Alexander Burchartz[1] · Simon Kolb[1] · Leon Klos[1] · Steffen C. E. Schmidt[1] ·
Birte von Haaren-Mack[1,2] · Claudia Niessner[1] · Alexander Woll[1]

[1] Institute for Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

[2] German Sport University Cologne, Cologne, Germany

How specific combinations of epoch length, non-wear time and cut-points influence physical activity

Processing accelerometer data from children and adolescents in the nationwide MoMo study

Supplementary Information

The online version of this article (https://doi.org/10.1007/s12662-023-00892-9) contains supplementary material, which is available to authorized users.

Background

When measuring physical activity (PA) with devices, especially accelerometers, the objective is to detect it by measuring the acceleration occurring at a specific point on the human body (Burchartz et al., 2020a). To report physical behavior (PB) in public health settings, epidemiological studies use times spent at certain levels of activity intensity throughout the day and the week as result. Usually, these intensities are divided into

Data availability

The datasets generated and analyzed during the current study are not publicly available due to the strict ethical standards required by the German Federal Office for the Protection of Data to which study investigators are obliged to comply, but are available from the corresponding author upon reasonable request.

Birte von Haaren-Mack was still employed at the German Sport University in Cologne at the time of writing and submission, but has since transferred to the KIT in Karlsruhe. The work was therefore written when she was still in Cologne, but is now currently in Karlsruhe.

sedentary behavior (SED), light (LPA), moderate (MPA), and vigorous (VPA) intensity PA, although there are few more complex evaluation possibilities (Pfeiffer et al., 2022). To convert the captured accelerations into these intensity levels, the collected data needs to be preprocessed. Many parameters like device selection, carrying position and recording parameters such as recording frequency or filters influence how the measurement is transformed into the final results and largely affect the following estimation of activity (Burchartz et al., 2020a; Rowlands et al., 2018). Understanding the PA behavior of children and adolescents is severely limited by the presence of many sets of intensity-based limits for different brands and models of accelerometers. Trost (2007) used the term "cut-point conundrum" already in the early days of device-based assessment of PA to describe the bewildering number of calibration studies for different individual populations with a wide variety of devices and settings that have been published without a generalized and standardized approach, which has made a comparison between individual studies difficult, if not impossible.

Today, there are even more approaches to analyze accelerometer data, and for these reasons, it is important to doc-

ument all decisions made in recording and processing the data to allow later comparison with other studies (Arvidsson et al., 2019; Burchartz et al., 2020a; Migueles et al., 2017). In an earlier review, Cain, Sallis, Conway, van Dyck, and Calhoon (2013) reported that only about half of all studies report all decisions used during the process of data processing. In a recent review analyzing the approach of accessibility and use of novel analytic techniques for classifying PA intensity using raw or count-based accelerometer data, Pfeiffer et al. (2022) found that less than half of the models developed between 2000 and 2021 are not even publicly available. Therefore, it is not surprising that so many models are not used by other researchers.

One of the first things to look at when preprocessing the data is the accuracy of different algorithms for detecting the time the accelerometer is not worn, the so-called non-wear time (NWT; Migueles et al., 2017). However, determining NWT has received little attention in the literature, although it sometimes accounts for a large portion of the activity data in the 24h activity behavior cycle (Syed, Morseth, Hopstock, & Horsch, 2020). To evaluate compliance with the study, determining the time during which the accelerometer was not worn is very

important. For an accurate calculation of summary statistics, such as minutes spent in SED or LPA, MPA, and VPA, NWT has a profound impact (Syed et al., 2020; Vanhelst et al., 2019).

Second, the effect of different thresholds for PA intensity classification and of choosing the right epoch length for the target population based on age is of high interest as long as proprietary counts are used (Dencker, Svensson, El-Naaman, Bugge, & Andersen, 2012), since classifying activity is done by accumulating counts in a specific time interval, i.e., the epoch length. For children and adolescents, an epoch length between 1 and 5 s or the shortest possible epoch length is recommended by Bornstein, Beets, Byun, and McIver (2011); Migueles et al. (2017), and Vanhelst et al. (2012) due to the sporadic activity of children. Some studies (Banda et al., 2016; Bornstein et al., 2011; Breau, Coyle-Asbil, Haines, Ma, & Vallis, 2022; Dias Moura et al., 2019; Leppänen et al., 2022; Logan, Duncan, Harris, Hinckson, & Schofield, 2016; Migueles et al., 2019; Xing, Huang, & Sit, 2021) already compared the effects of different cut.points for children in the estimates of PB. However, to our knowledge, the specific combination of algorithms used in MoMo have not yet been compared as a collective. Migueles et al. (2019) provide a comprehensive comparison of different cut-point algorithms in overweight children. In doing so, they demonstrate that it is currently not and probably never will be possible to determine the prevalence of meeting PA guidelines based on accelerometer data. This is based on the fact that the apparent differences found range from almost zero to almost all participants meeting the guidelines, depending on the algorithm used for the evaluation. Giurgiu et al. (2022) found that most validation studies did not meet recommended quality principles when performing validation of wearables in real-world conditions and therefore recommend using standardized protocols to document all technical decisions for reproducibility. This is another reason why it is important that this point of the analysis in MoMo be adequately documented. Ultimately, the comparability of results from different studies in

Germany stands or falls with the algorithms used to analyze the data.

Many cohort studies use sensors or generations of software that have been replaced by newer versions over time, making it difficult to draw conclusions about changes in current technology. In the early days of accelerometry, only vertical acceleration was measured by the devices. Nowadays, sensors are much smaller and cheaper than earlier devices and, thus, it has become state of the art to measure acceleration in all three axes. Although this enables for even more accurate measurement, it also comes with new challenges. Newer cut-point intensity classification based on vector magnitude (the magnitude of a three-dimensional vector as the length of the entire acceleration, respectively, the movement in all three axes) produces higher MVPA time compared to estimations based on older ENMO (Euclidean Norm Minus One, see Bakrania et al., 2016) or vertical axis algorithms, independent of the place of attachment (Migueles et al., 2019).

As a consequence of the factors described above that prevented reasonable comparability of previous studies, it is now recommended to analyze multiple cut-point algorithms and identify the one where the validation sample best fits the target population at hand (Breau et al., 2022; Burchartz et al., 2020a; Migueles et al., 2017; Syed et al., 2020). This means that from the large number of available validation studies, one selects the study or algorithm that best fits one's own study. It is important that the following criteria are particularly similar to the validation study: age frame, gender, device, location, recording frequency, filter, epoch length, valid days. By capturing accelerometer data from 2014–2020, the nationwide Motorik-Modul study (MoMo) collected representative data on the PB of children and adolescents in Germany in two waves (MoMo wave 2 and MoMo wave 3). The objective of this study is to transparently examine the acceleration data of MoMo wave 3 (2018–2022) under the various aspects of epoch lengths, NWT, and cut-point sets as a combined overall construct. In particular, the influence of five different epoch lengths on two different NWT

algorithms was examined, as well as the differences when interpreting intensity classifications by two sets of cut-points for different age groups. The results are intended to make German researchers understand to what extent their own evaluation changes if individual parameters in the evaluation differ. It should be made clear that their data can only be compared directly with the results of MoMo if they are processed in exactly the same way.

Methods

Study design

The German Health Interview and Examination Survey for Children and Adolescents (KiGGS) is part of the Federal Health Monitoring System conducted by the Robert Koch Institute (RKI) and consists of regularly conducted nationwide surveys among children, adolescents, and young adults aged 0–29 years and living in Germany since 2003. MoMo is a submodule of the KiGGS study and aims to assess physical fitness, PA, and determinants of PA in children and adolescents (Woll, Albrecht, & Worth, 2017).

The entire study sample was drawn from the German resident population aged 4–17 years using a two-stage cluster sampling approach. Furthermore, participants from the baseline study (2003–2006), wave 1 (2009–2012), and wave 2 (2015–2017) were invited back for wave 3 (2018–2021). A detailed description of the study design and sampling procedure can be found elsewhere (Hoffmann et al., 2018; Mauz et al., 2019; Woll et al., 2017). KiGGS and MoMo provide nationally representative data on PA and SED of children, adolescents, and young adults living in Germany. A positive vote of the Ethics Committee of the Karlsruhe Institute of Technology on 23 September 2014 is available for the study.

Sample description

For the current analysis, only cross-sectional data from participants aged 6–17 years from MoMo wave 3 (2018–2022) were used ($n = 1525$). All data

used in this study had been collected before the first coronavirus disease 2019 (COVID-19)-related lockdown in March 2020. The participation in the study was voluntary and the guardians of the participants gave their written consent.

Device-based measured PA data

The technical and methodological details of the present study have previously been published (Burchartz et al., 2021; Burchartz et al., 2020b). In summary, for the assessment of device-based measured PA, ActiGraph GT3X+ and wGT3X-BT accelerometers (ActiGraph, LLC, Pensacola, FL, USA) were used for eight consecutive days. Participants were instructed to place the accelerometer on the right hip and wear it continuously, except during water activities or while sleeping. Data were sampled at a frequency of 30 Hz, downloaded as ActiLife GT3X raw device files, and stored for offline analysis. The GT3X files were then accumulated in ActiGraph count-based AGD files with a 1 s epoch length and converted to a Matlab-readable format.

For the present study, we analyzed all accelerometry data using the Matlab software version R2017a (The MathWorks Inc., Natick, MA, USA) to automate the data processing workflow without having to rely on the ActiLife 6.13.4 graphical user interface software (ActiGraph, LLC, Pensacola, FL, USA). The recordings of the first day were not considered for data analysis because the participants received the devices at different times throughout the day. In addition, the first day served as an adaptation period for the participants. In total, data for 10,557 days were recorded in $n = 1525$ participants, with data available for all 7 days in 97.2% of the sample. Data were analyzed for individual days to investigate the effect of the parameters for example on wear time in the context of the valid day criteria. The analysis focuses on methodological differences rather than test subjects' results. This allows examination of how different the results can be when evaluating them using different methods.

The epoch length data set of 1 s was therefore reintegrated into another four data sets with epoch length of 5 s, 15 s, 30 s, and 60 s, respectively. Afterward, each of these five epoch length datasets were analyzed by specifying the two different non-wear criteria of Troiano et al. (2008) & Choi, Liu, Matthews, and Buchowski (2011) resulting in 10 different configurations. The Choi et al. (2011) algorithms uses a 90 min time window for consecutive zero/nonzero counts with an allowance of 2 min intervals of nonzero counts with an up/downstream 30 min consecutive zero counts window. For the Troiano et al. (2008) uses a minimum of 60 min time window for consecutive zero/nonzero counts with an allowance of up to two continuous minutes of activity counts ranging from 1–100. Those algorithms were chosen since in the standard software ActiLife offers only three wear time validation options to users: Troiano, Choi and a daily/hourly algorithm by ActiGraph (ActiGraph, 2020). Thus, for the inexperienced user there is only the possibility to choose from these two algorithms. Two vertical-axis cut-point algorithms for classification into different activity classes (Evenson, Catellier, Gill, Ondrak, & McMurray, 2008; Romanzini,

Petroski, Ohara, Dourado, & Reichert, 2014, ☐ Table 1) were then applied to each of these 10 epoch length/NWT datasets, resulting in 20 final datasets including all possible combinations of the specific epoch lengths, NWT, and cut-points. This decision is oriented to the ICAD specifications and sample specifications. Evenson is also used in ICAD as an evaluation routine, also only the variant with vertical axis. However, our age range is larger than that of Evenson (validated with 5–9 years). That is why Romanzini was used for older children in MoMo (validated with 10–15 years, recommended also from 12–19 years by Migueles et al. (2017)).

The Troiano et al. (2008), Choi et al. (2011), Evenson et al. (2008), and Romanzini et al. (2014) algorithms were implemented in Matlab programming language according to their published descriptions and, where available, the published code (Choi et al., 2011). The Troiano and Choi implementations were validated against their implementations in the ActiLife Software. Cut-point limits of the Evenson & Romanzini algorithms

A. Burchartz · S. Kolb · L. Klos · S. C. E. Schmidt · B. von Haaren-Mack · C. Niessner · A. Woll

How specific combinations of epoch length, non-wear time and cut-points influence physical activity. Processing accelerometer data from children and adolescents in the nationwide MoMo study

Abstract

This study assesses three factors that influence the quantification of children's and adolescents' physical activity (PA) using accelerometers: selection of (1) non-wear algorithm, (2) epoch length and (3) cut-points. A total of 1525 participants from MoMo wave 3 (2018–2022), aged 6–17 years, wore GT3X accelerometers (ActiGraph, LLC, Pensacola, FL, USA) during waking hours. Acceleration counts were reintegrated into lengths of 1, 5, 15, 30, and 60 s epochs. Two non-wear time algorithms and two sets of cut-points were applied to each epoch length. Differences were found in both the comparison of the non-wear time algorithms and the comparison of the cut-points when the different epoch lengths were considered. This may result in large differences in estimated sedentary behavior and PA values. We propose to pool the data by merging and combining multiple accelerometer datasets from different studies and evaluate them in a harmonized way in the future. In addition to the need for future validation studies using short epoch lengths for young children, we also propose to conduct meta-analyses. This allows the use of data from multiple studies to validate cut-points and to propose a consensual set of cut-points that can be used in different settings and projects. The high discrepancy between results when comparing different epoch lengths has to be considered when interpreting accelerometer data and is regarded a confounding variable when comparing levels of PA between studies.

Keywords

Guidelines · Physical behavior · Data analysis · Youth · Vertical axis

Table 1 Count cutoff ranges for 1 s epochs of vertical-axis intensity algorithms

Count cutoffs	Calibration population (in years)	SED	LPA	MPA	VPA
Evenson et al. (2008)	5–9	≤ 2	> 2 to ≤ 38	> 38 to ≤ 67	> 67
Romanzini et al. (2014)	10–15	≤ 3	> 2 to ≤ 40	> 40 to ≤ 54	> 54

SED sedentary behavior, *LPA* light physical activity, *MPA* moderate physical activity, *VPA* vigorous physical activity

Table 2 Valid days depending on the valid hour criterion, epoch length, and non-wear time algorithm

NWT	EL (in s)	8 h (in %)	10 h (in %)	12 h (in %)	13 h (in %)	14 h (in %)	15 h (in %)	16 h (in %)
Choi et al. (2011)	1							
	5							
	15	83.83[a]	77.84[a]	64.26[a]	51.62[a]	35.43[a]	21.76[a]	12.17[a]
	30							
	60							
Troiano et al. (2008)	1	83.77	77.83	63.74	51.11	34.63	21.06	11.85
	5	83.60	77.56	63.28	50.41	34.03	20.52	11.40
	15	83.30	77.10	62.19	49.20	32.84	19.48	10.63
	30	83.01	76.61	61.12	47.61	31.33	18.20	9.61
	60	83.12	76.72	61.50	48.04	31.67	18.44	9.48

NWT non-wear time algorithm, *EL* epoch length in s, *%* of $n = 21{,}114$ valid days overall
[a]The values of the different epoch lengths are uniform for the Choi algorithm within each valid hour criterion

were independently verified by two authors.

Statistics

All statistical tests were conducted using IBM SPSS 28 (IBM Corporation, Armonk, NY, USA). Descriptive analyzes, stratified by NWT algorithms, epoch length, and cut-point algorithms, were performed and means (*M*), standard deviations (*SD*), and percentages were reported, as well as the corresponding inference statistical parameters, including 95% confidence intervals for differences (*CI*).

Results

Regarding the wear time for each day of the week, there were no differences between the algorithms except in the tenth range (◧ Fig. 1, Supplement Table 4). While applying the Choi algorithm revealed more full-day NWT, applying the Troiano algorithm resulted in more NWT classification at less than 1 h. For both algorithms, the results indicated a wear time between 8 and 16 h (Choi = 73.5%,

Troiano = 74.6%) for almost three-quarters of the days. Wear times of 0–2 h were classified by both algorithms for only 9.4% of the days of the week. However, descriptive analysis of the NWT algorithms revealed that the mean wear time of Choi was independently constant in all epoch lengths and evaluation algorithms ($M_{Choi} = 825.01$, $SD_{Choi} = 165.26$). For Troiano, however, different epoch lengths had an influence on the result. Longer epoch length resulted in more NWT and therefore fewer valid days ($M_{Troiano1s} = 822.48$, $S_{Troiano1s} = 162.38$; $M_{Troiano5s} = 819.54$, $S_{Troiano5s} = 160.16$; $M_{Troiano15s} = 813.80$, $S_{Troiano15s} = 155.43$; $M_{Troiano30s} = 806.42$, $S_{Troiano30s} = 149.36$; $M_{Troiano60s} = 805.55$, $S_{Troiano60s} = 145.58$) (◧ Fig. 2).

The proportion of valid days in the data set depending on the criterion of valid hours required decreases rapidly above 10 h of wear time (◧ Table 2).

For the analysis of the two cut-point algorithms, only the Choi data set with more than 8 h of wear time was used. Differences in mean minutes per day were analyzed for all cut-point levels of both

algorithms and differences were found between the five epoch lengths (◧ Fig. 3).

For SED, a 196 min (29.5%) and a 152 min (22.4%) decrease were observed between the 1 s and 60 s epochs for the cut-points of Evenson et al. and Romanzini et al., respectively (◧ Fig. 3a). Differences in sedentary time were observed between each activity output for all epoch length. A decrease in SED time occurred as epoch length increased for both cut-points.

For LPA, an increase of 213 min (216.3%) and a decrease of 170 min (185.3%) were observed between the 1 s and 60 s epochs for Evenson et al. and Romanzini et al. cut-points, respectively (◧ Fig. 3b). Differences in LPA time were observed between each activity output for all epoch lengths. An increase in LPA time occurred as the epoch length increased for both cut-points.

An increase of 0.6 min (1.8%) and 1 min (7.9%) was observed between the 1 and 60 s epochs using Evenson et al. and Romanzini et al. cut-points, respectively (◧ Fig. 3c). There was an overall decrease in MPA between the 15 and 1 s as well as between the 15 and 60 s epochs. The decrease was 4 min (12.2%) and 3 min (10.6%) for Evenson et al. and 2 min (14.4%) and 1 min (7.7%) for the cut-points of Romanzini et al., respectively (◧ Fig. 3c). Differences in MPA time were observed between each activity output for all epoch length except for 5 s (Evenson et al.) with 30 s and 5 s (Romanzini et al.) with 60 s.

For VPA, a 17 min (58.9%) and a 19 min (47.2%) decrease were observed between the 1 s and 60 s epochs for the cut-points of Evenson et al. and Romanzini et al., respectively (◧ Fig. 3d). Differences were observed between the VPA times spent within and between all epoch length. A decrease in VPA time occurred as epoch length increased for both cut-points.

For moderate-to-vigorous physical activity (MVPA), 17 min (27.5%) and 18 min (30.9%) decreases were observed between the 1 s and 60 s epochs for Evenson et al. and Romanzini et al. Cut-points, respectively (◧ Fig. 4). The mean minutes of MVPA per day between both the cut-point algorithms and all
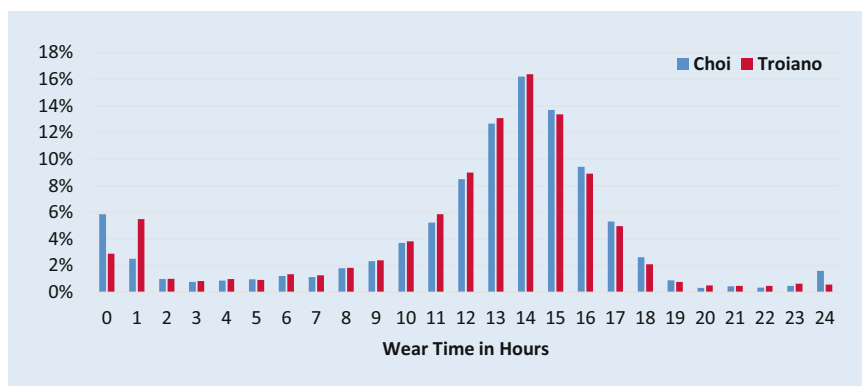
Fig. 1 ▲ Frequency distribution of wear time per hour, for both non-wear time (NWT) algorithms in % of datasets
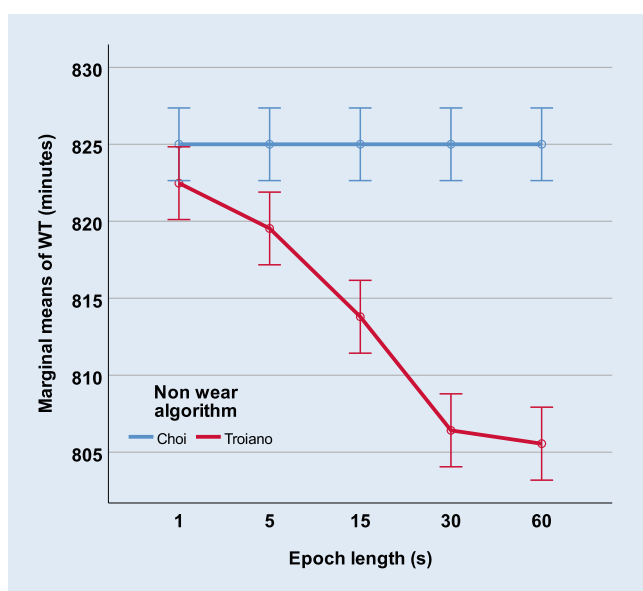

Fig. 2 ◄ Difference of wear time (WT) for different epoch lengths and non-wear time algorithms (with 95% confidence interval). Analysis for days with a wear time of more than 8 h

Calculating active days based on the WHO guidelines on PA (World Health Organization, 2010), which means more than 60 min of MVPA per day, there were some differences based on the epoch length and cut-point algorithm. Of the 10,557 recorded days, 4033 (38.2%) and 3686 (34.9%) were considered active days at a 1 s epoch length in Evenson and Romanzini cut-points, respectively. This number was reduced to 3189 (30.2%) and 2789 (26.4%) active days at an epoch length of 15 s, and 2257 (21.4%) and 1909 (18.1%) at an epoch length of 60 s for Evenson and Romanzini cut-points, respectively.

Discussion

The lack of a standardized preprocessing process for accelerometer data has challenged research efforts to gain a comprehensive understanding of children's and adolescents' PA and how they can be related to their health behaviors. It is recommended to analyze multiple accelerometer data preprocessing algorithms and identify those in which the validation sample best fits the sample at hand. Therefore, the present study aimed to examine the accelerometer data of MoMo wave 3 (2018–2022) concerning different data processing approaches used specifically in this scenario.

We sequentially assessed three factors that based on previous research may influence the quantification of children's and adolescents' PA using accelerometers: (1) non-wear time algorithm selection, (2) epoch length selection, and (3) cut-point selection. In our study, we identified variations when evaluating data analyzed with different epoch lengths. Specifically, when using different NWT algorithms and definitions of activity cut-points with epoch lengths that were not originally used to validate these algorithms, the resulting differences in the estimated SED and PA values turned out to be very large.

Both NWT algorithms used in this study were similar except for the tenth range (◻ **Fig. 1**, supplemental Table 4). Approximately 74% of recorded days had wear times between 8 and 16 h for both algorithms (Choi = 73.7%,
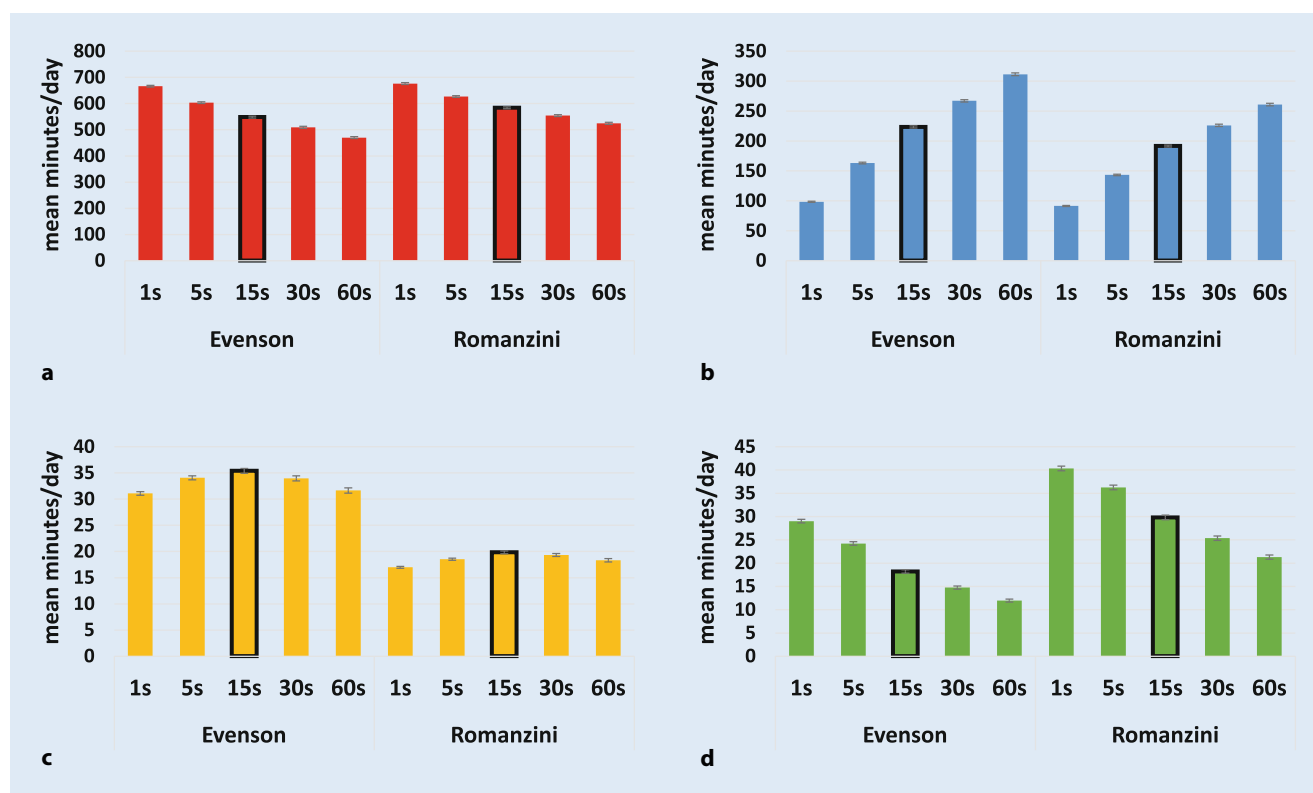
epoch length are statistically different. A decrease in MVPA time occurred as the epoch length increased for both cut-points.

For further analysis, only the data set with an epoch length of 15 s was analyzed since this was also used as a parameter in the validation study of the two cut-point algorithms (Evenson et al., 2008; Romanzini et al., 2014). When examining the intensities of individual days (without days with 0 min WT; $n = 9937$) for normal distribution, all tests were significant due to the high number of observations, but visual inspection of the Q-Q graphs showed that mainly some outliers in the high and low ranges are responsible (supplemental Fig. 1). Skewness and kurtosis showed that the distribution with SED was minimally left-skewed (Evenson: 0.19, Romanzini: 0.08) and val-

ues below the mean were more likely to be obtained (Evenson: 0.91, Romanzini: 0.99), and that marginal areas are more pronounced. For LPA, the distribution was almost symmetric (Evenson: −0.04, Romanzini: 0.10) for both algorithms but equally downward sloping (Evenson: −0.27, Romanzini: −0.27). MPA and VPA were no longer normally distributed for the two algorithms, with the majority of days having low values.

The results showed that MVPA was lower for Romanzini ($M = 42.89$; $SD = 34.19$) compared to Evenson ($M = 46.28$; $SD = 36.22$). Taking each step individually, MPA was lower for Romanzini ($M = 17.11$; $SD = 13.06$) compared to Evenson ($M = 30.53$; $SD = 22.95$). Whereas VPA was higher for Romanzini ($M = 25.78$; $SD = 23.74$) compared to Evenson ($M = 15.75$; $SD = 17.14$).

Fig. 3 ▲ Mean minutes spent in (**a**) sedentary time, (**b**) light, (**c**) moderate, and (**d**) vigorous physical activity per day interpreted using Evenson et al. and Romanzini et al. algorithms for five different epoch lengths (95% confidence interval). The 15 s epoch length bar is outlined in *black* as it is the one it was validated for
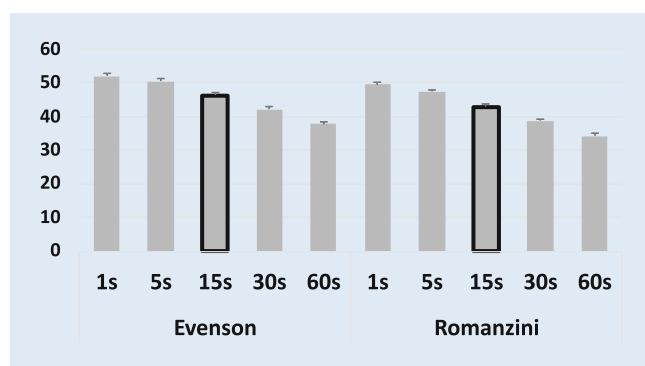

Fig. 4 ▲ Mean minutes spent in moderate-to-vigorous physical activity per day interpreted using Evenson et al. and Romanzini et al. algorithms for five different epoch lengths (95% confidence interval). The 15 s epoch length bar is outlined in *black* as it is the one it was validated for

Troiano = 74.7%). Although there were only small differences between the algorithms in general, different epoch lengths had a significant influence on the results when using the Troiano algorithm, as was also found by others (Banda et al., 2016).

The reason for the differences between the NWT algorithms may be due to how individual counts are accumulated during the allowance periods, regardless of whether the epoch length is 10 or 60 s. In the Troiano algorithm, the 2 min allowance period is filled with counts occurring during the epochs in this window. Counts occurring during a smaller epoch length like 10 s only contribute 10 s to the allowance window, while the same counts during a 60 s epoch contribute 60 s. Longer epoch lengths result in less NWT because the 2 min allowance window is reached faster for the Troiano algorithm.

Banda et al. (2016) on the other hand described the Choi algorithm accumulates all non-60 s epoch data into 60 s epoch data before applying the Choi NWT algorithm. Using a NWT algorithm that remains constant makes sense because the epoch length also affects further data processing. For example, the intensity algorithms for different age groups may differ in the specified epoch lengths.

Research studies have shown that the choice of NWT rules can impact the evaluation of SED time in adolescents over the age of 11 (Aadland et al., 2018; Vanhelst et al., 2019). It is recommended to use algorithms with shorter periods of continuous zeros for defining NWT to improve the accuracy in determining the activity levels. Vanhelst et al. (2019) recommend 30 min of continuous zeros because longer criteria, such as Choi et al. and Troiano et al. overestimate SED time for children under 10 years.

Two other studies suggested that a 20 min window is too short and recommended a longer window of at least

60 min of consecutive zeros as a realistic NWT criterion for younger children aged 7–13 years (Banda et al., 2016; Chinapaw et al., 2014). However, these studies had small numbers of participants, while the complete cross-section of the society is represented in MoMo. Furthermore, one of the goals in MoMo was to include as many different children as possible. Toftager et al. (2013) have shown that overweight children drop out of the data set more often when using shorter window sizes because their behavior is more sedentary and sometimes misclassified as NWT resulting in not reaching the required wearing time. Since Choi uses a 90 min window of consecutive zero/nonzero counts instead of, for example, a smaller window of 30 or 60 min, overweight children should be included more often in the MoMo dataset instead of dropping out.

The longitudinal design and the long-term study duration of almost 20 years of the MoMo study, with a total of 4 survey waves (baseline, wave 1–3), resulted in a very large age range and heterogeneity sample of participants. Although a large part of the data is drawn from a cross-sectional sample of participants aged 6–17 years in each wave, evaluations of the repeating and now adult participants aged 18–32 years are planned soon. The Choi algorithm has also been validated for an age range of 10–67 years (Choi et al., 2011). In addition, Aadland et al. (2018) recommended to standardize NWT algorithms across studies to avoid confounding and improve the comparability of children's PA levels. To be able to use validated algorithms for children, adolescents, and adults with different epoch lengths while still maintaining the comparability of the data, we decided to use Choi's algorithm for calculating the NWT in MoMo.

The study found that as wear time criteria became longer, there were fewer valid days for participants. A change from an 8 h to a 10 h wear time criteria resulted in 5% fewer valid days, which is acceptable, but a change to a 12 h wear time criteria resulted in a large 20% loss of participants, what is worth discussing, especially considering the diverse backgrounds of MoMo partic-

ipants. Therefore, the minimum wear time for MoMo and KiGGS accelerometer data sets was set to 8 h of recordings on 4 weekdays and 1 weekend day, which meets the requirements for inclusion in the International Children's Accelerometry Database (ICAD) (Sherar et al., 2011). Research suggests that measuring PA for 4 days is reasonable (Colley, Gorber, & Tremblay, 2010; Mâsse et al., 2005; Matthews, Ainsworth, Thompson, & Bassett, 2002; Toftager et al., 2013), although more days are even better (Chinapaw et al., 2014). To collect the highest possible number of valid data sets, participants should wear the accelerometer for 7 days following the day of the examination in the study center to ensure inclusion of weekdays and weekend days, which differ in PA levels (Burchartz et al., 2022; Chinapaw et al., 2014; Donaldson, Montoye, Tuttle, & Kaminsky, 2016; Matthews et al., 2002).

We found similar patterns of results for the two cut-point algorithms considered in MoMo for the average minutes per day classified in the intensity levels SED and PA. Both algorithms showed that SED and VPA decreased with longer epoch length, while LPA increased and MPA peaked at 15 s. Longer epoch lengths led to more SED being classified as LPA. This was shown in ◘ **Fig. 3** and supported by other studies (Banda et al., 2016; Logan et al., 2016). The wear time did not differ between the two algorithms, only the allocation of activity to intensity levels.

However, it must be acknowledged that although it is now very clear that small epoch lengths (e.g., 1 s as proposed by Aadland, Andersen, Anderssen, Resaland, & Kvalheim, 2020; Aadland & Nilsen, 2022; Bornstein et al., 2011; Migueles et al., 2017 and Vanhelst et al., 2012) are significantly better in interpreting activity behavior, there are still almost no new validation studies using short epoch lengths. For the future, these short epoch lengths are definitely preferable. However, in order to compare results with the original activity cut-point definitions (e.g., a cut-point validated at an epoch length of 15 s), the epoch lengths used there should be used, otherwise the results will be misleading.

Assuming that this is possible, at a minimum, open documentation of data processing decisions must be provided to help other studies interpret the results. This is even more important to nationally representative studies such as MoMo, since other studies with special subsamples (e.g., children with certain diseases) from Germany, for example, are using these results to compare their data.

Banda et al. (2016) and Logan et al. (2016) compared the Evenson and Romanzini cut-points and found similar results, but the comparability is complicated because they used different raw data sets (vertical axis with Romanzini vs. vector magnitude with Evenson). Breau et al. (2022) also compared different algorithms for young children and found significant variations in times spent in different intensities of PA. The choice of cut-points affects the proportion of children meeting guidelines, as shown in other studies (Banda et al., 2016; Breau et al., 2022; Logan et al., 2016). Thus, comparisons of movement intensities should not be made across studies with the same study population using different sets of cut-points. We again emphasize the need for additional studies validating WT algorithms and cut-points using smaller epoch lengths in children and adolescents, as already suggested elsewhere (Giurgiu et al., 2022).

On the other hand, Migueles et al. (2017) recommended using different cut-point criteria for different age groups. Due to the wide age range in the MoMo sample (6–32 years), this study divided participants into children (6–10 years) and adolescents (11–17 years) and compared the most suitable cut-points of Evenson (5–9 years) and Romanzini (10–15 years). These algorithms closely fit the age groups in MoMo and have similar validation protocols, both using an epoch length of 15 s. ◘ **Table 1** and ◘ **Fig. 5** also show similar cut-point ranges of the two algorithms. Although there are slight differences in the classification of SED and PA times, they are correlated with age. As children get older, their school curricular activities change from short to longer activity bouts, re-
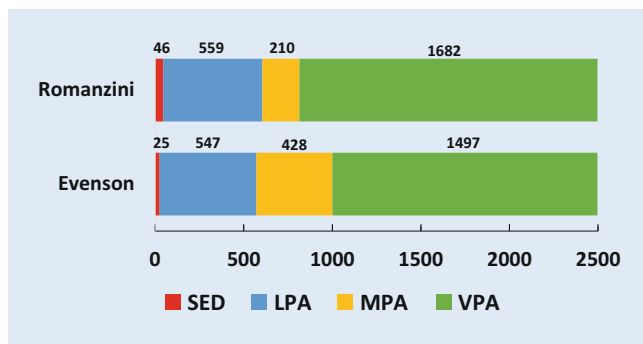
Fig. 5 ◄ Visualization of the cut-point ranges for 15 s epochs of vertical axis algorithms. *SED* sedentary behavior, *LPA* light physical activity, *MPA* moderate physical activity, *VPA* vigorous physical activity

sulting in more SED (Crane, Naylor, & Temple, 2018). This is also confirmed by PA behavior captured with MoMo questionnaires, which also showed more SED for older children (Burchartz et al., 2021). Differences in the classification of MPA and VPA compensate for each other and as they are normally considered together as MVPA (◘ **Fig. 4**). However, when examining individual classifications such as VPA, this factor should be kept in mind, especially since the percentage of VPA on an average day has been shown to remain stable across age groups, while the absolute time spent on VPA increases at the same time due to longer waking hours among adolescents (Burchartz et al., 2022).

The study investigated a large group of children and adolescents under free-living conditions and provides important results on NWT data processing, epoch length, and intensity classification. The study's comprehensive documentation improves replicability. However, some limitations should be acknowledged: (a) cut-points based on a vertical axis instead of vector magnitude are known to produce significantly lower estimations of time spent in MVPA (Migueles et al., 2019). The two cut-points under analysis in this study only used the vertical axis. The MVPA values of other vector magnitude cut-points might therefore be higher. Although recent studies (Breau et al., 2022; Leppänen et al., 2022) have investigated other cut-points, we are not aware of any study that has investigated the specific algorithms used in MoMo; (b) we did not identify sleep/wake states since the participants were told to remove the device for sleeping. Therefore, separate detection of sleep phases was not possible; and (c) we were

unable to evaluate the precision of the individual cut-point limits, as we did not have a comparison criterion measure available in the present study.

In device-based PA research, there is currently no consensus on the criteria for collecting and processing data, leading to inconsistent reporting methods that may be difficult for nonexperts to understand. One solution to this issue is to pool accelerometer data from different studies and process them in a standardized way to create comparable variables, as demonstrated by the ICAD initiative (Steene-Johannessen et al., 2020). By using standardized methods, the data can then be processed in a consistent manner to enable the creation of comparable accelerometer variables. Newer open-source metrics are increasingly being used to analyze accelerometer data (Leppänen et al., 2022). Previously, it was difficult to evaluate counts in device-based research because each manufacturer used different and undisclosed methods to calculate them. However, recent studies have shown that ActiGraph counts can correlate well with counts from other accelerometer brands using new methods (Brønd & Arvidsson, 2016; Brondeel et al., 2021; Clevenger et al., 2022). Nonetheless, there is a growing preference for using raw data to calculate certain metrics. ActiGraph has responded to this trend by releasing its proprietary algorithm for counting, which can now be used to convert raw data from other accelerometers into ActiGraph counts (Neishabouri et al., 2022). Open-source analyzes can further improve data analysis and make comparisons even between device models more feasible (Clevenger et al., 2022).

Conclusion

This study demonstrates the extent of variation in the results for physical activity (PA) levels across epoch lengths, different non-wear algorithms, and between different cut-point algorithms. We identified variations when evaluating data analyzed with different epoch lengths, specifically when different non-wear time (NWT) algorithms and definitions of activity cut-points are used with epoch lengths that were not originally used to validate these algorithms. As a consequence, the resulting differences in the results of estimated sedentary behavior (SED) and PA values can become very large.

Therefore, our results confirm previous studies and extend their findings to a sample of children and adolescents in Germany. We hope to have illustrated the extent to which the evaluation changes when individual parameters differ in the assessment. It can be further concluded from this that data can only be directly compared if they are processed in exactly the same way. The easiest way to do this is to pool raw data from different studies and analyze them in a harmonized manner. In addition to new validation studies with short epoch lengths for young children (e.g., 1 or 5 s), we also propose to conduct meta-analyses using data from multiple studies to validate cut-points to propose a consensual set of cut-points that can be used in different settings and projects.

Corresponding address

Alexander Burchartz
Institute for Sports and Sports Science, Karlsruhe Institute of Technology
Engler-Bunte-Ring 15, Building 40.40,
76131 Karlsruhe, Germany
alexander.burchartz@kit.edu

Declarations

References

Aadland, E., & Nilsen, A. K. O. (2022). Accelerometer epoch length influence associations for physical activity intensities with body mass index and locomotor skills in young children. *Journal of Sports Sciences*, *40*(14), 1568–1577.

Aadland, E., Andersen, L. B., Anderssen, S. A., & Resaland, G. K. (2018). A comparison of 10 accelerometer non-wear time criteria and logbooks in children. *BMC Public Health*, *18*(1), 323. https://doi.org/10.1186/s12889-018-5212-4.

Aadland, E., Andersen, L. B., Anderssen, S. A., Resaland, G. K., & Kvalheim, O. M. (2020). Accelerometer epoch setting is decisive for associations between physical activity and metabolic health in children. *Journal of Sports Sciences*, *38*(3), 256–263. https://doi.org/10.1080/02640414.2019.1693320.

ActiGraph (2020). What is the difference between the Wear Time Validation algorithms? https://actigraphcorp.my.site.com/support/s/article/What-is-the-difference-between-the-Wear-Time-Validation-algorithms. Accessed 16.12.2022

Arvidsson, D., Fridolfsson, J., Börjesson, M., Andersen, L. B., Ekblom, Ö., Dencker, M., & Brønd, J. C. (2019). Re-examination of accelerometer data processing and calibration for the assessment of physical activity intensity. *Scandinavian Journal of Medicine & Science in Sports*, *29*(10), 1442–1452. https://doi.org/10.1111/sms.13470.

Bakrania, K., Yates, T., Rowlands, A. V., Esliger, D. W., Bunnewell, S., Sanders, J., Davies, M. J., Khunti, K., & Edwardson, C. L. (2016). Intensity Thresholds on Raw Acceleration Data: Euclidean Norm Minus One (ENMO) and Mean Amplitude Deviation (MAD) Approaches. *PLOS ONE*, *11*(10), e164045. https://doi.org/10.1371/journal.pone.0164045.

Banda, J. A., Haydel, K. F., Davila, T., Desai, M., Bryson, S., Haskell, W. L., Matheson, D., & Robinson, T. N. (2016). Effects of varying epoch lengths, wear time algorithms, and activity cut-points on estimates of child sedentary behavior and physical activity from accelerometer data. *PLOS ONE*, *11*(3), e150534. https://doi.org/10.1371/journal.pone.0150534.

Bornstein, D. B., Beets, M. W., Byun, W., & McIver, K. L. (2011). Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *Journal of Science and Medicine in Sport*, *14*(6), 504–511. https://doi.org/10.1016/j.jsams.2011.05.007.

Breau, B., Coyle-Asbil, H. J., Haines, J., Ma, D. W. L., & Vallis, L. A. (2022). Actigraph cutpoints impact physical activity and sedentary behavior outcomes in young children. *Journal for the Measurement of Physical Behaviour*, *5*(2), 85–96. https://doi.org/10.1123/jmpb.2021-0042.

Brønd, J. C., & Arvidsson, D. (2016). Sampling frequency affects the processing of Actigraph raw acceleration data to activity counts. *Journal of Applied Physiology*, *120*(3), 362–369. https://doi.org/10.1152/japplphysiol.00628.2015.

Brondeel, R., Kestens, Y., Rahimipour Anaraki, J., Stanley, K., Thierry, B., & Fuller, D. (2021). Converting raw accelerometer data to activity counts using open-source code: implementing a MATLAB code in python and R, and comparing the results to Actilife. *Journal for the Measurement of Physical Behaviour*, *4*(3), 205–211. https://doi.org/10.1123/jmpb.2019-0063.

Burchartz, A., Anedda, B., Auerswald, T., Giurgiu, M., Hill, H., Ketelhut, S., Kolb, S., Mall, C., Manz, K., Nigg, C. R., Reichert, M., Sprengeler, O., Wunsch, K., & Matthews, C. E. (2020a). Assessing physical behavior through accelerometry—State of the science, best practices and future directions. *Psychology of Sport and Exercise*, *49*, 101703. https://doi.org/10.1016/j.psychsport.2020.101703.

Burchartz, A., Manz, K., Anedda, B., Niessner, C., Oriwol, D., Schmidt, S. C. E., & Woll, A. (2020b). Measurement of physical activity and sedentary behavior by accelerometry among a nationwide sample from the KiGGS and MoMo study: study protocol. *JMIR Research Protocols*, *9*(7), e14370. https://doi.org/10.2196/14370.

Burchartz, A., Oriwol, D., Kolb, S., Schmidt, S. C. E., Wunsch, K., Manz, K., Niessner, C., & Woll, A. (2021). Comparison of self-reported & device-based, measured physical activity among children in Germany. *BMC Public Health*, *21*(1), 1081. https://doi.org/10.1186/s12889-021-11114-y.

Burchartz, A., Oriwol, D., Kolb, S., Schmidt, S. C. E., von Haaren-Mack, B., Niessner, C., & Woll, A. (2022). Impact of weekdays versus weekend days on accelerometer measured physical behavior among children and adolescents: results from the MoMo study. *German Journal of Exercise and Sport Research*, *52*(2), 218–227. https://doi.org/10.1007/s12662-022-00811-4.

Cain, K. L., Sallis, J. F., Conway, T. L., van Dyck, D., & Calhoon, L. (2013). Using accelerometers in youth physical activity studies: a review of methods. *Journal of Physical Activity & Health*, *10*(3), 437–450. https://doi.org/10.1123/jpah.10.3.437.

Chinapaw, M. J. M., de Niet, M., Verloigne, M., de Bourdeaudhuij, I., Brug, J., & Altenburg, T. M. (2014). From sedentary time to sedentary patterns: Accelerometer data reduction decisions in youth. *PLOS ONE*, *9*(11), e111205. https://doi.org/10.1371/journal.pone.0111205.

Choi, L., Liu, Z., Matthews, C. E., & Buchowski, M. S. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise*, *43*(2), 357–364. https://doi.org/10.1249/MSS.0b013e3181ed61a3.

Clevenger, K. A., Mackintosh, K. A., McNarry, M. A., Pfeiffer, K. A., Montoye, A. H. K., & Brønd, J. C. (2022). Comparison of child and adolescent physical activity levels from open-source versus Actigraph counts. *Journal for the Measurement of Physical Behaviour*, *5*(2), 120–128. https://doi.org/10.1123/jmpb.2021-0057.

Colley, R., Gorber, C. S., & Tremblay, M. S. (2010). Quality control and data reduction procedures for accelerometry-derived measures of physical activity. *Health Reports*, *21*(1), 63–69.

Crane, J. R., Naylor, P.-J., & Temple, V. A. (2018). The physical activity and sedentary behaviour patterns of children in kindergarten and grade 2. *Children*. https://doi.org/10.3390/children5100131.

Dencker, M., Svensson, J., El-Naaman, B., Bugge, A., & Andersen, L. B. (2012). Importance of epoch length and registration time on accelerometer measurements in younger children. *The Journal of Sports Medicine and Physical Fitness*, *52*(2), 115–121.

Donaldson, S. C., Montoye, A. H. K., Tuttle, M. S., & Kaminsky, L. A. (2016). Variability of objectively measured sedentary behavior. *Medicine and Science in Sports and Exercise*, *48*(4), 755–761. https://doi.org/10.1249/MSS.0000000000000828.

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196.

Giurgiu, M., Kolb, S., Nigg, C., Burchartz, A., Timm, I., Becker, M., Rulf, E., Doster, A.-K., Koch, E., Bussmann, J. B., Nigg, C., Ebner-Priemer, U. W., & Woll, A. (2022). Assessment of 24-hour physical behaviour in children and adolescents via wearables: A systematic review of free-living validation studies. *BMJ Open Sport & Exercise Medicine*, *8*(2), e1267. https://doi.org/10.1136/bmjsem-2021-001267.

Hoffmann, R., Lange, M., Butschalowsky, H., Houben, R., Allen, J., Kuhnert, R., Schaffrath, R. A., & Gößwald, A. (2018). *Querschnitterhebung von KiGGS Welle 2: Teilnehmendengewinnung, Response und Repräsentativität*. https://doi.org: Robert-Koch-Institut. https://doi.org/10.17886/RKI-GBE-2018-015. Cross-sectional survey of KiGGS Welle 2—participant recruitment, response and representativeness

Leppänen, M. H., Migueles, J. H., Abdollahi, A. M., Engberg, E., Ortega, F. B., & Roos, E. (2022). Comparing estimates of physical activity in children across different cut-points and the associations with weight status. *Scandinavian Journal of Medicine & Science in Sports*, *32*(6), 971–983. https://doi.org/10.1111/sms.14147.

Logan, G. R. M., Duncan, S., Harris, N. K., Hinckson, E. A., & Schofield, G. (2016). Adolescent physical activity levels: discrepancies with accelerometer data analysis. *Journal of Sports Sciences*, *34*(21), 2047–2053. https://doi.org/10.1080/02640414.2016.1150599.

Mâsse, L. C., Fuemmeler, B. F., Anderson, C. B., Matthews, C. E., Trost, S. G., Catellier, D. J., & Treuth, M. (2005). Accelerometer data reduction: A comparison of four reduction algorithms on select outcome variables. *Medicine and Science in Sports and Exercise*, *37*(11 Suppl), S544–S554. https://doi.org/10.1249/01.mss.0000185674.09066.8a.

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and Science in Sports and Exercise*, *34*(8), 1376–1381. https://doi.org/10.1097/00005768-200208000-00021.

Mauz, E., Lange, M., Houben, R., Hoffmann, R., Allen, J., Gößwald, A., Hölling, H., Lampert, T., Lange, C., Poethko-Müller, C., Richter, A., Rosario, A. S., von Schenck, U., Ziese, T., & Kurth, B.-M. (2019). Cohort profile: Kiggs cohort longitudinal study on the health of children, adolescents and young adults in Germany. *International Journal of Epidemiology*, *49*(2), 375–375k. https://doi.org/10.1093/ije/dyz231.

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & 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*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0.

Migueles, J. H., Cadenas-Sanchez, C., Tudor-Locke, C., Löf, M., Esteban-Cornejo, I., Molina-Garcia, P., Mora-Gonzalez, J., Rodriguez-Ayllon, M., Garcia-Marmol, E., Ekelund, U., & Ortega, F. B. (2019). Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. *Scandinavian Journal of Medicine & Science in Sports*, *29*(4), 566–574. https://doi.org/10.1111/sms.13356.

Moura, D. I. R., Barbosa, A. O., Da Silva, I. C. M., Romanzini, M., Filho, P. A., & de Júnior, F. J. C. (2019). Impact of cutoff points on adolescent sedentary behavior measured by accelerometer. *Revista Brasileira De Atividade Física & Saúde*, *24*, 1–7. https://doi.org/10.12820/rbafs.24e0071.

Neishabouri, A., Nguyen, J., Samuelsson, J., Guthrie, T., Biggs, M., Wyatt, J., Cross, D., Karas, M., Migueles, J. H., Khan, S., & Guo, C. C. (2022). Quantification of acceleration as activity counts in ActiGraph wearable. *Scientific Reports*, *12*(1), 11958. https://doi.org/10.1038/s41598-022-16003-x.

Pfeiffer, K. A., Clevenger, K. A., Kaplan, A., van Camp, C. A., Strath, S. J., & Montoye, A. H. K. (2022). Accessibility and use of novel methods for predicting physical activity and energy expenditure using accelerometry: a scoping review. *Physiological Measurement*. https://doi.org/10.1088/1361-6579/ac89ca.

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European Journal of Sport Science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614.

Rowlands, A. V., Mirkes, E. M., Yates, T., Clemes, S., Davies, M. J., Khunti, K., & Edwardson, C. L. (2018). Accelerometer-assessed physical activity in epidemiology: are monitors equivalent? *Medicine and Science in Sports and Exercise*, *50*(2), 257–265. https://doi.org/10.1249/mss.0000000000001435.

Sherar, L. B., Griew, P., Esliger, D. W., Cooper, A. R., Ekelund, U., Judge, K., & Riddoch, C. (2011). International children's accelerometry database (ICAD): Design and methods. *BMC Public Health*, *11*, 485. https://doi.org/10.1186/1471-2458-11-485.

Steene-Johannessen, J., Hansen, B. H., Dalene, K. E., Kolle, E., Northstone, K., Møller, N. C., Grøntved, A., Wedderkopp, N., Kriemler, S., Page, A. S., Puder, J. J., Reilly, J. J., Sardinha, L. B., van Sluijs, E. M. F., Andersen, L. B., van der Ploeg, H. P., Ahrens, W., Flexeder, C., Standl, M., & Ekelund, U. (2020). Variations in accelerometry measured physical activity and sedentary time across Europe—harmonized analyses of 47,497 children and adolescents. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 38. https://doi.org/10.1186/s12966-020-00930-x.

Syed, S., Morseth, B., Hopstock, L. A., & Horsch, A. (2020). Evaluating the performance of raw and epoch non-wear algorithms using multiple accelerometers and electrocardiogram recordings. *Scientific Reports*, *10*(1), 5866. https://doi.org/10.1038/s41598-020-62821-2.

Toftager, M., Kristensen, P. L., Oliver, M., Duncan, S., Christiansen, L. B., Boyle, E., Brønd, J. C., & Troelsen, J. (2013). Accelerometer data reduction in adolescents: effects on sample retention and bias. *International Journal of Behavioral Nutrition and Physical Activity*, *10*, 140. https://doi.org/10.1186/1479-5868-10-140.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3.

Trost, S. G. (2007). State of the art reviews: measurement of physical activity in children and adolescents. *American Journal of Lifestyle Medicine*, *1*(4), 299–314. https://doi.org/10.1177/1559827607301686.

Vanhelst, J., Béghin, L., Duhamel, A., Bergman, P., Sjöström, M., & Gottrand, F. (2012). Comparison of uniaxial and triaxial accelerometry in the assessment of physical activity among adolescents under free-living conditions: the HELENA study. *BMC Medical Research Methodology*, *12*(1), 26. https://doi.org/10.1186/1471-2288-12-26.

Vanhelst, J., Vidal, F., Drumez, E., Béghin, L., Baudelet, J.-B., Coopman, S., & Gottrand, F. (2019). Comparison and validation of accelerometer wear time and non-wear time algorithms for assessing physical activity levels in children and adolescents. *BMC Medical Research Methodology*, *19*(1), 72. https://doi.org/10.1186/s12874-019-0712-1.

Woll, A., Albrecht, C., & Worth, A. (2017). Motorik-Module (MoMo)—the KiGGS wave 2 module to survey motor performance and physical activity. *Journal of Health Monitoring*, *2*(S3), 66–73. https://doi.org/10.17886/RKI-GBE-2017-110.

World Health Organization (2010). *Global recommendations on physical activity for health*.

Xing, R., Huang, W. Y., & Sit, C. H.-P. (2021). Validity of accelerometry for predicting physical activity and sedentary time in ambulatory children and young adults with cerebral palsy. *Journal of Exercise Science and Fitness*, *19*(1), 19–24. https://doi.org/10.1016/j.jesf.2020.06.006.