	Pearson's *r*	Bias ± SE	Confidence Interval for Equivalence Test	
					Lower	Upper
Metric						
VM (counts·5-s^{-1})	0.1 ± 1.1	0.1 ± 0.1	0.999	-0.1 ± 0.0	-0.1	-0.1
MAD (m*g*)	0.6 ± 0.3	1.7 ± 0.5	0.999	-0.6 ± 0.0	-0.6	-0.6
MIMS	5.1 ± 5.5	0.0 ± 0.0	0.999	-5.0 ± 0.3	-5.4	-4.6
Hangii VM counts·s^{-1} cut-points						
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.999	-0.2 ± 0.0	0.2	0.2
Light (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	-0.2 ± 0.0	-0.2	-0.2
MVPA (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	-0.0 ± 0.0	-0.0	-0.0
Aittasalo MAD·5-s^{-1} cut-points						
Sedentary (%)	0.2 ± 0.1	0.3 ± 0.1	0.999	0.2 ± 0.0	0.2	0.2
Light (%)	0.2 ± 0.1	0.8 ± 0.4	0.999	-0.2 ± 0.0	-0.2	-0.2
MVPA (%)	0.0 ± 0.0	2.8 ± 1.8	0.999	-0.0 ± 0.0	-0.0	-0.0
Crouter VM two-regression model						
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.999	0.0 ± 0.0	0.0	0.0
Light (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	-0.0 ± 0.0	-0.0	-0.0
Montoye artificial neural network						
Sedentary (%)	0.0 ± 0.0	7.7 ± 32.2	0.990	-0.0 ± 0.0	-0.0	0.0
Light (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	-0.0 ± 0.0	-0.0	-0.0
Romanzini VM counts·15-s^{-1} cut-points						
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.999	0.0 ± 0.0	0.0	0.0
Light (%)	0.0 ± 0.0	0.2 ± 0.1	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	-0.0 ± 0.0	-0.0	-0.0
Brønd VM counts·min^{-1} cut-points						
Sedentary (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	0.0 ± 0.0	-0.0	0.0
Light (%)	0.1 ± 0.1	0.1 ± 0.1	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	-0.0 ± 0.0	-0.0	-0.0

SD: standard deviation; SE: standard error; VM: Vector Magnitude; MAD: Mean Amplitude Deviation; MIMS: Monitor-Independent Movement Summary; MVPA: Moderate-to-Vigorous Physical Activity

Table 3. Mean absolute differences, correlations, and observed confidence intervals for the equivalence tests for accelerometer metrics and percentage of time spent in each activity intensity using 30 vs. 80 Hz sampling frequency

	Mean Absolute Difference ± SD	Mean Absolute Percentage Difference ± SD	Pearson's *r*	Bias ± SE	Confidence Interval for Equivalence Test	
					Lower	Upper
Metric						
VM (counts·5-s^{-1})	1.1 ± 0.6	1.4 ± 0.6	0.999	-1.1 ± 0.0	-1.1	-1.0
MAD (m*g*)	0.7 ± 0.3	1.9 ± 0.5	0.999	-0.7 ± 0.0	-0.7	-0.7
MIMS	3.8 ± 9.5	0.0 ± 0.0	0.999	-3.2 ± 0.5	-4.0	-2.5
Hangii VM counts·s^{-1} cut-points						
Sedentary (%)	0.1 ± 0.0	0.1 ± 0.1	0.999	0.2 ± 0.0	0.2	0.2
Light (%)	0.1 ± 0.1	0.5 ± 0.5	0.999	-0.2 ± 0.0	-0.2	-0.2
MVPA (%)	0.1 ± 0.1	1.6 ± 1.0	0.999	-0.0 ± 0.0	-0.1	-0.0
Aittasalo MAD·5-s^{-1} cut-points						
Sedentary (%)	0.2 ± 0.1	0.3 ± 0.2	0.999	0.2 ± 0.0	0.2	0.2
Light (%)	0.2 ± 0.1	0.9 ± 0.4	0.999	-0.2 ± 0.0	-0.2	-0.2
MVPA (%)	0.0 ± 0.0	2.9 ± 1.9	0.999	-0.0 ± 0.0	-0.1	-0.0
Crouter VM two-regression model						
Sedentary (%)	0.1 ± 0.1	0.1 ± 0.1	0.999	0.1 ± 0.0	0.1	0.1
Light (%)	0.2 ± 0.1	0.9 ± 0.7	0.999	0.1 ± 0.0	0.1	0.2
MVPA (%)	0.2 ± 0.1	1.6 ± 1.0	0.999	-0.2 ± 0.0	-0.2	-0.2
Montoye artificial neural network						
Sedentary (%)	0.0 ± 0.0	19.4 ± 46.0	0.970	0.0 ± 0.0	-0.0	0.0
Light (%)	0.1 ± 0.1	0.2 ± 0.1	0.999	0.1 ± 0.0	0.1	0.1
MVPA (%)	0.1 ± 0.1	0.2 ± 0.1	0.999	-0.1 ± 0.0	-0.1	-0.1
Romanzini VM counts·15-s^{-1} cut-points						
Sedentary (%)	0.1 ± 0.0	0.1 ± 0.1	0.999	0.1 ± 0.0	0.1	0.1
Light (%)	0.1 ± 0.1	0.7 ± 0.6	0.999	0.1 ± 0.0	0.1	0.1
MVPA (%)	0.2 ± 0.1	2.2 ± 1.4	0.999	-0.2 ± 0.0	-0.2	-0.2
Brønd VM counts·min^{-1} cut-points						
Sedentary (%)	0.1 ± 0.1	0.2 ± 0.2	0.999	0.1 ± 0.0	0.1	0.1
Light (%)	0.2 ± 0.2	0.4 ± 0.3	0.999	0.1 ± 0.0	0.1	0.1
MVPA (%)	0.2 ± 0.1	1.6 ± 1.2	0.999	-0.2 ± 0.0	-0.2	-0.2

SD: standard deviation; SE: standard error; VM: Vector Magnitude; MAD: Mean Amplitude Deviation; MIMS: Monitor-Independent Movement Summary; MVPA: Moderate-to-Vigorous Physical Activity

Table 4. Mean absolute differences, correlations, and observed confidence intervals for the equivalence tests for accelerometer metrics and percentage of time spent in each activity intensity using 30 vs. 90 Hz sampling frequency

	Mean Absolute Difference ± SD	Mean Absolute Percentage Difference ± SD	Pearson's *r*	Bias ± SE	Confidence Interval for Equivalence Test	
					Lower	Upper
Metric						
VM (counts·5-s^{-1})	0.2 ± 0.2	0.3 ± 0.2	0.999	-0.2 ± 0.0	-0.2	-0.2
MAD (m*g*)	0.7 ± 0.3	2.0 ± 0.5	0.999	-0.7 ± 0.0	-0.7	-0.7
MIMS	3.2 ± 8.3	0.0 ± 0.0	0.999	-1.5 ± 0.4	-2.2	-0.8
Hangii VM counts·s^{-1} cut-points						
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.999	0.0 ± 0.0	-0.0	-0.0
Light (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	-0.0 ± 0.0	-0.0	-0.0
MVPA (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	-0.0 ± 0.0	-0.0	-0.0
Aittasalo MAD·5-s^{-1} cut-points						
Sedentary (%)	0.2 ± 0.1	0.3 ± 0.2	0.999	0.2 ± 0.0	0.2	0.3
Light (%)	0.2 ± 0.1	0.9 ± 0.4	0.999	-0.2 ± 0.0	-0.2	-0.2
MVPA (%)	0.0 ± 0.0	2.9 ± 1.9	0.999	-0.0 ± 0.0	-0.1	-0.0
Crouter VM two-regression model						
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.999	0.0 ± 0.0	0.0	0.0
Light (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.3 ± 0.3	0.999	-0.0 ± 0.0	-0.0	-0.0
Montoye artificial neural network						
Sedentary (%)	0.0 ± 0.0	10.0 ± 34.6	0.985	-0.0 ± 0.0	-0.0	0.0
Light (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.1 ± 0.1	0.999	-0.0 ± 0.0	-0.0	-0.0
Romanzini VM counts·15-s^{-1} cut-points						
Sedentary (%)	0.0 ± 0.0	0.0 ± 0.0	0.999	0.0 ± 0.0	0.0	0.0
Light (%)	0.0 ± 0.0	0.2 ± 0.2	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.4 ± 0.4	0.999	-0.0 ± 0.0	-0.0	-0.0
Brønd VM counts·min^{-1} cut-points						
Sedentary (%)	0.1 ± 0.1	0.1 ± 0.2	0.999	0.0 ± 0.0	0.0	0.0
Light (%)	0.1 ± 0.1	0.2 ± 0.1	0.999	0.0 ± 0.0	0.0	0.0
MVPA (%)	0.0 ± 0.0	0.4 ± 0.4	0.999	-0.0 ± 0.0	-0.0	-0.0

SD: standard deviation; SE: standard error; VM: Vector Magnitude; MAD: Mean Amplitude Deviation; MIMS: Monitor-Independent Movement Summary; MVPA: Moderate-to-Vigorous Physical Activity

Table 5. Mean absolute differences, correlations, and observed confidence intervals for the equivalence tests for accelerometer metrics and percentage of time spent in each activity intensity using 30 vs. 100 Hz sampling frequency

	Mean Absolute Difference ± SD	Mean Absolute Percentage Difference ± SD	Pearson's *r*	Bias ± SE	Confidence Interval for Equivalence Test	
					Lower	Upper
Metric						
VM (counts·5-s^{-1})	2.0 ± 1.1	2.8 ± 1.0	0.999	2.0 ± 0.1	1.9	2.1
MAD (m*g*)	0.7 ± 0.3	1.8 ± 0.5	0.999	0.7 ± 0.0	0.6	0.7
MIMS	6.7 ± 5.3	0.0 ± 0.0	0.999	-5.9 ± 0.3	-6.4	-5.4
Hangii VM counts·s^{-1} cut-points						
Sedentary (%)	0.1 ± 0.1	0.2 ± 0.1	0.999	-0.1 ± 0.0	-0.1	-0.1
Light (%)	0.2 ± 0.2	1.1 ± 1.0	0.999	-0.2 ± 0.0	-0.2	-0.2
MVPA (%)	0.3 ± 0.2	3.3 ± 2.0	0.999	0.3 ± 0.0	0.3	0.3
Aittasalo MAD·5-s^{-1} cut-points						
Sedentary (%)	0.2 ± 0.1	0.3 ± 0.2	0.999	-0.2 ± 0.0	-0.2	-0.2
Light (%)	0.2 ± 0.1	0.9 ± 0.4	0.999	0.2 ± 0.0	0.2	0.2
MVPA (%)	0.0 ± 0.0	2.7 ± 1.8	0.999	0.0 ± 0.0	0.0	0.0
Crouter VM two-regression model						
Sedentary (%)	0.1 ± 0.1	0.2 ± 0.1	0.999	-0.1 ± 0.0	-0.1	-0.1
Light (%)	0.3 ± 0.2	1.9 ± 1.3	0.998	-0.3 ± 0.0	-0.3	-0.3
MVPA (%)	0.4 ± 0.3	3.2 ± 1.8	0.999	0.4 ± 0.0	0.4	0.4
Montoye artificial neural network						
Sedentary (%)	0.0 ± 0.0	21.6 ± 49.1	0.969	-0.0 ± 0.0	-0.0	-0.0
Light (%)	0.2 ± 0.1	0.3 ± 0.2	0.999	-0.2 ± 0.0	-0.2	-0.1
MVPA (%)	0.2 ± 0.1	0.3 ± 0.2	0.999	0.2 ± 0.0	0.1	0.2
Romanzini VM counts·15-s^{-1} cut-points						
Sedentary (%)	0.1 ± 0.1	0.2 ± 0.1	0.999	-0.1 ± 0.0	-0.1	-0.1
Light (%)	0.3 ± 0.2	1.6 ± 1.1	0.998	-0.3 ± 0.0	-0.3	-0.3
MVPA (%)	0.4 ± 0.3	4.4 ± 2.7	0.998	0.4 ± 0.0	0.4	0.4
Brønd VM counts·min^{-1} cut-points						
Sedentary (%)	0.1 ± 0.1	0.3 ± 0.3	0.999	-0.1 ± 0.0	-0.1	-0.1
Light (%)	0.3 ± 0.2	0.7 ± 0.5	0.999	-0.3 ± 0.0	-0.3	-0.2
MVPA (%)	0.4 ± 0.2	3.0 ± 1.9	0.999	0.4 ± 0.0	0.4	0.4

SD: standard deviation; SE: standard error; VM: Vector Magnitude; MAD: Mean Amplitude Deviation; MIMS: Monitor-Independent Movement Summary; MVPA: Moderate-to-Vigorous Physical Activity

When collapsed to min·day^{-1} of MVPA (Table 6), mean absolute differences across the different classification methods were 0.1 to 0.3 min (30 vs. 60 Hz), 0.3 to 1.3 min (30 vs. 80 Hz), 0.1 to 0.3 min (30 vs. 90 Hz), and 0.3 to 2.5 min (30 vs. 100 Hz). Maximum differences in MVPA in min·day^{-1} were 0.4 to 1.5 min (30 vs. 60 Hz), 1.6 to 8.6 min (30 vs. 80 Hz), 1.1 to 2.5 min (30 vs. 90 Hz), and 1.6 to 14.3 min (30 vs. 100 Hz).

The magnitude of differences attributed to sampling rate varied between participants. For example, mean absolute differences in mean VM across participants were 0.0 to 0.5 counts·5-s^{-1} (30 vs. 60 Hz), 0.1 to 4.1 counts·5-s^{-1} (30 vs. 80 Hz), 0.0 to 1.1 counts·5-s^{-1} (30 vs. 90 Hz), 0.3 to 7.4 counts·5-s^{-1} (30 vs. 100 Hz). Bland-Altman plots are found in Supplementary Figures 1-36. For the comparison of 30 Hz with 80 or 100 Hz sampling rates, bias increased as mean VM counts increased. As mean MAD and MIMS increased, bias increased only slightly, and the pattern was consistent across sampling rates. Most often, bias increased as time spent in MVPA (regardless of model or cut-points used) increased.

Table 6. Mean absolute difference and maximum difference in min·day^{-1} of moderate-to-vigorous physical activity by sampling rate across the different classification methods

Classification Method	Mean Absolute Difference ± SD				Maximum Absolute Difference			
	30 vs. 60	30 vs. 80	30 vs. 90	30 vs. 100	30 vs. 60	30 vs. 80	30 vs. 90	30 vs. 100
Hangii VM counts·s^{-1} cut-points	0.1 ± 0.1	0.9 ± 0.6	0.1 ± 0.1	1.9 ± 1.2	0.4	5.8	1.1	10.6
Aittasalo MAD·5-s^{-1} cut-points	0.3 ± 0.2	0.3 ± 0.2	0.3 ± 0.2	0.3 ± 0.2	1.5	1.6	1.6	1.6
Crouter VM two-regression model	0.1 ± 0.1	1.3 ± 0.8	0.3 ± 0.2	2.5 ± 1.5	1.0	6.9	1.4	11.7
Montoye artificial neural network	0.2 ± 0.2	0.6 ± 0.4	0.3 ± 0.2	0.9 ± 0.5	1.0	2.2	1.4	3.4
Romanzini VM counts·15-s^{-1} cut-points	0.1 ± 0.1	1.2 ± 0.8	0.2 ± 0.2	2.5 ± 1.5	0.8	6.6	1.5	14.3
Brønd VM counts·min^{-1} cut-points	0.2 ± 0.2	1.2 ± 0.8	0.3 ± 0.3	2.3 ± 1.3	1.5	8.6	2.5	11.1

SD: standard deviation; VM: Vector Magnitude; MAD: Mean Amplitude Deviation

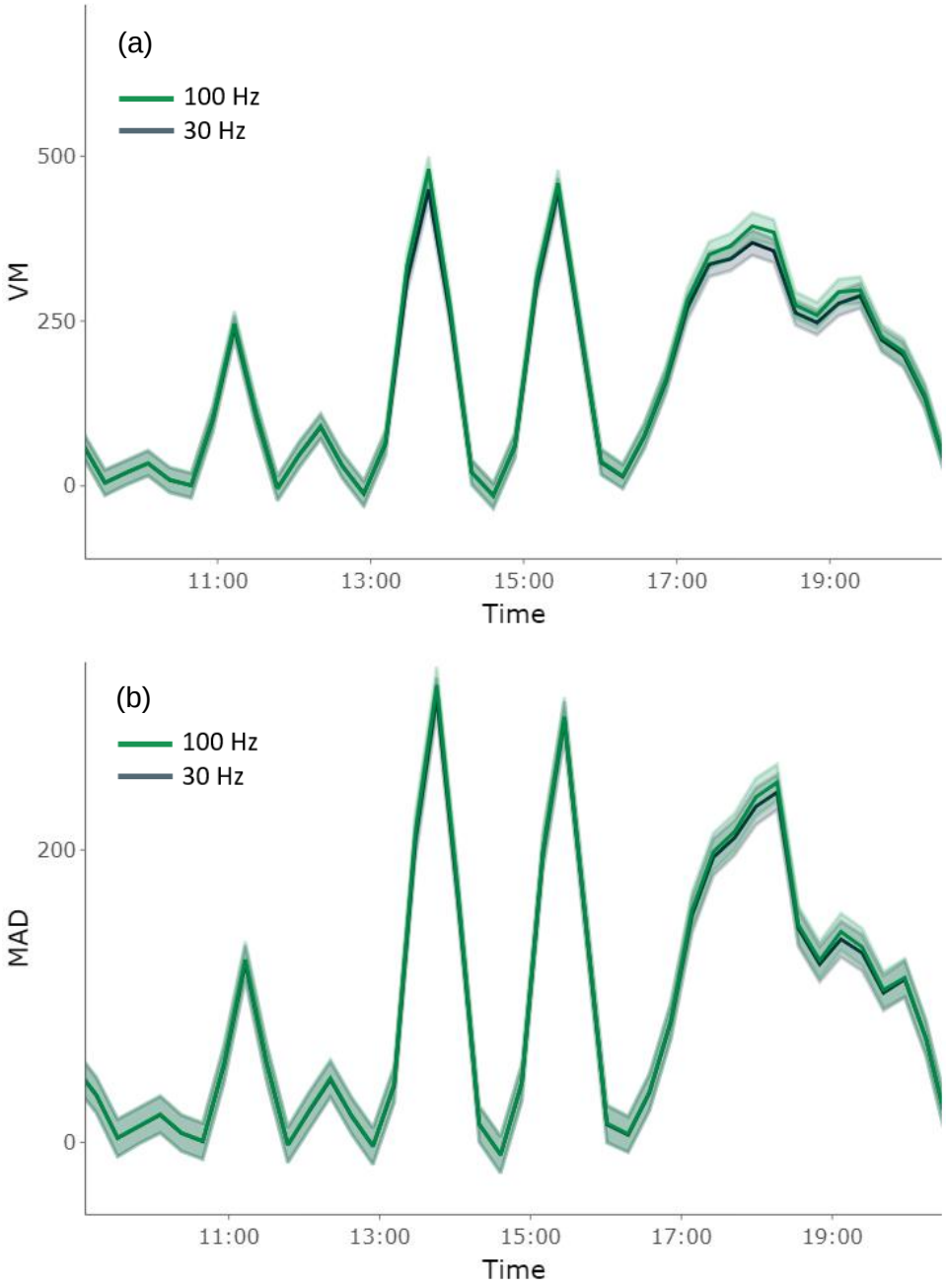
Epoch-level comparisons are reported in Table 7. Correlations were classified as high, ($r \geq 0.997$) and mean absolute percentage differences were less than 10% for all comparisons. Maximum epoch-level absolute differences in VM were 291.5 (30 vs. 60 Hz), 876.0 (30 vs. 80 Hz), 426.1 (30 vs. 90 Hz), and 1595.8 (30 vs. 100 Hz) counts·5-s^{-1}. Maximum epoch-level absolute differences in MAD were 325.6 (30 vs. 60 Hz), 336.5 (30 vs. 80 Hz), 344.5 (30 vs. 90 Hz), and 346.2 (30 vs.100 Hz) m*g*. Maximum epoch-level absolute differences in MIMS were 52.5 (30 vs. 60 Hz), 80.5 (30 vs. 80 Hz), 66.8 (30 vs. 90 Hz), and 59.8 (30 vs. 100 Hz). An example comparison of VM, MAD, and MIMS between 100 and 30 Hz sampling rate over one day from a single participant is shown in Figure 1.

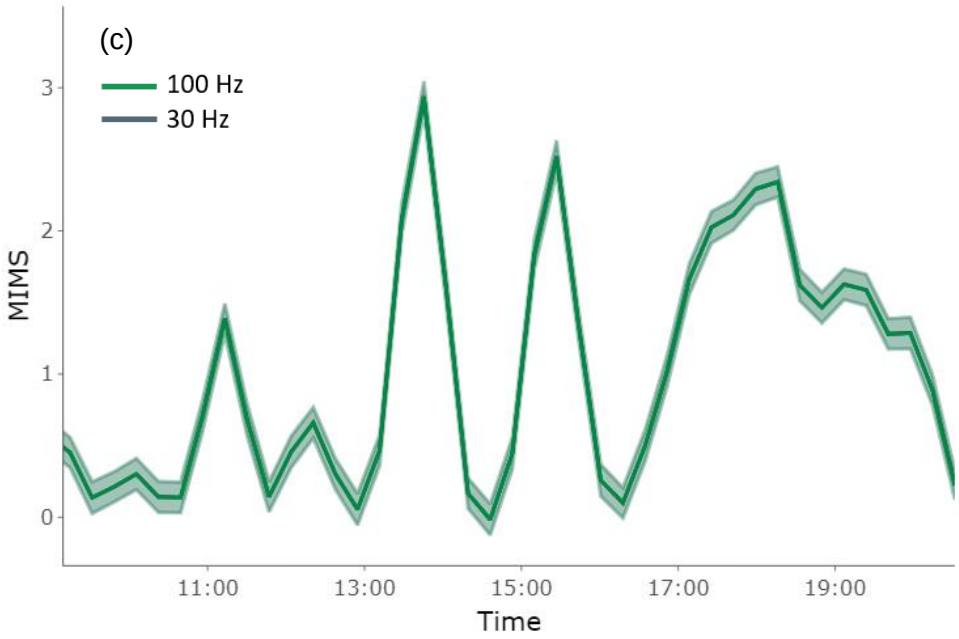
Table 7. Epoch-level (per 5-s) differences by sampling frequency in Vector Magnitude (VM), Mean Amplitude Deviation (MAD), and Monitor-Independent Movement Summary (MIMS) units

Comparison	VM (counts·5-s^{-1})			MAD (m*g*)			MIMS		
	Mean Absolute Difference ± SD	**Mean Absolute Percentage Difference ± SD**	***r***	**Mean Absolute Difference ± SD**	**Mean Absolute Percentage Difference ± SD**	***r***	**Mean Absolute Difference ± SD**	**Mean Absolute Percentage Difference ± SD**	***r***
30 vs. 60 Hz	0.6 ± 1.8	2.9 ± 14.9	0.999	0.7 ± 2.4	6.5 ± 6.7	0.999	0.0 ± 0.0	0.9 ± 10.7	0.999
30 vs. 80 Hz	2.0 ± 7.7	7.5 ± 23.9	0.999	0.8 ± 2.6	8.3 ± 9.6	0.999	0.0 ± 0.0	0.9 ± 10.4	0.999
30 vs. 90 Hz	0.8 ± 3.3	4.0 ± 17.9	0.999	0.8 ± 2.6	8.3 ± 9.6	0.999	0.0 ± 0.0	0.9 ± 10.4	0.999
30 vs. 100 Hz	2.9 ± 11.5	9.7 ± 27.2	0.997	0.8 ± 2.6	7.7 ± 9.4	0.999	0.0 ± 0.0	0.8 ± 10.1	0.999

SD: standard deviation; VM: Vector Magnitude; MAD: Mean Amplitude Deviation; MIMS: Monitor-Independent Movement Summary

Figure 1. Comparison of (a) Vector Magnitude (VM; counts·5-s^{-1}), (b) Mean Amplitude Deviation (MAD; m*g*), and (c) Monitor-Independent Movement Summary (MIMS) units between 100 and 30 Hz sampling rates for one day from one participant. Data are smoothed and shaded regions represent variability.




(c)
100 Hz
30 Hz
MIMS
3
2
1
0
11:00
13:00
15:00
17:00
19:00
Time

Discussion

In the present study, we compared three accelerometer metrics and six classification methods between a 30 Hz sampling frequency and 60, 80, 90, and 100 Hz sampling frequencies for data collected from free-living children wearing ActiGraph hip-worn monitors. While differences were largest for the comparison of VM counts between 30 and 100 Hz sampling frequencies, and smallest between 30 and 60 or 90 Hz, all outcomes were strongly correlated and statistically equivalent across sampling frequencies and resulted in minimal differences in outcomes (i.e., min·day^{-1} of MVPA; mean of ≤3 min). However, the impact of using a 100 vs. 30 Hz sampling rate when using VM counts to classify PA levels was large for a small number of participants (~5%). As differences may accrue over longer wear periods, this should be considered in future research.

Prior research has suggested that use of a sampling rate other than 30 Hz (or a multiple thereof) results in additional counts being recorded due to signals bypassing ActiGraph's filter (Brønd and Arvidsson, 2016). Comparing 100 and 30 Hz sampling rates in a prior lab-based pediatric study resulted in a mean absolute difference of 36.9 counts·15-s^{-1} (6.0%; Clevenger *et al.*, 2019), while a study of free-living adults reported differences of 44.2 counts·min^{-1} (6.0%; Clevenger *et al.*, 2021). Differences in the present study for the 100 and 30 Hz comparison were smaller (2.0 counts·5-s^{-1}; 2.8%), which translate to 6.0 counts·15-s^{-1} or 24.0 counts·min^{-1}. These smaller differences are likely due to our focus on children instead of adults (Clevenger *et al.*, 2021) or free-living behavior instead of laboratory-based activities, which often consist of a greater proportion of higher-intensity activities (Clevenger *et al.*, 2019). As the primary issue with using a sampling rate other than 30 Hz (or a multiple thereof) is that high frequency signals bypass ActiGraph's filter, this seems to have occurred to a lesser degree in this sample and/or setting.

Only the prior study in free-living adults explored the impact of using a 100 vs. 30 Hz sampling rate on the MAD metric, finding a mean absolute difference of 2.2 m*g* (5.8%), which was larger than the present study (0.7 m*g*; 1.8%) (Clevenger *et al.*, 2021). Smaller differences in the present study compared to prior research may be due to our focus on a pediatric sample instead of adults. There are two possible explanations for sampling rate differences in an acceleration-based metric like MAD. First, this may be due to differences in recorded acceleration by sampling rate, which occurs even with other monitor brands, such as Axivity or GENEA (Small *et al.*, 2021; Zhang *et al.*, 2012). Second, there may be internal processing of ActiGraph's 'raw' data as indicated by other researchers (Brønd and Arvidsson, 2016; John *et al.*, 2013). As interest in use of raw acceleration has increased over time, understanding how metrics like MAD are affected by data collection decisions is of interest.

A strength of the present study is the inclusion of MIMS units, which are an open-source acceleration-based metric used for national surveillance in the United States (Belcher *et al.*, 2021). While MIMS are intended to be sampling rate agnostic (John *et al.*, 2019), this has not been confirmed but could be affected if ActiGraph raw data are not truly raw (John *et al.*, 2013). The present study supports that MIMS are unaffected by sampling rate, with a mean absolute percentage difference of 0.0% and minimal bias. Researchers should continue to explore the properties of MIMS (e.g., reliability) and develop models for classifying PA intensity.

Differences by sampling rate were magnified at the epoch level versus when data were collapsed to mean values per participant over the entire wear period. Compared to the epoch-level comparison of 100 and 30 Hz sampling rates in free-living adults (Clevenger *et al.*, 2021), the present study demonstrated stronger correlation coefficients and lower mean absolute percentage differences for both VM (r=0.808 vs. 0.997; 40.6 vs. 9.7%) and MAD (r=0.744 vs. 0.999; 38.3 vs. 7.7%). Notably, the present study employs a much large sample size (n=445) than the prior study (n=20). However, in line with prior research, we report that differences attributable to sampling rate were larger for some participants, as illustrated by maximum

differences well above the group mean differences (Table 6). The Bland-Altman plots support that differences due to sampling rate tend to increase in more active participants, similar to Brønd et al. (Brønd and Arvidsson, 2016), and Clevenger et al. (Clevenger *et al.*, 2021). This raises some concern that sampling rate differences in future studies may be larger depending on the sample characteristics.

While we found that 60, 80, 90, and 100 Hz data were statistically equivalent when collapsed to mean values per participant over the entire wear period, it is also important to consider whether any of these differences are practically meaningful. Differences in min·day^{-1} of MVPA per 10-h wear day were, on average, 1 minute or less when using a 60, 80, or 90 Hz sampling rate and less than 3 min when using a 100 Hz sampling rate compared to a 30 Hz sampling rate. However, we note that maximum differences were sometimes considerably larger. Specifically, up to 6.3% of participants (depending on the classification method) had an absolute difference in MVPA of ≥5 min·day^{-1}. While these differences may seem small, we note that differences of this magnitude are in line with changes in daily MVPA reported in prior interventions (Kriemler *et al.*, 2011; Dobbins *et al.*, 2013). These estimates are also for a 10-h wear day and differences may accrue over longer wear periods. For example, the maximum difference of 14.3 min·day^{-1} of MVPA becomes a difference of 17.2 min·day^{-1} over a 12-h wear day or 100.1 min over an entire 7-day wear week (at 10-h per day). Similarly, up to 12.8% of participants had an absolute difference in MVPA of ≥5 min·day^{-1} over a 12-h wear day.

Of note, PA measurement is only affected when using VM counts to classify activity intensity as the maximum difference when using MAD cut-points was only 1.6 min. While there is no established method of classifying activity intensity using MIMS, the small epoch-level differences and minimal bias indicate this may also be a promising metric for classifying activity intensity across a range of sampling rates. Additionally, no participants had more than a 5 min·day^{-1} difference in MVPA when using a sampling rate of 60 or 90 Hz compared to 30 Hz. Thus, for most participants, the impact of using a sampling rate of 60 or 90 Hz compared to 30

Hz is negligible when measuring overall PA levels while some caution should be exercised when using a sampling of 80 or 100 Hz. In future research employing sampling rates other than 30 Hz (or multiples thereof), we recommend use of a non-count metric like MAD to classify activity, or downsampling to 30 Hz using a similar process to the present study, if the goal is to maintain comparability with prior research employing a 30 Hz sampling rate.

The majority of prior research has compared a 100 to 30 Hz sampling rate, which makes sense given that these are the two most commonly used sampling rates for ActiGraph devices (Migueles *et al.*, 2017). In addition, we compared 30 Hz data to sampling rates of 60, 80, 90 Hz. In line with the findings of Brønd et al. (2016), we found that 30, 60, and 90 Hz data are not exactly the same but that differences were smaller when comparing 30 Hz with 60 or 90 Hz than comparing 30 Hz with 80 or 100 Hz. ActiGraph accelerometers can also sample at 40, 50, and 70 Hz but these were not included in the present study due to their limited use in prior research (Migueles *et al.*, 2017). Brønd et al. (2016) reported that an ActiGraph sampling at 40 Hz recorded the highest counts in a mechanical experiment (even more than 100Hz) and had a different pattern of bias during a semi-structured activity involving walking and running. As the present study can only support equivalence of 60, 80, 90, and 100 Hz sampling rates with 30 Hz data, additional research will need to be done if researchers wish to use other sampling rates.

Another limitation of the present study is that we did not investigate the effect of sampling rate on other metrics, like vertical axis counts or mean acceleration, or outcomes like activity type. Future research may also replicate this analysis using other accelerometer wear-locations as wrist-worn ActiGraph monitors may be less affected by sampling rate (Clevenger *et al.*, 2019). In addition to the wrist, other wear-locations, such as the thigh or ankle, may also be of interest but were not included in present or prior studies. Our use of a minimum wear-time of a single 10-h·day^{-1}, which we selected to retain the largest number of participants in our sample, is a potential limitation. While this wear-time criteria should not affect our findings, which focused on differences attributable to ActiGraph sampling frequency and not getting unbiased

estimates of habitual PA participation, there may be concern that requiring fewer wear hours per day or fewer wear days results in lower estimates of PA participation. Finally, our goal was not to optimize ActiGraph's activity count algorithm. While the algorithm for calculating activity counts was recently published (Neishabouri *et al.*, 2022), it is still uncertain which specific step leads to this discrepancy by sampling rate or if modifications would address this issue.

This study provides the first large-scale investigation into the impact of chosen sampling rate on processing of ActiGraph accelerometer data, including both acceleration- and count-based metrics, several methods for classifying PA intensity, and multiple sampling rates. Our findings support that the investigated outcomes (percentage of time spent in each activity intensity and mean VM and MAD, and total MIMS over the wear period) are equivalent between 30 and 60, 80, 90, or 100 Hz sampling rates. However, the largest differences were found when comparing a 80 or 100 Hz sampling rate with the standard 30 Hz. Due to large individual differences and epoch-level differences, and increasing differences with activity level, we recommend that researchers employing count-based metrics continue using a 30 Hz sampling rate or multiple thereof (60, 90 Hz) if the goal is to maintain comparability with prior research which almost exclusively uses a 30 Hz sampling rate. As a prior review showed that approximately 20% of studies did not report the sampling rate used (Migueles *et al.*, 2017), we note that including this information in reporting of future trials is needed.

References

Aittasalo M, Vähä-Ypyä H, Vasankari T, Husu P, Jussila A-M and Sievänen H 2015 Mean amplitude deviation calculated from raw acceleration data: a novel method for classifying the intensity of adolescents' physical activity irrespective of accelerometer brand *BMC Sports Science, Medicine and Rehabilitation* **7** 1-7

Belcher B R, Wolff-Hughes D L, Dooley E E, Staudenmayer J, Berrigan D, Eberhardt M S and Troiano R P 2021 US Population-referenced Percentiles for Wrist-Worn Accelerometer-derived Activity *Medicine and Science in Sports and Exercise*

Bland J M and Altman D 1986 Statistical methods for assessing agreement between two methods of clinical measurement *The Lancet* **327** 307-10

Brønd J C, Aadland E, Andersen L B, Resaland G K, Andersen S A and Arvidsson D 2019 The ActiGraph counts processing and the assessment of vigorous activity *Clinical Physiology and Functional Imaging* **39** 276-83

Brønd J C, Andersen L B and Arvidsson D 2017 Generating ActiGraph counts from raw acceleration recorded by an alternative monitor *Medicine and Science in Sports and Exercise* **49** 2351-60

Brønd J C and Arvidsson D 2016 Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts *Journal of Applied Physiology* **120** 362-9

Cain K L, Sallis J F, Conway T L, Van Dyck D and Calhoon L 2013 Using accelerometers in youth physical activity studies: a review of methods *Journal of Physical Activity and Health* **10** 437-50

Clevenger K A, Brønd J C, Arvidsson D, Montoye A H, Mackintosh K A, McNarry M A and Pfeiffer K A 2021 Impact of ActiGraph Sampling Rate and Intermonitor Comparability on Measures of Physical Activity in Adults *Journal for the Measurement of Physical Behaviour* **4** 287-97

Clevenger K A, Pfeiffer K A, Mackintosh K A, McNarry M A, Brønd J, Arvidsson D and Montoye A H 2019 Effect of sampling rate on acceleration and counts of hip-and wrist-worn ActiGraph accelerometers in children *Physiological Measurement* **40** 095008

Clevenger K A, Pfeiffer K A and Montoye A H 2020 Cross-generational comparability of raw and count-based metrics from ActiGraph GT9X and wGT3X-BT accelerometers during free-living in youth *Measurement in Physical Education and Exercise Science* **24** 194-204

Crouter S E, Horton M and Bassett Jr D R 2012 Use of a 2-regression model for estimating energy expenditure in children *Medicine and Science in Sports and Exercise* **44** 1177

Dobbins M, Husson H, DeCorby K and LaRocca R L 2013 School‐based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18 *Cochrane Database of Systematic Reviews* **2** CD007651

Freedson P S, Melanson E and Sirard J 1998 Calibration of the computer science and applications, inc. accelerometer *Medicine and Science in Sports and Exercise* **30** 777-81

Hanggi J M, Phillips L R S and Rowlands A V 2013 Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph *Journal of Science and Medicine in Sport* **16** 40-4

Hibbing P 2018 TwoRegression: Process Data from Wearable Research Devices Using Two-Regression Algorithms. R package.

John D and Freedson P 2012 ActiGraph and Actical physical activity monitors: a peek under the hood *Medicine and Science in Sports and Exercise* **44** S86

John D, Sasaki J, Staudenmayer J, Mavilia M and Freedson P S 2013 Comparison of raw acceleration from the GENEA and ActiGraph™ GT3X+ activity monitors *Sensors* **13** 14754-63

John D, Tang Q, Albinali F and Intille S 2019 An Open-Source Monitor-Independent Movement Summary for Accelerometer Data Processing *Journal for the Measurement of Physical Behaviour* **2** 268-81

Kriemler S, Meyer U, Martin E, van Sluijs E M, Andersen L B and Martin B W 2011 Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update *British Journal of Sports Medicine* **45** 923-30

Lakens D 2017 TOSTER: Two one-sided tests (TOST) equivalence testing. R package.

Migueles J H, Cadenas-Sanchez C, Ekelund U, Nyström C D, Mora-Gonzalez J, Löf M, Labayen I, Ruiz J R and Ortega F B 2017 Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations *Sports Medicine* **47** 1821-45

Montoye A H, Clevenger K A, Mackintosh K A, McNarry M A and Pfeiffer K A 2019 Cross-Validation and Comparison of Energy Expenditure Prediction Models Using Count-Based and Raw Accelerometer Data in Youth *Journal for the Measurement of Physical Behaviour* **2** 237-46

Neishabouri A, Nguyen J, Samuelsson J, Guthrie T, Biggs M, Wyatt J, Cross D, Karas M, Migueles JH, Khan S 2022 Quantification of acceleration as activity counts in ActiGraph wearables *Scientific Reports* **12** 11958

Romanzini M, Petroski E L, Ohara D, Dourado A C and Reichert F F 2014 Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents *European Journal of Sport Science* **14** 91-9

Sasaki J E, da Silva K S, da Costa B G G and John D 2016 *Computer-Assisted and Web-based Innovations in Psychology, Special Education, and Health*. Elsevier pp 33-60

Small S, Khalid S, Dhiman P, Chan S, Jackson D, Doherty A and Price A 2021 Impact of Reduced Sampling Rate on Accelerometer-based Physical Activity Monitoring and

Machine Learning Activity Classification *Journal for the Measurement of Physical Behaviour* **4** 298-310

Troiano R P, McClain J J, Brychta R J and Chen K Y 2014 Evolution of accelerometer methods for physical activity research *British Journal of Sports Medicine* **48** 1019-23

Vähä‐Ypyä H, Vasankari T, Husu P, Suni J and Sievänen H 2015 A universal, accurate intensity‐based classification of different physical activities using raw data of accelerometer *Clinical Physiology and Functional Imaging* **35** 64-70

Van Domelen D R and Pittard W S 2014 Flexible R functions for processing accelerometer data, with emphasis on NHANES 2003–2006 *R Journal* **6** 52-62

Wijndaele K, Westgate K, Stephens S K, Blair S N, Bull F C, Chastin S F, Dunstan D W, Ekelund U, Esliger D W and Freedson P S 2015 Utilization and harmonization of adult accelerometry data: review and expert consensus *Medicine and Science in Sports and Exercise* **47** 2129-39

Zhang S, Murray P, Zillmer R, Eston R G, Catt M and Rowlands A V 2012 Activity classification using the GENEA: optimum sampling frequency and number of axes. *Medicine and Science in Sports and Exercise* **44** 2228-34