ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/315111025

Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations

Article *in* Sports Medicine · March 2017

DOI: 10.1007/s40279-017-0716-0

CITATIONS
379

READS
813

9 authors, including:

Jairo H Migueles
University of Granada
69 PUBLICATIONS **785** CITATIONS
SEE PROFILE

Cristina Cadenas-Sanchez
Universidad de Cádiz
115 PUBLICATIONS **1,450** CITATIONS
SEE PROFILE

Christine Delisle Nyström
Karolinska Institutet
46 PUBLICATIONS **951** CITATIONS
SEE PROFILE

Jose Mora-Gonzalez
University of North Carolina at Charlotte
68 PUBLICATIONS **993** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

MUBI (MUévete BIen) View project

ActiveBrains View project

All content following this page was uploaded by Cristina Cadenas-Sanchez on 12 March 2020.

The user has requested enhancement of the downloaded file.

Sports Med
DOI 10.1007/s40279-017-0716-0

SYSTEMATIC REVIEW

Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations

Jairo H. Migueles[1] · Cristina Cadenas-Sanchez[1] · Ulf Ekelund[2,3] · Christine Delisle Nyström[4] · Jose Mora-Gonzalez[1] · Marie Löf[4,5] · Idoia Labayen[6] · Jonatan R. Ruiz[1,4] · Francisco B. Ortega[1,4]

© Springer International Publishing Switzerland 2017

Abstract

Background Accelerometers are widely used to measure sedentary time, physical activity, physical activity energy expenditure (PAEE), and sleep-related behaviors, with the ActiGraph being the most frequently used brand by researchers. However, data collection and processing criteria have evolved in a myriad of ways out of the need to answer unique research questions; as a result there is no consensus.

Objectives The purpose of this review was to: (1) compile and classify existing studies assessing sedentary time, physical activity, energy expenditure, or sleep using the ActiGraph GT3X/+ through data collection and processing criteria to improve data comparability and (2) review data collection and processing criteria when using GT3X/+ and provide age-specific practical considerations based on the validation/calibration studies identified.

Methods Two independent researchers conducted the search in PubMed and Web of Science. We included all original studies in which the GT3X/+ was used in laboratory, controlled, or free-living conditions published from 1 January 2010 to the 31 December 2015.

Results The present systematic review provides key information about the following data collection and processing criteria: placement, sampling frequency, filter, epoch length, non-wear-time, what constitutes a valid day and a valid week, cut-points for sedentary time and physical activity intensity classification, and algorithms to estimate PAEE and sleep-related behaviors. The information is organized by age group, since criteria are usually age-specific.

Conclusion This review will help researchers and practitioners to make better decisions before (i.e., device placement and sampling frequency) and after (i.e., data processing criteria) data collection using the GT3X/+ accelerometer, in order to obtain more valid and comparable data.

PROSPERO registration number CRD42016039991.

Electronic supplementary material The online version of this article (doi:10.1007/s40279-017-0716-0) contains supplementary material, which is available to authorized users.

✉ Jairo H. Migueles
jairohm@ugr.es

[1] PROFITH "PROmoting FITness and Health through physical activity" Research Group, Department of Physical Education and Sports, Faculty of Sport Sciences, University of Granada, Ctra. Alfacar s/n, 18011 Granada, Spain

[2] Department of Sport Medicine, Norwegian School of Sport Sciences, Oslo, Norway

[3] MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Addenbrooke's Hospital Hills Road, Cambridge, UK

[4] Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden

[5] Department of Clinical and Experimental Medicine, Faculty of the Health Sciences, Linköping University, Linköping, Sweden

[6] Department of Nutrition and Food Science, University of the Basque Country, UPV-EHU, Vitoria-Gasteiz, Spain

Published online: 16 March 2017

Key Points

Methods regarding data collection and processing criteria when using ActiGraph GT3X are summarized, and age-specific practical considerations that will be useful for researchers and practitioners are provided.

The tabulated data generated will facilitate comparisons between studies using ActiGraph GT3X and aid in the selection of the most appropriate method to use for each specific research purpose.

1 Introduction

Health benefits of physical activity (PA) across a person's lifespan have been widely reported [1–3]. The use of accelerometers to assess sedentary time (SED) and PA [4–7] has become an objective and feasible alternative to self-report methods such as questionnaires, which are characterized by their poor reliability and validity, especially in younger populations [8–10]. Accelerometers are wearable devices that measure accelerations of the body segment to which the monitor is attached. The signal is usually filtered and pre-processed by the monitor to obtain activity counts, i.e., accelerations due to body movement. The amount and intensity of daily SED and PA may be obtained by classifying activity counts accumulated in a specific time interval (epoch length) with a set of cut-points, i.e., intensity thresholds for PA intensity classification [11–15]. Physical activity energy expenditure (PAEE) or sleep-related behaviors may also be estimated by applying algorithms to objectively-determined activity counts [16–21]. New methods to estimate these variables from raw acceleration signals (gravity units) instead of activity counts have been developed recently [22–24].

Among the commercially available brands, the ActiGraph (Pensacola, FL, USA) accelerometers are the most frequently used by researchers, accounting for >50% of published studies [25]. This review only considered the latest generation of ActiGraph devices, i.e., GT3X, GT3X+, and wGT3X-BT (hereinafter referred to as GT3X/+). The continuous change in the features of these devices makes it difficult to compare data between studies.

The first ActiGraph accelerometers available were uniaxial (i.e., they could only detect vertical axis accelerations) and consequently cut-points and algorithms were developed to assess SED, PA intensity, PAEE, and sleep-related behaviors from vertical axis accelerations [11, 17, 21]. In mid-2009, ActiGraph released the triaxial GT3X, which detected accelerations in three axes (i.e., vertical, medio-lateral and antero-posterior axes). The transition from uniaxial to triaxial devices implied new calibration processes, and the algorithms developed for the vertical axis were not applicable to vector magnitude (i.e., the square root of the sum of squared activity counts from the three axes) [7, 13, 18, 20, 26–28].

Due to the extremely rapid development in this field, there is an overwhelming amount of data collection and processing criteria decisions, and there is no consensus about which approaches to use. Consequently, it is difficult for researchers and practitioners to make the right decisions about which criteria should be used in a given situation. This is important as the chosen criteria have a huge impact on the outcome. In order to address this problem, some studies have compared certain GT3X/+ outcomes estimated by different cut-points and algorithms [4, 29–31] in an attempt to recommend which decisions are the most accurate; however, this information is still scarce.

It is important to note that algorithms validated in a specific age group might not be valid for other age groups due to different PA patterns, so whenever possible, data collection and processing criteria should be age-specific. Accelerometer methods can be grouped into two categories: (1) data collection protocols, which are decisions that need to be made a priori such as device placement or sampling frequency; and (2) data processing criteria, which involve decisions that can be made a posteriori such as filters, epoch length, non-wear-time definition, cut-points, and algorithms. The present review will address all of these criteria separately and specifically by age group. In this review we aimed to: (1) compile and classify existing studies assessing sedentary time, physical activity, energy expenditure, or sleep using the ActiGraph GT3X/+ by data collection and processing criteria to improve data comparability, and (2) review data collection and processing criteria when using GT3X/+ and provide age-specific practical considerations based on the validation/calibration studies identified. Both objectives were approached separately for the following age groups: preschoolers, children/adolescents, adults, and older adults. Although there is a large amount of information included in this review, we believe that it is useful for readers to have a single article that summarizes the most important accelerometer methods for each age group separately. This will allow readers to go directly to a specific criteria for the age group they are interested in (e.g., PAEE in preschoolers). In this review, we provide a section with examples of how the information presented can be used in practical terms, as well as a table with practical considerations.

2 Methods

2.1 Study Design

The present review focuses on 11 key methodological issues related to GT3X/+ data collection and processing criteria: (1) device placement, (2) sampling frequency, (3) filter, (4) epoch length, (5) non-wear-time definition, (6) what constitutes a valid day and a valid week, (7) registration period protocol, (8) SED and PA intensity classification, (9) PAEE algorithms, (10) sleep algorithms, and (11) step counting. Available information was classified into two different types of studies: (1) any cross-sectional, longitudinal, or intervention study which used the GT3X/+ device and met the inclusion criteria indicated in Sect. 2.3 (objective 1); and (2) studies focused on validation, calibration, or comparison of functions related to data collection or processing criteria (objective 2). Therefore, the practical considerations provided for each age group are based on the results from the validation/calibration studies (see Table 1). Furthermore, we provide a summary of all data extracted from the validation/calibration papers included in this review by age group in the Electronic Supplementary Material Appendix S1. Inclusion/exclusion criteria and analytical methods were specified in advance and registered in the PROSPERO (http://www.crd.york.ac.

Table 1 Summary of practical considerations by age group

Age group/criterion	Preschoolers	Children and adolescents	Adults	Older adults
Placement	Hip[a] and wrist	Hip[a] and wrist	Hip[a] and wrist	Hip[a] and wrist
Sampling frequency	90–100 Hz	90–100 Hz	90–100 Hz	90–100 Hz
Filter[b]	Normal	Normal	Normal	Low-frequency extension
Epoch length[b]	1–15 s	1–15 s	60 s[c]	60 s[c]
Non-wear-time definition	Not clear[c]	Not clear[c]	Not clear[c]	Choi et al. [45] algorithm
Valid day[d]	≥10 h	≥10 h	≥10 h	≥10 h
Valid week	≥4 days	≥4 days	≥4 days	≥4 days
Registration period protocol	24 h	24 h	24 h	24 h
SED/PA intensity classification[e,f]				
Dominant wrist	No data found	Crouter et al. [37]	Staudenmayer et al. [24]	No data found
Non-dominant wrist	Johansson et al. [34] (2–3 y)	Hildebrand et al. [22] Chandler et al. [15]	Hildebrand et al. [22]	No data found
Hip	Costa et al. [35] (2–3 y) Jimmy et al. [57] (4–6 y)	Hänggi et al. [13] (7–11 years) Romanzini et al. [38] (12–19 years)	Sasaki et al. [7]	Aguilar–Farias et al. [26] Santos–Lozano et al. [27]
PAEE algorithm[e]				
Non-dominant wrist	No data found	Hildebrand et al. [22]	Ellis et al. [42]	No data found
Hip	Butte et al. [71] (2–3 years)	Crouter et al. [20] (7–11 years)	Hildebrand et al. [22]	Santos–Lozano et al. [27]
Sleep algorithm[e]	No data found	Sadeh et al. [16]	Sadeh et al. [16] (20–30 years) Cole et al. [17] (>30 years)	Cole et al. [17]

Hz hertz, *SED* sedentary time, *PA* physical activity, *PAEE* physical activity energy expenditure

Note These recommendations should be considered with caution. We strongly recommend reading Sect. 4 for an understanding of the specific considerations for each age group

[a] There are no algorithms currently available to estimate sleep-related behaviors from data obtained from hip-worn devices

[b] Criterion that could highly affect the output. In these cases, when estimations of PA, PAEE or sleep are the variables of interest, the same criterion as selected in the validation study is recommended. If acceleration metrics are the variables of interest (e.g., counts), the recommendation is provided in this table

[c] There is a need to thoroughly test this criterion

[d] Number of hours per awake period in a day

[e] Citation of the study in which the cut-points or algorithms were validated

[f] Cut-points are specified in Table 5

uk/PROSPERO/) international database of systematic reviews (CRD42016039991) [32]. The study is conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [33].

2.2 Search Strategy

We searched PubMed and Web of Science for studies using the ActiGraph GT3X/+ model and classified the studies into the following age groups: preschoolers (2–5 years), children (6–11 years), adolescents (12–18 years), adults (19–59 years), and older adults (≥60 years). We combined (using the Boolean operator "OR") the following search terms: GT3X, GT3X+, and ActiGraph. Although we wanted to limit the search to GT3X/+, the word ActiGraph was entered in the search because we found that some studies specified the brand (i.e., ActiGraph) instead of the model (i.e., GT3X/+) in the title/abstract/keywords. Since the GT3X/+ models were launched in mid-2009, we limited the dates of the search to 1 January 2010 to the 31 December 2015 and conducted the final search on 3 January 2016. We contacted authors of those studies where the data processing and collection information was unavailable in the published article. In a final step, we extended the search to the IEEE (Institute of Electrical and Electronics Engineers) Xplore database, in case we had missed any relevant studies.

2.3 Inclusion/Exclusion Criteria

We included all original studies (cross-sectional, longitudinal, or intervention studies) in which the GT3X/+ was used in a laboratory, or under controlled or free-living conditions. Protocol studies, reviews, editorials, and abstract or congress communications were excluded, as well as studies conducted in people with mobility problems or in periods of life in which their mobility could have been markedly altered (e.g., pregnancy).

Two authors working independently (JHM and CCS) read the articles and checked whether they met the inclusion/exclusion criteria. They obtained 76% agreement on the papers selected for the review before consensus and 100% agreement after discrepancies were resolved in a consensus meeting. Risk of bias assessment was also conducted independently by JHM and CCS in order to assess the quality of studies (see Electronic Supplementary Material Appendix S2).

3 Results

A total of 940 articles were identified (Fig. 1), of which 444 were excluded after reading the title and abstract and 261 articles were additionally excluded after reading the full text and did not meet the inclusion/exclusion criteria stated above. Finally, a total of 235 studies were considered eligible for the current systematic review. Of them, 78 were validation/calibration studies. Methods and results of these validation/calibration studies are summarized in the Electronic Supplementary Material Appendix S1. Detailed information about the methods and results for the rest of studies (i.e., those using GT3X/+ that were not validation/calibration studies) included in this review is available upon request.

Forty-four percent ($N = 103$) of the included studies were conducted in adults (46% validation/calibration studies), 34% ($N = 81$) in youth (30% validation/calibration studies), 22% ($N = 51$) in older adults (11% validation/calibration studies), and 10% ($N = 24$) in preschoolers (13% validation/calibration studies).

Studies including two or more age groups are summarized in both age group sections in this review. Table 2 presents the criteria used for data collection and processing by age group. A list of references for each of the criteria is found in Electronic Supplementary Material Appendix S3. The information provided in Table 2 and Electronic Supplementary Material Appendix S3 allows researchers to make comparisons between studies that have used the same data collection and processing criteria.

Figure 2 shows the percentage of studies that did not report key methodological issues by age group. Of the studies reviewed, 15–20% did not report criteria such as sampling frequency, epoch length, and a non-wear-time definition, and 60–80% of studies did not report information on the filter used.

Table 3 presents the studies that compared the differences in several outcomes when the GT3X/+ device was simultaneously worn on the wrist and hip. The optimal place to attach the GT3X/+ should be chosen based on reliability, validity, and compliance. Table 4 shows the references for the studies sorted by age group and placement site that have developed SED and PA cut-points, PAEE prediction equations, and sleep algorithms. Table 5 shows the intensity cut-points used in the included studies together with the pre-processing criteria used in the study which developed each set of cut-points. Therefore, the practical considerations provided for each age group are based on the results from the validation/calibration studies (see Table 1).

In the following sub-sections, we will focus only on information from validation/calibration studies presented in Electronic Supplementary Material Appendix S1. Sections 3.1, 3.2, and 3.3 correspond to data collection protocols (i.e., pre-processing stage) and Sects. 3.4–3.10 correspond to processing criteria (i.e., processing stage).

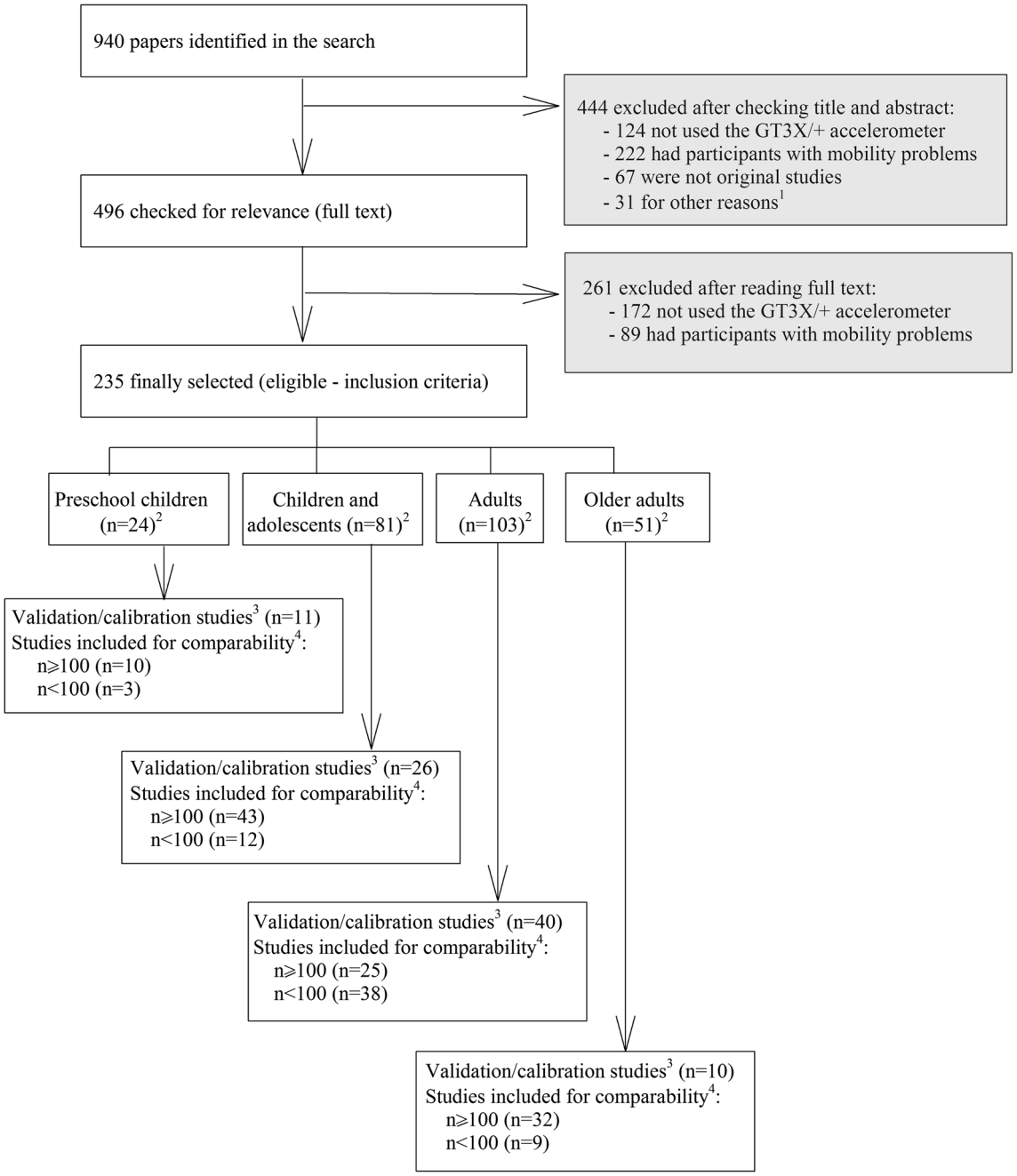

Fig. 1 Flowchart of the literature search and study selection process. [1]Studies using accelerometers for other purposes (e.g., accelerometers attached to dogs). [2]Studies that included two age ranges were counted in both age groups. [3]Studies focused on validation, calibration or comparison of functions related to data collection or processing criteria. [4]All cross-sectional, longitudinal, or intervention studies, which used the GT3X/+ device and met the inclusion criteria

3.1 Device Placement

3.1.1 Preschoolers

In young preschoolers Johansson et al. [34] reported receiver operating characteristic area under the curve (ROC-AUC) data for intensity thresholds between 0.88 to 0.98 using a left wrist-mounted GT3X+. Similarly, a ROC-AUC of 0.90–0.94 was reported by Costa et al. [35] using a hip placement, suggesting high potential for both placements to correctly classify PA intensity in preschoolers.

3.1.2 Children and Adolescents

A higher compliance for wrist-worn versus hip-worn devices has been reported in children/adolescents [23].

Table 2 Summary for the criteria used for data collection protocols and data processing from articles reviewed by age group (see Electronic Supplementary Material Appendix S1 for the criteria used by each of the studies listed in this table)

Reference	Preschoolers ($n = 24$) n (%)	Children and adolescents ($n = 81$) n (%)	Adults ($n = 103$) n (%)	Older adults ($n = 51$) n (%)
Placement				
Hip	22 (92)	73 (90)	87 (84)	44 (86)
Non-dominant wrist	2 (8)	6 (7)	8 (8)	5 (10)
Dominant wrist	0 (0)	1 (1)	6 (6)	5 (10)
Others[a]	0 (0)	2 (2)	21 (20)	2 (4)
Not reported	0 (0)	5 (6)	6 (6)	0 (0)
Sampling frequency				
30 Hz	16 (67)	53 (65)	70 (68)	39 (76)
40 Hz	0 (0)	0 (0)	2 (2)	0 (0)
50 Hz	0 (0)	0 (0)	2 (2)	0 (0)
60 Hz	2 (8)	2 (2)	6 (6)	2 (4)
70 Hz	0 (0)	0 (0)	1 (1)	0 (0)
80 Hz	1 (4)	6 (7)	9 (9)	5 (10)
90 Hz	0 (0)	0 (0)	2 (2)	0 (0)
100 Hz	1 (4)	4 (5)	6 (6)	1 (2)
Not reported	4 (17)	18 (23)	15 (15)	5 (10)
Filter				
Normal	8 (34)	14 (17)	25 (24)	6 (12)
Low-frequency extension	2 (8)	11 (14)	15 (15)	6 (12)
Not reported	14 (58)	53 (65)	67 (65)	40 (80)
Epoch length				
1 s	1 (4)	8 (10)	15 (15)	6 (12)
2 s	0 (0)	1 (1)	2 (2)	0 (0)
3 s	0 (0)	1 (1)	0 (0)	0 (0)
5 s	6 (25)	6 (7)	1 (1)	0 (0)
10 s	0 (0)	16 (20)	6 (6)	1 (2)
15 s	13 (54)	28 (35)	8 (8)	3 (6)
20 s	0 (0)	0 (0)	0 (0)	1 (2)
30 s	1 (4)	3 (4)	0 (0)	0 (0)
45 s	0 (0)	1 (1)	0 (0)	0 (0)
60 s	6 (25)	17 (21)	52 (50)	38 (74)
Not reported	0 (0)	5 (6)	16 (16)	3 (6)
Non-wear-time definition[b]				
10-0-0	3 (13)	6 (7)	7 (7)	4 (8)
20-0-0	3 (13)	21 (26)	3 (3)	0 (0)
20-0-2	0 (0)	0 (0)	1 (1)	0 (0)
30-0-0	2 (8)	5 (6)	1 (1)	0 (0)
30-0-1	0 (0)	3 (4)	0 (0)	0 (0)
60-0-0	3 (13)	5 (6)	16 (16)	3 (6)
60-0-2	0 (0)	7 (9)	15 (15)	12 (24)
60-30-2	0 (0)	0 (0)	0 (0)	1 (2)
90-0-0	0 (0)	1 (1)	5 (5)	1 (2)
90-0-2	0 (0)	0 (0)	3 (3)	2 (4)
90-30-2	0 (0)	0 (0)	2 (2)	14 (27)
120-0-0	0 (0)	0 (0)	1 (1)	0 (0)
180-0-0	0 (0)	0 (0)	0 (0)	1 (2)
Not reported	6 (25)	7 (9)	15 (15)	6 (12)

Table 2 continued

Reference	Preschoolers ($n = 24$) n (%)	Children and adolescents ($n = 81$) n (%)	Adults ($n = 103$) n (%)	Older adults ($n = 51$) n (%)
Cut-points for sedentary time (cpm and vector used) [original reference]				
25 cpm VA [26]	0 (0)	0 (0)	1 (1)	1 (2)
50 cpm VA [79]	0 (0)	0 (0)	1 (1)	0 (0)
60 cpm VA [35]	1 (4)	0 (0)	0 (0)	0 (0)
100 cpm VA [11, 12, 14, 40, 49, 61, 95–98]	4 (17)	31 (38)	40 (39)	31 (61)
100 cpm VM	0 (0)	1 (1)	1 (1)	0 (0)
120 cpm VM [13]	0 (0)	2 (2)	0 (0)	0 (0)
148 cpm VA [67]	3 (13)	1 (1)	0 (0)	0 (0)
150 cpm VA [74]	0 (0)	2 (2)	5 (5)	1 (2)
150 cpm VM [79]	0 (0)	1 (1)	2 (2)	1 (2)
184 cpm VA [38]	0 (0)	1 (1)	0 (0)	0 (0)
200 cpm VA [79, 99]	0 (0)	0 (0)	1 (1)	1 (2)
200 cpm VM [26]	0 (0)	0 (0)	1 (1)	2 (4)
240 cpm VA [71]	1 (4)	0 (0)	0 (0)	0 (0)
250 cpm VA [79]	0 (0)	0 (0)	1 (1)	0 (0)
274 cpm VA [66]	2 (8)	0 (0)	0 (0)	0 (0)
384 cpm VM [35]	1 (4)	0 (0)	0 (0)	0 (0)
500 cpm VA [100]	0 (0)	1 (1)	0 (0)	0 (0)
720 cpm VM [38]	0 (0)	1 (1)	0 (0)	0 (0)
796 cpm VA [68]	2 (8)	1 (1)	0 (0)	0 (0)
820 cpm VM [71]	1 (4)	0 (0)	0 (0)	0 (0)
1068 cpm VA [34]	1 (4)	0 (0)	0 (0)	0 (0)
1204 cpm VA [65]	2 (8)	0 (0)	0 (0)	0 (0)
1260 cpm VM [37]	0 (0)	1 (1)	0 (0)	0 (0)
1452 cpm VA [65]	2 (8)	0 (0)	0 (0)	0 (0)
1488 cpm VA [101]	2 (8)	0 (0)	0 (0)	0 (0)
1592 cpm VA [65]	3 (13)	0 (0)	0 (0)	0 (0)
1932 cpm VM [15]	0 (0)	1 (1)	0 (0)	0 (0)
2652 cpm VM [34]	2 (8)	0 (0)	0 (0)	0 (0)
3300 cpm VM [37]	0 (0)	1 (1)	0 (0)	0 (0)
3660 cpm VM [15]	0 (0)	1 (1)	0 (0)	0 (0)
Cut-points for physical activity intensity classification [original reference]				
Aguilar-Farías et al. [26]	0 (0)	0 (0)	1 (1)	3 (6)
Aittasalo et al. [78]	0 (0)	1 (1)	0 (0)	0 (0)
Andersen et al. [100]	0 (0)	1 (1)	0 (0)	0 (0)
Butte et al. [71]	1 (4)	0 (0)	0 (0)	0 (0)
Chandler et al. [15]	0 (0)	1 (1)	0 (0)	0 (0)
Copeland et al. [14]	0 (0)	0 (0)	0 (0)	9 (18)
Costa et al. [35]	1 (4)	0 (0)	0 (0)	0 (0)
Crouter et al. [37]	0 (0)	1 (1)	0 (0)	0 (0)
Davis et al. [99]	0 (0)	0 (0)	0 (0)	1 (2)
Evenson et al. [12]	8 (34)	36 (45)	0 (0)	0 (0)
Freedson et al. [11]	0 (0)	1 (1)	30 (29)	14 (27)
Freedson et al. [74]	0 (0)	8 (10)	0 (0)	0 (0)
Grydeland et al. [102]	0 (0)	1 (1)	0 (0)	0 (0)
Hänggi et al. [13]	0 (0)	2 (2)	0 (0)	0 (0)
Hildebrand et al. [22]	0 (0)	1 (1)	0 (0)	0 (0)

Table 2 continued

Reference	Preschoolers ($n = 24$) n (%)	Children and adolescents ($n = 81$) n (%)	Adults ($n = 103$) n (%)	Older adults ($n = 51$) n (%)
Jimmy et al. [57]	1 (4)	1 (1)	0 (0)	0 (0)
Johansson et al. [34]	2 (8)	0 (0)	0 (0)	0 (0)
Matthews et al. [103]	0 (0)	1 (1)	6 (6)	8 (16)
Mattocks et al. [75]	0 (0)	1 (1)	0 (0)	0 (0)
Metzger et al. [98]	0 (0)	0 (0)	3 (3)	1 (2)
Pate et al. [67]	4 (17)	1 (1)	0 (0)	0 (0)
Pruitt et al. [104]	0 (0)	0 (0)	0 (0)	1 (2)
Pulakka et al. [70]	2 (8)	0 (0)	0 (0)	0 (0)
Puyau et al. [68]	2 (8)	2 (2)	0 (0)	0 (0)
Reilly et al. [66]	3 (13)	0 (0)	0 (0)	0 (0)
Romanzini et al. [38]	0 (0)	1 (1)	0 (0)	0 (0)
Santos-Lozano et al. [27]	0 (0)	1 (1)	1 (1)	1 (2)
Sasaki et al. [7]	0 (0)	0 (0)	6 (6)	2 (4)
Sirard et al. [65]	3 (13)	0 (0)	0 (0)	0 (0)
Treuth et al. [95]	0 (0)	4 (5)	0 (0)	0 (0)
Troiano et al. [61]	0 (0)	0 (0)	8 (8)	4 (8)
Trost et al. [21]	0 (0)	2 (2)	0 (0)	0 (0)
Vähä-Ypyä et al. [80]	0 (0)	0 (0)	1 (1)	0 (0)
Van Cauwenberghe et al. [101]	3 (13)	0 (0)	0 (0)	0 (0)
Vanhelst et al. [73]	0 (0)	1 (1)	0 (0)	0 (0)
Zhu et al. [72]	0 (0)	2 (2)	0 (0)	0 (0)
Zisko et al. [105]	0 (0)	0 (0)	0 (0)	1 (2)
Physical activity energy expenditure algorithms [original reference]				
Butte et al. [71]	1 (1)	0 (0)	0 (0)	0 (0)
Crouter et al. [106]	0 (0)	1 (1)	0 (0)	0 (0)
Crouter et al. [107]	0 (0)	0 (0)	2 (2)	0 (0)
Crouter et al. [20]	0 (0)	3 (4)	0 (0)	0 (0)
Ellis et al. [28]	0 (0)	0 (0)	1 (1)	0 (0)
Evenson et al. [12]	0 (0)	1 (1)	0 (0)	0 (0)
Freedson et al. [74]	0 (0)	3 (4)	0 (0)	0 (0)
Hildebrand et al. [22]	0 (0)	1 (1)	0 (0)	0 (0)
Liu et al. [108]	0 (0)	1 (1)	0 (0)	0 (0)
Mattocks et al. [75]	0 (0)	1 (1)	0 (0)	0 (0)
Pate et al. [67]	1 (4)	0 (0)	0 (0)	0 (0)
Puyau et al. [68]	1 (4)	3 (4)	0 (0)	0 (0)
Santos-Lozano et al. [27]	0 (0)	1 (1)	1 (1)	1 (2)
Schmitz et al. [109]	0 (0)	1 (1)	0 (0)	0 (0)
Stec et al. [43]	0 (0)	0 (0)	1 (1)	0 (0)
Treuth et al. [95]	0 (0)	3 (4)	0 (0)	0 (0)
Trost et al. [21]	0 (0)	3 (4)	0 (0)	0 (0)
WET	0 (0)	1 (1)	1 (1)	1 (2)
WET + Freedson et al. [11]	0 (0)	1 (1)	3 (3)	2 (4)
WET + Sasaki et al. [7]	0 (0)	1 (1)	1 (1)	1 (2)
Zakeri et al. [81]	1 (4)	0 (0)	0 (0)	0 (0)
Zhu et al. [82]	0 (0)	1 (1)	0 (0)	0 (0)
Sleep algorithm [original reference]				
Barreira et al. [19]	0 (0)	1 (1)	0 (0)	0 (0)
Cole-Kripke et al. [17]	0 (0)	1 (1)	1 (1)	3 (6)

Table 2 continued

Reference	Preschoolers ($n = 24$) n (%)	Children and adolescents ($n = 81$) n (%)	Adults ($n = 103$) n (%)	Older adults ($n = 51$) n (%)
Sadeh et al. [16]	1 (4)	2 (2)	4 (4)	0 (0)
Tudor-Locke et al. [18]	0 (0)	4 (5)	0 (0)	0 (0)

cpm counts per minute, *Hz* hertz, *s* seconds, *VA* vertical axis, *VM* vector magnitude, *WET* work energy theory

Note 1 Studies using several criteria have been considered in each criterion, thus, in these cases percentages do not have to sum to 100%

Note 2 Criteria used for the data collection and processing validated with other devices but have been applied to GT3X/+ data have been also considered

[a] Other placements used for different aims from physical activity intensity classification, physical activity energy expenditure, or sleep estimation (e.g., physical activity type identification, light sensor validation, etc.)

[b] Non-wear-time definition expressed as: minimum minutes of 0 cpm—minimum minutes for before and after allowance windows—maximum minutes of allowance

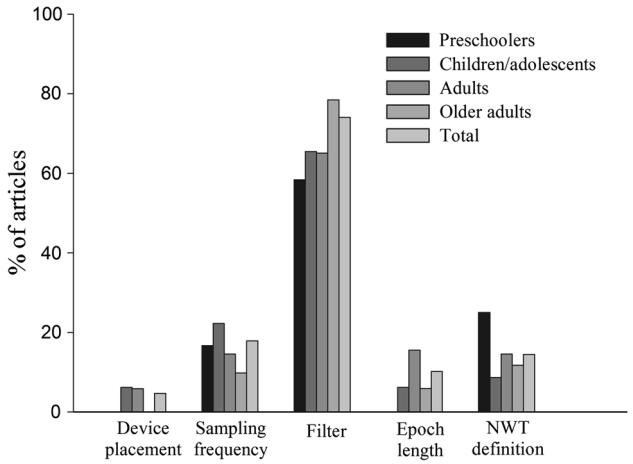

Fig. 2 Percentage of the 235 included papers that did not report key methodological issues, separated by age group. *NWT* non-wear-time

However, similar wear-time was achieved in protocols using 24-h waist-worn compared to 24-h wrist-worn accelerometers [36].

With regard to cut-points to classify SED and PA intensity, non-dominant wrist placement achieved a lower ROC-AUC (0.64–0.89) [15] compared to the dominant wrist (0.83–0.94) [37] and hip (>0.90) for all cut-points [13, 38]. Furthermore, Hildebrand et al. [22] found a greater percentage of the explained variance when using algorithms from the hip compared to the wrist (78% for hip; 71% for wrist).

Previously developed sleep algorithms for the wrist placement were tested on the hip and wrist by Hjorth et al. [39]. They obtained a classification accuracy between 86.6 and 89.9% for the algorithms tested (developed with GT1M) [16, 17] in hip compared to wrist measurement. Hip placement overestimated total sleep time compared to the wrist (60.1 vs. 73.8 min per day for wrist and hip, respectively). Finally, Tudor-Locke et al. [18] developed an algorithm to identify bedtime for the hip-worn GT3X/+, and Barreira et al. [19] refined and validated these in a free-living environment against self-reported participant logs, where they obtained a non-significant absolute difference of 9 ± 36 min.

3.1.3 Adults

Minimal differences between contralateral hips were found for vector magnitude activity counts (effect size: 0.016, $p = 0.619$) and wear-time (effect size: 0.040, $p = 0.213$) [40]. The reliability of the GT3X/+ attached to the hip, wrist, and ankle was studied by Ozemek et al. [41], who found high correlations from 0.824 to 0.998 in vector magnitude between pairs of devices under simulated activities of daily living.

Staudenmayer et al. [24] demonstrated greater accuracy for physical activity classification when the device was placed on the wrist compared to previously developed cut-points with the accelerometer placed on the hip. Additionally, they found that newly developed algorithms could also categorize behaviors in a laboratory setting (e.g., sitting, standing, riding in a vehicle, walking, and running) better for the wrist compared to the hip placement. Ellis et al. [28] achieved better performance with a wrist model to predict household simulated activities; however, the hip model outperformed the wrist model on locomotion prediction (i.e., slow walk, brisk walk, and jogging) as well as PAEE estimation. In contrast, Ellis et al. [42] and Hildebrand et al. [22] obtained a higher accuracy (5% more on average) and a larger explained variance (81% for hip vs. 75% for wrist), respectively, for the hip compared to the wrist placement to classify PA type and intensity. Stec et al. [43] found a significant correlation between vector magnitude activity counts and net energy expenditure from hip but not from wrist worn accelerometers during resistance exercise (Pearson correlations for hip = 0.50, $p = 0.005$; and wrist = −0.25, $p = 0.18$).

Table 3 Summary of studies comparing hip- and wrist-worn GT3X/+ accelerometers

References	Age group	Aims (principal outcomes studied in italics)	Main findings/conclusions
Ellis et al. [28]	Adults	To compare GT3X/+ worn on the right hip and the non-dominant wrist, and the added value of heart rate data, for predicting *PA type* and *PAEE estimation*	In estimating PAEE, both device positions produced comparable results. The wrist GT3X/+ was superior predicting activities with significant arm movement, while the hip GT3X/+ was superior for predicting locomotion
Fairclough et al. [23]	Children	To compare right hip and non-dominant wrist *compliance*, and to compare *PA* derived from wrist and hip raw data	Wrist placement was associated with superior compliance compared with the hip. Raw accelerations were significantly higher for the wrist compared with the hip
Hildebrand et al. [22]	Children and adults	To compare raw GT3X/+ output from the right hip and the non-dominant wrist and to develop *PAEE* equations for each placement	The output from the wrist monitor was higher during more intense activities but similar or lower during sedentary activities. Hip PAEE equation showed a higher accuracy
Hjorth et al. [39]	Children	To compare GT3X/+ *sleep scoring* from the right hip and the non-dominant wrist using existing algorithms	Hip-worn and wrist-worn GT3X/+ cannot be used interchangeably for estimating sleep-related behaviors
Ozemek et al. [41]	Adults	To test the reliability of GT3X+ placed on the hip, dominant wrist and ankle in measuring *activity counts* recorded by axis 1, 2, 3 and VM during daily living	GT3X/+ worn on the hip, wrist and ankle showed a high test–retest agreement across all axes and VM. Specifically, lower variability in activity counts was observed in hip placement compared to wrist- or ankle-worn accelerometers
Slater et al. [86]	Adults	To examine the GT3X/+ validity for *sleep scoring* from right hip and left wrist compared to polysomnography using the same algorithm	The wrist-worn GT3X + provided more valid measures of sleep but with only moderate capability to detect periods of wake during the sleep period. With Sadeh's algorithm[16] GTX3+ Actigraph worn on the hip does not provide valid or accurate measures of sleep
Staudenmayer et al. [24]	Adults	To develop algorithms for dominant wrist to estimate: *METs-hours, minutes in PA intensities, minutes in sedentary activities vs. not and minutes in locomotion vs. not*, validate them against indirect calorimetry and compare them against previously developed hip algorithms	The wrist models, applied to 15-s epoch, estimated METs better than a previously developed model that used counts per minute measured at the hip
Stec et al. [43]	Adults	To estimate the optimal placement (right wrist, right hip, or right ankle) to attach the GT3X/+ for *PAEE estimation* during resistance exercise	The hip-worn GT3X obtained better results for estimating PAEE in resistance exercise
Tudor-Locke et al. [44]	Adults	To compare GT3X/+ *step outputs* obtained from right hip and non-dominant wrist	In laboratory conditions, the hip detected more steps than the wrist independently of the filter selected. In free-living, the wrist produced a higher average step counts than the hip. The hip step counts were more accurate than the wrist in controlled conditions

METs metabolic equivalents, *PAEE* energy expenditure, *PA* physical activity, *VM* vector magnitude

With regard to step counting, Tudor-Locke et al. [44] found higher accuracy for step counting from hip compared to wrist devices in controlled conditions against direct observation. No data about placement comparisons were found in adults for sleep-related behavior estimations.

3.1.4 Older Adults

The hip has been the most commonly used placement for studies in older adults. Only one study by Choi et al. [45] placed the GT3X on the dominant wrist to validate their non-wear-time algorithm. However, we found no data comparing different device placement in this age group for any of the accelerometer outcomes.

3.2 Sampling Frequency

Due to an insufficient number of studies this section data from all age groups are combined. GT3X records accelerations at a sampling frequency of 30 Hz However, with the release of GT3X+, the manufacturer allowed users to select the sampling frequency between 30 and 100 Hz Brønd and Arvidsson [46] demonstrated that sampling frequency had an effect on activity counts (i.e., a difference

Table 4 Studies developing cut-points for sedentary time and physical activity intensity classification, PAEE, and sleep algorithms used in the articles reviewed and ordered by age group

Age group	Hip		Dominant wrist		Non-dominant wrist	
	VM	VA	VM	VA	VM	VA
Preschoolers						
Cut-points	[35, 57, 70, 71]	[35, 57, 65–67, 70, 71, 101]	NF	NF	[34][a]	[34][a]
PAEE	[71, 81]	[67]	NF	NF	NF	NF
Sleep	NF	NF	NF	NF	NF	NF
Children and adolescents						
Cut-points	[13, 22, 27, 38, 57, 78]	[12, 38, 57, 61, 68, 72–75, 95–97, 100, 102]	[37]	[37]	[15, 22]	[15]
PAEE	[20, 22, 27, 82]	[12, 20, 21, 68, 74, 75, 95, 106]	NF	NF	[22]	NF
Sleep	NF	[18, 19]	NF	[16]	NF	[16]
Adults						
Cut-points	[7, 27, 80]	[11, 61, 79, 97, 98, 103]	[24][b]	NF	[22]	NF
PAEE	[7, 22, 27, 28, 43]	[11, 27, 107]	NF	NF	[22, 28]	NF
Sleep	NF	NF	NF	[16, 17]	NF	[16, 17]
Older adults						
Cut-points	[26, 27, 105]	[14, 26, 27, 61, 97–99, 104, 105]	NF	NF	NF	NF
PAEE	[27]	NF	NF	NF	NF	NF
Sleep	NF	NF	NF	[17]	NF	[17]

NF not found, *PAEE* physical activity energy expenditure, *VA* vertical axis, *VM* vector magnitude

[a] Cut-points developed on the left wrist

[b] Algorithm developed with machine learning, not usual cut-points

of +90 counts/min (cpm) for a slow walk, +180 cpm for a fast walk, +103 cpm for a slow run, and +1601 cpm for a fast run at a sampling frequency of 40 Hz compared to 30 Hz). Since the filtering process was developed for 30 Hz, sampling frequencies in multiples of 30 produce the most accurate estimates. Particularly, these authors [46] observed that 30, 60, or 90 Hz produced similar activity counts whereas sampling frequencies at 40, 50, 70, 80, or 100 Hz offset the filter resulting in an increased number of activity counts.

3.3 Valid Day and Valid Week

We cannot present the information in this section separately for each age group due to the lack of studies. As Toftager et al. [47] reported, increasing the requirements for what is considered a valid day (i.e., the number of hours per day) and a valid week (i.e., the number of valid days with valid data) led to a decrease in sample size and therefore the study's power.

In the National Health And Nutrition Examination Survey (NHANES) 2003–2006, where participants wore the accelerometers during waking hours, only 40–70% of them achieved a minimum of 10 h/day of wear-time for ≥6 days, while in the NHANES 2011–2012, where participants were asked to wear the accelerometers 24 h/day, the wear-time achieved was 21–22 h/day for ≥6 days [48].

Aadland et al. [49] examined how many days were needed to obtain an intraclass correlation coefficient (ICC) of 0.80 with different hours per day wear-time criteria (≥8 h/day, ≥10 h/day, and ≥12 h/day). ICCs for a single day did not differ much for all variables when the wear-time criteria increased (i.e., ICCs = 0.20–0.53 for ≥8 h/day, ICCs = 0.21–0.53 for ≥10 h/day, ICCs = 0.23–0.52 for ≥12 h/day). The number of days needed for an ICC of 0.80 decreased with a more demanding wear-time criterion [from 8.3 to 6.4 days for SED, from 4.4 to 3.7 days for light PA, and from 8.5 to 7.0 days for moderate-to-vigorous PA (MVPA), all adjusted for wear-time]. Although the registration period is usually 1 week, 2 weeks were analyzed in the aforementioned study. Also, Donaldson et al. [50] reported that 4 days of measurement would be comparable to 1 week for estimating SED ($r^2 = 0.91$).

3.4 Filter

3.4.1 Preschoolers

No data about the influence of the filter selected (i.e., normal vs. LFE) were found in preschoolers.

Table 5 Cut-points for SED, LPA, MPA, VPA, and VVPA activity used in the articles reviewed and ordered by age-group in which they were validated

Age group/reference	Placement	Filter	Epoch	Vector	SED	LPA	MPA	VPA	VVPA
Preschoolers									
Butte et al. [71]	Right hip	Normal	60 s	VA	≤240	2120–4449	NA	NA	≥4450
				VM	≤820	3908–6111	NA	NA	≥6112
Costa et al. [35]	Right hip	LFE	5 s	VA	≤5	NA	≥165	NA	NA
				VM	≤96.12	NA	≥361.94	NA	NA
Jimmy et al. [57]	Right hip	Normal	5 s	VA	NA	NA	≥133	134–193	194–233
				VM	NA	NA	≥246	247–316	317–381
Johansson et al. [34]	Left wrist	Normal	5 s	VA	≤89	90–439	NA	≥440	NA
				VM	≤221	222–729	NA	≥730	NA
Pate et al. [67][a]	Right hip	NR	15 s	VA	≤37	38–419	420–841	≥842	NA
Pulakka et al. [70]	Right hip	Normal	15 s	VA	NA	NA	≥35	NA	NA
				VM	NA	NA	≥208	NA	NA
Reilly et al. [66][a]	Right hip	NR	60 s	VA	≤274	NA	NA	NA	NA
Sirard et al. [65][a]									
3 years old	Right hip	NR	15 s	VA	≤301	302–614	615–1230	≥1231	NA
4 years old					≤363	364–811	812–1234	≥1235	NA
5 years old					≤398	399–890	891–1254	≥1255	NA
Van Cauwenberghe et al. [101][b]	Right hip	NR	15 s	VA	≤372	373–584	585–880	≥881	NA
Children and adolescents									
Aittasalo et al. [78][b]	Hip	NR	Raw: 100 Hz	VM	≤26.9 mg	27–332 mg	332–558 mg	≥558 mg	NA
Andersen et al. [100][a]	Hip	NR	60 s	VA	≤499	500–1999	2000–2999	3000–4499	4500–32,767
Chandler et al. [15]	Non-DW	Normal	5 s	VA	≤161	162–529	530–1461	≥1462	NA
				A2	≤132	133–445	446–998	≥999	NA
				A3	≤113	114–372	373–776	≥777	NA
				VM	≤305	306–817	818–1968	≥1969	NA
Crouter et al. [37] (ROC analysis)	DW	LFE	5 s	VA	≤105	NA	262–564	≥565	NA
				VM	≤275	NA	416–777	≥778	NA
Crouter et al. [37] (regression analysis)	DW	LFE	5 s	VA	≤35	36–360	361–1129	≥1130	NA
				VM	≤100	101–609	610–1809	≥1810	NA
Evenson et al. [12][a]	Right hip	NR	15 s	VA	≤25	26–573	574–1002	≥1003	NA
Freedson et al. [74][a]	Right hip	NR	60 s	VA	≤149	150–499	500–3999	4000–7599	≥7600
Grydeland et al. [102]	Right hip	LFE	60 s	VA	≤100	101–2000	2001–6000	≥6001	NA
Hänggi et al. [13]	Right hip	Normal	1 s	VM	≤2	3–56	≥56	NA	NA
Hildebrand et al. [22][b]	Right hip	NR	Raw: 60 Hz	VM	NA	≤142 mg	142–464 mg	≥464 mg	NA
	Non-DW	NR			NA	≤201 mg	201–707 mg	≥707 mg	NA
Jimmy et al. [57]	Right hip	Normal	5 s	VA	NA	NA	NA	5 METs: 193	NA
								6 METs: 233	
				VM	NA	NA	NA	5 METs: 316	NA
								6 METs: 381	
Matthews et al. [97][a]	Right hip	NR	60 s	VA	≤100	NA	NA	NA	NA

Table 5 continued

Age group/reference	Placement	Filter	Epoch	Vector	SED	LPA	MPA	VPA	VVPA
Mattocks et al. [75][a]	Right hip	NR	60 s	VA	NA	NA	3581–6129	≥6130	NA
Puyau et al. [68][a]	Right hip	NR	15 s	VA	≤199	200–799	800–2049	≥2050	NA
Ridgers et al. [96][b]	Right hip	NR	60 s	VA	≤100	NA	NA	NA	NA
Romanzini et al. [38]	Hip	NR	15 s	VA	≤46	47–606	607–817	≥818	NA
				VM	≤180	181–756	757–1111	1112	NA
Santos-Lozano et al. [27]	Right hip	Normal	60 s	VM	NA	NA	≤2114	2115–6548	≥11,490
Treuth et al. [95][a]	Right hip	NR	30 s	VA	≤50	51–1499	1500–2600	>2600	NA
	Right hip	NR	60 s	VA	≤100	101–2999	3000–5200	>5200	NA
Troiano et al. [61]	Right hip	NR	60 s	VA	≤100	101–2019	2020–5998	≥5999	NA
Vanhelst et al. [73][b]	Hip	NR	60 s	VA	≤400	401–1900	1901–3918	≥3919	NA
Zhu et al. [72]	Right hip	NR	60 s	VA	NA	NA	2860–3839	≥3840	NA
Adults									
Freedson et al. [11][a]	Right hip	NR	60 s	VA	≤99	100–759	760–5724	5725–9498	≥9499
Hildebrand et al. [22][b]	Right hip	NR	Raw: 60 Hz	VM	NA	≤69 mg	69–259 mg	≥260 mg	NA
	Non-DW	NR			NA	≤101 mg	101–429 mg	≥430 mg	NA
Kozey-Keadle et al. [79]	Right hip	LFE	60 s	VA	≤150	NA	NA	NA	NA
Matthews et al. [110][a]	Right hip	NR	60 s	VA	NA	NA	760–5998	≥5999	NA
Matthews et al. [97][a]	Right hip	NR	60 s	VA	≤100	NA	NA	NA	NA
Metzger et al. [98][a]	Right hip	NR	60 s	VA	< 100	100–2019	NA	NA	NA
Santos-Lozano et al. [27]	Right hip	Normal	60 s	VM	NA	NA	≤3208	3209–8565	≥11,593
Sasaki et al. [7]	Right hip	Normal	60 s	VM	NA	≤2690	2691–6166	6167–9642	≥9643
Troiano et al. [61][a]	Right hip	NR	60 s	VA	≤100	101–2019	2020–5998	≥5998	NA
Older adults									
Aguilar-Farias et al. [26]	Right hip	LFE	1 s	VA	< 1	NA	NA	NA	NA
			15 s		≤9	NA	NA	NA	NA
			60 s		≤24	NA	NA	NA	NA
			1 s	VM	< 1	NA	NA	NA	NA
			15 s		≤69	NA	NA	NA	NA
			60 s		≤199	NA	NA	NA	NA
Copeland et al. [14][a]	Right hip	NR	60 s	VA	≤99	100–1040	≥1040	NA	NA
Davis et al. [99][a]	Right hip	NR	60 s	VA	≤199	200–1999	2000–3999	≥4000	NA
Kozey-Keadle et al. [79]	Right hip	LFE	60 s	VA	≤150	NA	NA	NA	NA
Matthews et al. [97][a]	Right hip	NR	60 s	VA	≤100	NA	NA	NA	NA
Metzger et al. [98]	Right hip	LFE	60 s	VA	≤149	150–2019	≥2020	NA	NA
Pruitt et al. [104]	Right hip	NR	60 s	VA	NA	NA	NA	NA	NA
Santos-Lozano et al. [27]	Right hip	Normal	60 s	VM	NA	NA	≤2751	2752–9359	≥9360
Troiano et al. [61][a]	Right hip	NR	60 s	VA	≤100	101–2019	2020–5998	≥5998	NA
Zisko et al. [105]									
Men	Right hip	NR	60 s	VA	≤55	56–266	267–1971	1972–3878	≥3879
				VM	≤610	611–1652	1653–3016	3017–4581	≥4582
Women				VA	≤59	60–212	213–1217	1218–3157	≥3158
				VM	≤464	465–1076	1077–2424	2425–4078	≥4079

A2 axis 2, *A3* axis 3, *DW* dominant wrist, *METs* metabolic equivalents, *LFE* low-frequency extension filter, *LPA* light physical activity, *MPA* moderate physical activity, *NA* not applicable, *NR* not reported, *ROC* receiver operating characteristic, *s* seconds, *SED* sedentary time, *VA* vertical axis, *VM* vector magnitude, *VPA* vigorous physical activity, *VVPA* very vigorous physical activity

[a] Cut-points developed with the 7164 or the GT1M model (ActiGraph, Pensacola, FL, USA)

[b] Cut-points not expressed in counts, but in other units

3.4.2 Children and Adolescents

Hjorth et al. [39] used normal and LFE filters on GT3X+ data from hip-mounted accelerometers during the night. Total activity count was (on average) 2815 counts per night period higher with the LFE filter compared to the normal one. Assuming 8 h of sleep, this means approximately 6 cpm more when the LFE filter is enabled. Therefore, total sleep time was 9 min per night higher with the normal filter compared to the LFE filter when using a hip mounted GT3X+.

3.4.3 Adults

Lyden et al. [51] found the normal filter more accurate compared to the LFE filter when it was used to identify SED and breaks in SED with the GT3X attached to the hip against direct observation. Ried-Larsen et al. [52] and Cain et al. [53] observed less SED and more minutes in each PA intensity with the LFE filter enabled. Non-wear-time estimation was similar between filters in the study published by Cain et al. [53]. For sleep-related behaviors, Cellini et al. [54] found similar results for total sleep time and sleep efficiency with both filters in a short sleeping time of 2 h. For step counting under free-living conditions, the use of LFE filter increased the step count by an average of approximately 6000 steps per day [44].

3.4.4 Older Adults

Wanner et al. [55] observed a mean difference of $+37.8 \pm 19.5$ cpm when enabling the LFE filter compared to the normal filter. Therefore, less time in SED and more time in all PA intensities were observed with the LFE filter due to the influence on activity counts. The normal filter appears to be more accurate than the LFE filter when compared with the NL-100 pedometer (Lee's Summit, MO, USA) [56].

3.5 Epoch Length

3.5.1 Preschoolers

We did not find any information about the influence of epoch length on accelerometer output in preschoolers. However, several studies used a 5-s epoch based on the belief that the activity pattern of very young children is intermittent and therefore shorter epoch lengths might be suitable to capture very short bouts of movement [34, 35, 57, 58].

3.5.2 Children and Adolescents

Aibar et al. [59] compared the effect of different epoch lengths (3–60 s) on PA intensity, and found a progressive decrease in the time spent in MVPA intensity as the epoch length increased. Furthermore, they found that smaller epoch lengths increased the resolution of the measure, thus increasing the time spent in vigorous PA intensity [59]. Therefore, they suggested using shorter epoch lengths (e.g., 3–15 s) in children.

3.5.3 Adults

No information about the influence of epoch length was found for adults.

3.5.4 Older Adults

No information about the influence of epoch length was found for older adults.

3.6 Non-Wear-Time Definition

3.6.1 Preschoolers

No information about non-wear-time definition was found for preschoolers.

3.6.2 Children and Adolescents

Toftager et al. [47] showed that the longer the non-wear-time duration the greater the number of participants that were included in the analyses. Furthermore, as the non-wear-time duration increased the average cpm decreased (e.g., average PA level: 641 cpm with strings of 10 min of consecutive zeros compared to 570 cpm with strings of 90 min of consecutive zeros) [47]. Since Toftager et al. [47] compared different non-wear-time definitions without a criterion method, it is not possible to conclude which algorithm was more valid.

3.6.3 Adults

Peeters et al. [60] compared six different definitions of non-wear-time (i.e., 20, 60, and 90 min with and without allowance of 2 min of small accelerations). It was observed that 20 min of 0 cpm without allowing for interruptions resulted in a lower misclassification (5.9%) and a similar ROC-AUC (0.94) than 60 min (6.7%, ROC-AUC = 0.94) and 90 min (7.4%, ROC-AUC = 0.93) [60]. However, in these conditions, more participants did not meet the non-wear-time criteria (32 out of 34 participants, i.e., 6% sample loss) compared to 60- or 90-min algorithms (33 and 34 out of 34 participants, i.e., 3 and 0% sample loss, respectively).

3.6.4 Older Adults

Keadle et al. [30] compared the Troiano et al. algorithm [61], which uses a minimum of 60 min of 0 cpm with an allowance of 2 min of interruptions, with the Choi et al. algorithm [62], which uses a minimum of 90 min of 0 cpm with the same allowance as the Troiano algorithm plus two 30-min windows of 0 cpm before and after that allowance. They concluded that the algorithm by Choi et al. [62] was the best to identify wear-time compared with diary records of the participants. The same conclusion was obtained in a later study by Choi et al. [45], especially when this algorithm was implemented for wrist-worn accelerometers, because the wrist placement is more sensitive to detect non-wear-time than the hip [45].

3.7 Registration Period Protocol: Waking Hours Versus 24 Hours

Due to an insufficient number of studies this section combines all age groups. Recent large-scale studies such as NHANES (2011–2012) and the International Study of Childhood Obesity, Lifestyle and Environment (ISCOLE 2012–2013) [63] have used a 24-h protocol. Tudor-Locke et al. [36] found higher wear-time compliance with 24-h protocols compared to waking-hour protocols, with this finding being consistent across different countries.

3.8 Sedentary Time (SED) and Physical Activity (PA) Intensity Classification

3.8.1 Preschoolers

Two studies comparing several cut-points developed from the vertical axis accelerations from hip-mounted devices were found [58, 64]. Janssen et al. [64] supported the use of the Evenson et al. [12] SED cut-point due to the higher classification accuracy compared to other cut-points [65–69], and recommended that the Pate et al. [67] cut-points are the best option for MVPA (all of them were developed with former models of ActiGraph). However, Kahan et al. [58] observed, in a small sample size (n = 12), that Sirard et al. [65] cut-points showed the best agreement with direct observation for SED and MVPA compared to other cut-points [12, 67, 69] developed with former models of ActiGraph.

Four studies developed cut-points for SED and PA intensity for vector magnitude counts from accelerometers worn on the hip [35, 57, 70, 71]. Butte et al. [71] developed their cut-points using a 60-s epoch considering energy expenditure cut-points established through smoothing splines and ROC curves. The cut-points developed by Costa et al. [35] used 5- and 15-s epochs and were validated against direct observation. Jimmy et al. [57] developed and validated their cut-points utilizing a 5-s epoch based on indirect calorimetry. Pulakka et al. [70] developed one cut-point to differentiate sedentary/light activities from MVPA (they did not differentiate moderate from vigorous PA, as done in the aforementioned cut-points) and validated it against direct observation. All of these studies obtained high ROC-AUC (0.89–0.98 for all cut-points).

Finally, Johansson et al. [34] was the only study that developed and validated cut-points for vector magnitude from a wrist-worn accelerometer against direct observation in young preschoolers (15–36 months) obtaining a ROC-AUC of 0.89–0.98.

3.8.2 Children and Adolescents

Zhu et al. [72] compared a set of cut-points for estimating SED and PA intensity developed using the vertical axis, with the accelerometer worn on the hip in a sample of Chinese children. The authors observed a better accuracy with the cut-points proposed by Evenson et al. [12], Vanhelst et al. [73], and those internally developed [72] than with the rest of the cut-points tested [68, 74, 75] (all these cut-points were developed with former models of ActiGraph).

Five studies developed cut-points for vector magnitude counts from the hip [13, 38, 57, 76]. Peterson et al. [76] suggested that 150 cpm from hip mounted accelerations is the most accurate SED cut-point compared with direct observation. Hänggi et al. [13] developed their cut-points using a 1-s epoch in comparison with indirect calorimetry and obtained a ROC-AUC of 0.96 for SED, light PA and moderate physical activity. These cut-points [13] obtained better correlations with other brands of accelerometers than other vertical axis based cut-points [77]. Jimmy et al. [57] developed cut-points utilizing a 5-s epoch against indirect calorimetry and attained a ROC-AUC ranging from 0.89 to 0.94 for all intensities. Romanzini et al. [38] validated cut-points using a 15-s epoch against indirect calorimetry and obtained a ROC-AUC of 0.93–0.99. Finally, Santos-Lozano et al. [27] validated cut-points utilizing a 60-s epoch against indirect calorimetry and found the lowest ROC-AUC (0.6–0.8).

Vector magnitude cut-points from the wrist placement were developed in three studies [15, 22, 37]. Chandler et al. [15] validated cut-points for the non-dominant wrist using a 5-s epoch against direct observation and attained a ROC-AUC ranging between 0.64 and 0.89. A higher ROC-AUC was obtained by Crouter et al. [37] using cut-points for the dominant wrist which were developed in a 5-s epoch against indirect calorimetry (ROC-AUC of 0.83–0.94). It is important to highlight that Crouter et al. [37] applied linear

regression models to the dominant wrist and obtained non-significant differences between the accelerometer outputs and indirect calorimetry (mean biases ranged from 2.2 to 8.4% for all cut-points).

Finally, we found two studies using metrics extracted directly from raw data instead of activity counts by ActiGraph. Aittasalo et al. [78] developed a method based on amplitude of accelerations from the hip's raw accelerations. These cut-points were validated against heart-rate monitoring using an ordinal logistic regression and showed a correlation coefficient of 0.97. However, these results must be interpreted carefully since only walking and running at different intensities were used during the development of the intensity cut points. Hildebrand et al. [22] used a linear regression analysis to establish the relation between an accelerometer metric based on raw data (i.e., Euclidean Norm Minus One) and energy expenditure measured through indirect calorimetry. Then, from the developed regression equations, they defined two sets of cut-points for the hip and the non-dominant wrist. They obtained correctly classified values between 96 and 97% for SED and light PA, 33 and 55% for moderate PA, and 68 and 80% for vigorous PA.

3.8.3 Adults

Kozey-Keadle et al. [79] tested some vertical axis-based cut-points and determined that 150 cpm using the vertical axis from hip accelerations was the most accurate SED cut-point compared with direct observation. Santos-Lozano et al. [27] validated cut-points for PA intensity against indirect calorimetry and obtained a ROC-AUC between 0.6 and 0.8. Sasaki et al. [7] used a linear regression model to establish the relation between ActiGraph vector magnitude counts from the hip and energy expenditure measured by indirect calorimetry. The mean differences between the metabolic equivalents (METs) predicted by the cut-points derived from the regression model and the actual METs were −0.3, −0.4, and 0.7 at moderate, vigorous, and very vigorous intensities, respectively.

Three studies developed cut-points from raw data metrics. Vähä-Ypyä et al. [80] developed an amplitude-domain method for raw hip accelerations. The cut-points were validated against heart-rate monitoring and showed an excellent agreement (ROC-AUC = 0.99 for all cut-points); however, they were not used during free-living conditions. Hildebrand et al. [22] validated regression models for the hip and the non-dominant wrist against indirect calorimetry and defined cut-points from the regression equations generated. They obtained correctly classified values between 93 and 96% for SED and light PA, 54 and 59% for moderate PA, and 89 and 92% for vigorous PA. Finally, Staudenmayer et al. [24] developed a classifier for PA intensity based on decision trees for the dominant wrist and they obtained 75% of values correctly classified using indirect calorimetry. Within this context, they reported preliminary results that their model performs well in a free-living environment [24].

3.8.4 Older Adults

Keadle et al. [30] observed that cut-points using the vertical axis or vector magnitude are not comparable. Unfortunately, they could not report which cut-points were the most accurate since they did not determine a criterion to compare the outcomes [30]. Aguilar-Farias et al. [26] validated SED cut-points utilizing vector magnitude counts acquired from the hip with 1-, 15-, and 60-s epochs against ActivPAL^{3TM} (Pal Technologies Ltd., Glasgow, UK) and found a high classification accuracy (ROC-AUC of 0.82, 0.85, and 0.86 for 1-, 15-, and 60-s epochs, respectively). Santos-Lozano et al. [27] validated moderate, vigorous and very vigorous PA cut-points against indirect calorimetry and obtained a ROC-AUC of 0.7 for all intensities examined.

3.9 Physical Activity Energy Expenditure (PAEE) Algorithms

Ten studies that developed PAEE algorithms were found [7, 20, 22, 27, 28, 43, 64, 71, 81, 82]. Due to an insufficient number of studies that used doubly labelled water or room calorimetry as criteria, this section combines age groups. Only two studies validated their algorithms against doubly labelled water or room calorimetry in preschoolers. Butte et al. [71] developed cross-sectional time series and multivariate adaptive regression splines to predict PAEE using both GT3X+ and heart-rate monitoring. They validated the algorithms under controlled conditions using room calorimetry and in free-living conditions utilizing doubly labelled water [71]. The multivariate adaptive regression splines obtained a better prediction of PAEE against room calorimetry, i.e., inter-method mean difference equal to 0.006 ± 0.085 kcal/min; however, the cross-sectional time series model achieved a better prediction in free-living conditions, using doubly labelled water (mean difference 41 ± 97 kcal/day) [71]. Zakeri et al. [81] used the same two statistical methods described above with GT3X/+ and heart-rate monitoring. They obtained better prediction with the cross-sectional time series model against room calorimetry (i.e., 0.001 ± 0.070 kcal/min), but they did not validate the method in a free-living environment [81]. All these studies were carried out with hip-worn GT3X/+.

3.10 Sleep-Related Behaviors

3.10.1 Preschoolers

We did not find any study comparing different sleep algorithms in this age group. Only Meredith-Jones et al. [83] used the Sadeh et al. [16] sleep algorithm to identify sleep time in preschoolers. However, this algorithm was developed in an older sample (10–25 years), and the results should be interpreted cautiously.

3.10.2 Children and Adolescents

Hjorth et al. [39] compared the performance of existent sleep algorithms from the hip versus the non-dominant wrist placements. Despite the fact that these algorithms were developed for wrist accelerations, they obtained good accuracy (86.6–89.9%) [16, 17] at both placements; however, the hip-worn device overestimated total sleep time compared to the wrist (60.1–73.8 min/day). These findings may be affected by the fact that bedtime was reported by participants using diaries and not through an algorithm.

Tudor-Locke et al. [18] developed an algorithm to detect bedtime for the hip-worn GT3X/+, and Barreira et al. [19] refined and validated it in a free-living environment against self-reported bedtime. They obtained a non-significant absolute difference of 9 ± 36 min of bedtime per night [19].

3.10.3 Adults

Cellini et al. [54] found an accuracy of 82.8% for classifying epoch-by-epoch sleep or awake status against polysomnography (i.e., the gold standard to measure sleep patterns) using the Sadeh et al. [16] sleep algorithm. However, they found an overestimation of total sleep time (i.e., inter-method mean difference equal to 8.80 min) and sleep efficiency (i.e., inter-method mean difference equal to 14.53%), as well as an underestimation of sleep onset latency (ICC = 0.56) and awakenings after sleep onset (ICC = 0.54) during a 2-h sleep protocol. Rosenberger et al. [84] observed a mean difference of 4 min of sleep time for the Sadeh et al. [16] sleep algorithm compared to the Z-machine (portable monitor to measure brain activity which is relatively comparable to polysomnography [85]). Equally, Slater et al. [86] reported good accuracy of the Sadeh et al. [16] algorithm to detect total sleep time and moderate validity for awakenings after sleep onset against polysomnography from the wrist, but not from the hip. Finally, Zinkhan et al. [87] tested the performance of the Cole et al. [17] algorithm for the hip, even though it was developed for wrist accelerations. They observed a limited agreement with total sleep time measured by polysomnography (mean difference of 81.1 min/night).

3.10.4 Older Adults

No data about sleep algorithms were found in the papers reviewed for older adults.

3.11 Step Counting

Only data on step counting estimated by ActiLife software are available in adults. Tudor-Locke et al. [44] found a higher accuracy for step counting from the hip-mounted monitors over wrist-mounted ones under controlled conditions against direct observation. Under free-living conditions, the wrist-worn accelerometer detected more steps than the hip-worn one independently of the filter used. See Sect. 3.4 for more information about how different filters influence step counting [44].

4 Discussion

The use of objective methods when assessing SED, PA, PAEE, and sleep in various research settings has increased enormously as indicated by the large number of articles included in this review. Accelerometry has several advantages over questionnaires and self-report methods [4, 10]; however, data collection and processing criteria have a large impact on the interpretation of the data. Thus, predetermined decisions about data collection and processing in relation to the study participants and the objective of the study are important when planning research in this area.

A major finding of this review is that many of the studies did not report on key methodological issues. Therefore, data cannot be accurately compared between studies and readers may have difficulties interpreting how various methodological decisions may have influenced the main findings/conclusions. We recommend that future studies should report the complete set of criteria included in the present review in order to improve data comparability and reproducibility.

The next subsections provide practical considerations for every criterion based on the critical information extracted from the validation/calibration studies reviewed.

4.1 Device Placement

Traditionally, cut-points to assess SED and PA variables, as well as PAEE estimates, were developed with the device placed on the hip, while algorithms to assess sleep-related

behaviors were more commonly developed based on wrist accelerations. Ideally, researchers may want to collect accelerometer data using a 24-h protocol with one accelerometer attached to either the hip or the wrist and be able to estimate SED, PA, and sleep-related behaviors. However, lack of validated algorithms in some age groups [15, 18, 19, 22, 24, 37] preclude this.

We decided not to differentiate between the right or left hip because no significant differences were found by Aadland et al. [40]. We consider that both the hip and the wrist are feasible places to attach the GT3X/+ . Better compliance for wrist- compared to hip-worn devices has been reported in children and adolescents [23], but similar wear-time was found in large-scale studies in adults using hip and wrist placements [36]. We therefore cannot confirm the general belief supporting better compliance for wrist-worn devices. More studies are needed to investigate compliance differences between wrist- and hip-worn devices as well as the extent to which these differences influence the validity and reliability of accelerometer outcomes.

There are only a few studies directly comparing two placement sites using the GT3X/+ and they have consistently shown more accurate classification of SED and PA intensity as well as estimates of PAEE when the accelerometer was worn on the hip compared to the wrist [13, 15, 22, 28, 37, 38, 43]. However, one study found a better performance for the wrist-worn device for PA intensity classification [24] (see Table 3). Step count also differs greatly depending on the device placement, i.e., more steps (>2500) were counted when wearing the accelerometer on the wrist compared to the hip in free-living conditions [44]. When studied under controlled conditions, hip placement has shown more accurate step counting than wrist placement from a speed of 54 m/min and upwards (at lower speeds, accuracy was better in the wrist) [44]. The lower accuracy for the wrist-worn devices could be due to the fact that accelerations such as brushing ones teeth might be interpreted as steps when the device is placed on the wrist, but not on the hip; nevertheless, this is just a hypothesis that needs to be confirmed by data under free-living conditions.

With regard to sleep algorithms, Hjorth et al. [39] compared the functioning of two algorithms applied to hip data against wrist data, finding an overestimation of the sleep time and a high accuracy (86.6 and 89.9 for each algorithm) from the hip compared to the wrist. However, it is important to note that these investigators imputed sleep and wake time manually from logs kept by the participants. The use of logs by the participants might explain the high accuracy achieved using a wrist-developed sleep algorithm on the hip.

4.2 Sampling Frequency

Our recommendation is to use the highest sampling frequency possible, as we cannot anticipate future data processing needs. However, given the issues associated with other sampling frequencies other than 30 Hz or its multiples as described in Sect. 3.2, sampling frequencies in multiples of 30 Hz seem to produce more accurate estimates when processing the signal using the methods proposed by ActiGraph. Therefore, the most reasonable conclusion for the time being would be to use 90 Hz when researchers are using the manufacturer methods, and 100 Hz when researchers are filtering and processing the signal on their own.

4.3 Valid Day and Valid Week

To ensure that data are representative of an entire day, it is necessary to establish how many hours of wear-time are required. It is also necessary to set how many valid days are needed to be representative of the total assessment period, which is usually 1 week (7 complete days). Wear-time criteria for a valid day depend on the registration period protocol, i.e., waking hours or 24 h. In studies in which the accelerometer is worn for 24-h periods to assess both physical activity and sleep-related behaviors, the number of hours required for a day to be considered valid has to be larger than studies in which the accelerometer was taken off at night.

Similarly, increasing requirements for a valid day and a valid week provides more reliable data (more information can be found in Table 5 from the study by Aadland et al. [49]); however, it results in greater sample loss. Our recommendation is to test different criteria to get the best compromise between sample size (and therefore optimal statistical power) and reliability of the measure. However, a minimum of 4 days of valid data is recommended as was suggested in a previous systematic review [88].

4.4 Filter

When movements (accelerations) occur at too low or high frequencies, ActiGraph interprets that this acceleration might not be compatible with human movement and should therefore be excluded from the analyses (e.g., if someone is using a drill). The GT3X/+ filtering process to exclude this kind of acceleration is implemented in the ActiLife software (ActiGraph, Pensacola, FL, USA). This software allows users to choose between two different filters when processing the data: normal (default) and low-frequency extension (LFE) filters.

The algorithms for these filters are proprietary information. It is known that a normal filter detects

accelerations from a frequency range of 0.25–2.5 Hz, while the LFE filter establishes a lower threshold to capture slower movements; however, it is unknown exactly how much lower this threshold is. A weighting function is applied to the accelerations between the range of 0.25–2.5 Hz, so that the full weight (i.e., 1.0) is given to a frequency of acceleration of 0.75 Hz, and lower weighting is given to higher and lower movement frequencies progressively [89]. Accelerations at a frequency greater than 2.5 Hz are removed by the filter, although it is important to highlight that accelerations up to 3.4 Hz can be produced by the human body when performing physical activity at vigorous intensities when the device is attached to the hip (higher frequencies are achieved in the wrist) [90, 91]. Therefore, ActiGraph's filtering process might remove accelerations associated with vigorous PA, and consequently, minutes in vigorous intensities might be misclassified as moderate PA [46, 92]. As the filter used has a large impact on the accelerometer outputs, it is alarming that 74% of the studies reviewed did not report this key information (Fig. 2).

When selecting a set of cut-points or an algorithm to estimate a variable from activity counts, our recommendation is to select the same filter that was used in the validation study for the cut-points or algorithm employed (Table 5 shows the filters used in all the cut-points identified). If cut-points or algorithms are not used, then researchers can decide which filter to use, we suggest using the LFE filter when low movements are of greater importance (e.g., when analyzing SED, sleep or PA in older adults). From the studies discussed above, researchers and practitioners should be aware that enabling the LFE filter compared with the normal filter will result in decreased SED, greater time in PA at all intensities and an increase in the number of steps per day.

4.5 Epoch Length

Activity counts produced by filtering raw accelerations need to be summed into specific time intervals or epoch in order to estimate PAEE, time spent in SED, in various levels of PA intensity, as well as estimating sleep/wake state, and this is usually done by applying specific intensity cut-points and algorithms.

Given that epoch length influences activity counts, it is important to use the same epoch length that was used in the validation study for the cut-points or algorithms (see Table 5). Epoch length should also be taken into account when comparing data from different studies. In young people (from preschoolers to adolescents), shorter epochs (1–15 s) are recommended to capture short bouts of activity occurring frequently in these age groups. In adults and older adults there are currently no data comparing the effect of epoch length on the outcomes studied. However, our own unpublished data suggest that selecting a 1-s versus a 60-s epoch length has a marked impact on the accelerometer outcomes, i.e., +45–60 min/day in MVPA using a 1-s compared to a 60-s epoch. This large impact on the accelerometer outputs warrants further research on this topic in order to obtain more comparable and accurate data. Considering sleep measurements there is an overall consensus for using a 60-s epoch (probably due to the stable movement pattern during sleep), as all sleep algorithms have been validated using that epoch length [16–19].

4.6 Non-Wear-Time Definition

In free-living studies, accelerometers are usually removed during water based activities, e.g., swimming or showering and when sleeping (in some studies). As a result, individuals might forget to wear the accelerometer for a day(s) or part of day. Consequently, non-wear-time must be identified (e.g., by a diary or algorithms) and excluded from data before analysis. Otherwise, this time is likely categorized as SED. Generally, algorithms to detect non-wear-time consist of intervals of time with consecutive 0 cpm with or without an allowance of several minutes in which small accelerations are allowed, with optional windows of 0 cpm before and after this allowance. Toftager et al. [47] studied the effect of different non-wear-time definitions and concluded that the most accurate algorithm might differ among subgroups of children/adolescents. For example, studies focused on overweight adolescents might need to set a longer time of consecutive 0 cpm, since they have higher SED that can be misclassified as non-wear-time.

More studies are needed to examine the accuracy of different non-wear-time detection algorithms in all age groups. Based on the reviewed studies, we cannot recommend a non-wear-time definition for preschoolers, children, or adolescents. For adults, 20 min of consecutive 0 cpm without allowance showed the lowest misclassification error; however, it may result in slightly more loss of data (6% of the sample size [60]). As the accuracy between 20 and 60 min of consecutive 0 cpm was similar (i.e., the ROC-AUC was virtually identical = 0.94), we suggest using 60 min of consecutive 0 cpm without allowing for interruptions in counts in this period for adults, to avoid the risk of misclassification of non-wear-time as SED. In older adults, we recommend the Choi et al. algorithm [45], which consists of 90 min of 0 cpm with an allowance of 2 min of activity when it is placed between two 30-min windows of 0 cpm. This algorithm outperformed other algorithms on the detection of non-wear-time [45] compared with the non-wear-time reported by participants.

4.7 Registration Period Protocol: Waking Hours versus 24 Hours

In line with recent and large-scale studies [63], we suggest registration periods of 24 h instead of waking hours (more recording time, therefore more valid data). This is mainly due to an interest in assessing sleep-related behaviors and better compliance.

4.8 SED and PA Intensity Classification

Traditionally, SED and PA intensity have been estimated based on the number of activity counts accumulated in a certain period (epoch length). Cut-points are the thresholds of activity counts used to categorize activity as SED, light, moderate, vigorous or very vigorous PA. Table 5 presents the cut-point values (expressed as counts per time unit) that are currently available for SED, and for light, moderate, vigorous, and very vigorous PA by age group. It is important to keep in mind that although the GT3X/+ is a triaxial accelerometer, the data are provided separately for the three axes plus the vector magnitude, so that it is still possible to use the data registered only by the vertical axis and apply it to the previously developed algorithms for the vertical axis.

When applying cut-points to a specific data set, it is recommended to follow the same data collection and processing criteria which were used in the original validation/calibration study (see Table 5). All derived intensity thresholds are influenced by the activities chosen when performing the calibration studies. Thus, it is impossible to recommend the most appropriate set of intensity thresholds for free-living assessment. Also, different generations of ActiGraph devices have shown to be comparable under controlled conditions [7, 93], but not in a free-living environment [7, 52, 53, 94]. This suggests that if a certain cut-point was developed, for instance using the vertical axis from the GT1M, that cut-points may not be used for data collected with the GT3X/+ vertical axis, since they are not fully comparable. Therefore, our recommendations are based on cut-points developed only with GT3X/+ accelerometers. This review shows the need for future meta-analytic studies summarizing cut-points for each age group in order to obtain a set of cut-points with a wide range of activities influencing its development. Finally, across the studies reviewed, we have observed a widely accepted criterion to define PA intensity in the studies validating cut-points against indirect calorimetry, i.e., 1–1.5 MET for SED, 1.5–3 MET for light PA, 3–6 MET for moderate PA, and >6 MET for vigorous PA.

The criteria considered for cut-point recommendations are: (1) the cut-points cover the whole activity spectrum (i.e., SED, light PA, moderate PA, and vigorous PA), (2) calorimetry as an objective criterion is better than direct observation; (3) for young populations, cut-points developed in short epochs; (4) the number and type of activities included in the study that derived the cut-points; and (5) results obtained in comparison with the criterion.

Preschoolers experience a rapid anatomical development and their patterns of PA change dramatically during the first years of life; therefore, the age of the sample is very important in preschoolers. Thus, for the hip placement we recommend Costa et al. [35] cut-points for early preschoolers (2–3 years old) and Jimmy et al. [57] cut-points for older preschoolers (4–6 years old). They have been developed in short epochs which enables the devices to capture small bouts of vigorous PA, which is typical for this age group, while obtaining high accuracy in their validation. For young preschoolers (15–36 months) using the wrist placement, we recommend Johansson et al. [34] cut-points developed using a 5-s epoch because they obtained similar accuracy to hip-developed cut-points.

For children, we recommend using the Hänggi et al. [13] cut-points developed in 1-s epoch for the hip. For adolescents, the Romanzini et al. [38] cut-points developed utilizing a 15-s epoch appears appropriate. Both of these obtained excellent classification accuracy (ROC-AUC >0.90 for all cut-points) and cover almost the whole spectrum of PA intensities. For the dominant wrist, and working with counts data, we recommend Crouter et al. [37] cut-points and for the non-dominant wrist Chandler et al. [15] cut-points. If a researcher is interested in working directly with raw data, Hildebrand et al. [22] cut-points seem to be the best options since they were validated against indirect calorimetry and they obtained relatively high accuracy, except for moderate and vigorous PA (33–80%).

For adults, 150 cpm measured using the vertical axis from hip accelerations are the best option to estimate SED [79]. For PA intensity classification, we recommend Sasaki et al. [7] cut-points developed utilizing a linear regression equation. Staudenmayer et al. [24] and Hildebrand et al. [22] cut-points are the only alternative at the moment to estimate PA from the dominant and the non-dominant wrists, respectively, considering that raw data metrics have to be used to apply them, not activity counts.

For older adults, we only found the SED cut-points proposed by Aguilar-Farias et al. [26] and the PA cut-points by Santos-Lozano et al. [27]. By combining these cut-points we can assess the whole spectrum of PA levels, which is the only option at present.

4.9 PAEE Algorithms

PAEE can be estimated using algorithms applied to GT3X/+ data. Since the same movement can produce different energy expenditure depending on the characteristics of the

individuals, caution is advised when interpreting PAEE estimated from accelerometry. It is worth highlighting that PAEE algorithms developed in a laboratory or a controlled setting are influenced by the activities selected in the study, while only studies under free-living conditions using doubly labelled water as a criterion to test validity can actually measure PAEE. Thus, in this review we have only focused on studies using doubly labelled water and room calorimetry as a criterion.

The criteria considered for PAEE algorithm recommendations were: (1) free-living studies are better than lab/controlled studies; (2) for young populations, algorithms developed in short epochs; (3) whether cross-validation was performed; and (4) results obtained in comparison with the criterion.

As noted in Sect. 4.8, different generations of ActiGraph devices are not fully comparable in free-living conditions [7, 52, 53, 94], thus, our recommendations are based on PAEE algorithms developed only with GT3X/+ accelerometers. PAEE has been expressed differently across studies, which needs to be considered when choosing a suitable algorithm. Overall, our conclusion is that more validation studies during free-living conditions utilizing doubly labelled water are needed in all age groups.

For preschoolers, we recommend to use the algorithm proposed by Butte et al. [71] as it is the only one validated in free-living conditions against doubly labelled water and they obtained a high accuracy. There are no algorithms for wrist accelerations in this age group. Likewise, we do not recommend algorithms for the rest of age-groups since none of them were developed using doubly labelled water or room calorimetry as a criterion.

4.10 Sleep-Related Behaviors

The ActiGraph GT3X/+ can identify sleep-related behaviors from movement/non-movement patterns by applying sleep algorithms to activity counts. The overall conclusion is that more studies developing and validating sleep algorithms for the wrist and the hip mounted ActiGraphs are needed in all age groups. However, based on the aforementioned information and the ages of the samples in the validation studies for sleep algorithms, we recommend use of the Barreira et al. [19] algorithm in children and adolescents when the accelerometer is attached to the hip to document bedtime that is not reported by the participants. Also, we recommend the Sadeh et al. [16] algorithm if it is placed on the wrist to score the sleep time and the rest of sleep-related behaviors. Moreover, we recommend the Sadeh et al. [16] algorithm for young adults (i.e., up to 30 years of age) and Cole-Kripke et al. [17] algorithm for older adults (i.e., >30 years of age) (in both cases with the accelerometer placed on the wrist).

4.11 Step Counting

We have reviewed studies using the default step counting function by ActiLife. In this regard, we recommend using the normal filter when the step count is a variable of interest, as it has been demonstrated to be more comparable to other criterion devices than the LFE filter (see Sect. 3.4).

5 Limitations and Strengths

Several limitations need to be acknowledged. Studies with earlier models than GT3X/+ (e.g., GT1M) have not been included in our review, so our recommendations are limited to the triaxial ActiGraph models (GT3X/+). Another limitation is that for certain age groups and for some accelerometer criteria analyzed, the number of studies was small; therefore, the recommendation should be revisited when more studies on those topics are available. In addition, the field of accelerometry is rapidly developing and continuously changing. Therefore, reviews are needed every few years in order to update the recommendations provided in this review. Another major issue is that proprietary algorithms used by the manufacturer when processing the data to obtain activity counts are unavailable to the public and these affect outputs. Future work using the raw acceleration signal (i.e., before any filtering is applied) should overcome this problem. Finally, another limitation is that some of our recommendations are based on few studies and should be interpreted cautiously. Thus, further studies such as a formal meta-analysis may provide the most optimal intensity thresholds for the different intensity thresholds.

The strengths of this review are: (1) the inclusion of a large number of studies, summarizing the methodologies used in each of them, which will allow for more accurate comparability of the data; (2) the separate sections for the validation/calibration studies in order to provide guidance and recommendations to researchers and practitioners; (3) the inclusion of all age groups in one single review, which will allow researchers to find/read the information about the age group they are working with/interested in; and (4) the set of tables included in this review were developed to assist researchers in their decision making process (see the examples included in Sect. 5).

6 Practical Implications

This review will help researchers and practitioners to make better decisions when designing their study and processing the data from the GT3X/+ accelerometer in order to obtain the most accurate and comparable information. Here, we provide some hypothetical examples illustrating how the

information presented in the tables in this review can be used.

- A researcher intends to evaluate accelerometry in a new study and needs to know where to place the accelerometer. Table 3 summarizes the most important results obtained when comparing the outputs from the GT3X/+ attached to the hip versus the wrist and provides recommendations depending on the variables to be analyzed by age group.
- A researcher has collected accelerometry data and wishes to compare the data with those from other studies to generate an accurate and meaningful discussion. Table 2 lists the criteria used for data collection protocols and data processing in studies. Electronic Supplementary Material Appendix S3 lists all articles that have been used for each of these criteria.
- A researcher has collected accelerometry data with the device placed on the wrist (for example) and wishes to know which cut-points, PAEE or sleep algorithms can be applied to those data. Tables 4 and 5 will help the researcher answer these questions.
- A researcher has decided to apply a specific set of cut-points based on the characteristics of his/her sample but is uncertain which exact setting was used in the original study (and is aware that it is recommended that the same settings be used to ensure the cut-points are applied correctly to the new data). Table 5 lists all criteria needed to correctly apply these cut-points (i.e., placement, filter, vector and epoch).

7 Conclusion

We suggest that researchers who assess SED, PA, PAEE, sleep-related behaviors, and/or steps using GT3X/+ select the specific placement, sampling frequency, filter, epoch length, non-wear-time definition, valid days and valid week criteria, SED and PA intensity classification, PAEE, and sleep algorithms depending on the population's age (i.e., preschoolers, children and adolescents, adults, or older adults). Likewise, when selecting a specific cut-point or algorithm, it is important to apply the same criteria as in the original validation/calibration study. Moreover, this review has identified some issues in the studies using the GT3X/+ during the last 5 years, such as that many studies do not report all of the criteria used in their analyses (see Fig. 2). Future studies are recommended to report the criteria as summarized in the present review.

Although ideally researchers should select all the data collection and processing criteria before the assessment period, it is important to note that only the placement and sampling frequency criteria have to be decided a priori (i.e., before the measurement period), while the rest of processing decisions can be made a posteriori. This is important since new and better analytical methods might emerge after a study was planned, and they should be considered and tested, at least as sensitivity analyses. The preliminary evidence comparing wrist and hip placements seems to support the idea that a similar compliance can be achieved wearing the accelerometer on the wrist or on the hip, while wearing it on the hip might produce more accurate estimates of PAEE and better SED and PA intensity classifications; however, these notions need to be confirmed or refuted in future studies. We recommend recording raw data for complete days (i.e., 24-h periods), so that collected data will have the maximum potential for future analyses. The summary tables presented in this systematic review will help researchers to make better decisions on how to design and process the GT3X/+ data.

Acknowledgements We are deeply thankful to Patty Freedson, Professor (University of Massachusetts/Amherst, USA) and Catrine Tudor-Locke, PhD (University of Massachusetts/Amherst, USA), for their comments on an earlier draft. This is part of a PhD Thesis conducted in the Biomedicine Doctoral Studies at the University of Granada, Spain.

Compliance with ethical standards

Funding This review was conducted under the umbrella of the ActiveBrains project (DEP2013-47540). Jairo H. Migueles is supported by a Grant from the Spanish Ministry of Education, Culture and Sport (FPU15/02645). Cristina Cadenas-Sanchez is supported by a Grant from the Spanish Ministry of Economy and Competitiveness (BES-2014-068829). Jose Mora-Gonzalez is supported by a Grant from the Spanish Ministry of Education, Culture and Sport (FPU14/06837). Francisco B. Ortega and Jonatan R. Ruiz are supported by Grants from the Spanish Ministry of Science and Innovation (RYC-2011-09011 and RYC-2010-05957, respectively). Ulf Ekelund is supported by Grants from the Research Council of Norway (249932/F20) and the UK Medical Research Council (MC_UU_12015/3). Additional funding was obtained from the University of Granada, Plan Propio de Investigación 2016, Excellence actions: Units of Excellence; Unit of Excellence on Exercise and Health (UCEES). In addition, funding was provided by the SAMID III network, RETICS, funded by the PN I + D+I 2017-2021 (Spain), ISCIII- Sub-Directorate General for Research Assessment and Promotion, the European Regional Development Fund (ERDF) (Ref. RD16/0022) and the EXERNET Research Network on Exercise and Health in Special Populations (DEP2005-00046/ACTI).

Conflict of interest Jairo H. Migueles, Cristina Cadenas-Sanchez, Ulf Ekelund, Christine Delisle Nyström, Jose Mora-Gonzalez, Marie Löf, Idoia Labayen, Jonatan R. Ruiz, and Francisco B. Ortega declare that they have no conflicts of interest relevant to the content of this review.

References

1. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ. 2006;174:801–9.

2. Fletcher G, Balady G, Blair S, et al. Statement on exercise: benefits and recommendations for physical activity programs for all Americans. Circulation. 1996;94:857–62.
3. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act. 2010;7:40.
4. Warren JM, Ekelund U, Besson H, et al. Assessment of physical activity—a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil. 2010;17:127–39.
5. Bassett DR, Rowlands A, Trost SG. Calibration and validation of wearable monitors. Med Sci Sports Exerc. 2012;44:32–8.
6. Rothney MP, Brychta RJ, Meade NN, et al. Validation of the ActiGraph two-regression model for predicting energy expenditure. Med Sci Sports Exerc. 2010;42:1785–92.
7. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport. 2011;14:411–6.
8. Baranowski T, Dworkin RJ, Cieslik CJ, et al. Reliability and validity of self-report of aerobic activity—Family Health Project. Res Q Exerc Sport. 1984;55:309–17.
9. Sallis JF. Self-report measures of children's physical activity. J Sch Health. 1991;61:215–9.
10. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med. 2001;31:439–54.
11. Freedson PS, Melanson E, Sirard JR. Calibration of the computer science and applications, Inc. accelerometer. Med Sci Sports Exerc. 1998;30:777–81.
12. Evenson KR, Catellier DJ, Gill K, et al. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008;26:1557–65.
13. Hänggi JM, Phillips LRS, Rowlands AV. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. J Sci Med Sport. 2013;16:40–4. doi:10.1016/j.jsams.2012.05.012.
14. Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. J Aging Phys Act. 2009;17:17–30.
15. Chandler JL, Brazendale K, Beets MW, et al. Classification of physical activity intensities using a wrist-worn accelerometer in 8- to 12-year-old children. Pediatr Obes. 2015;11(2):120–7. doi:10.1111/ijpo.12033.
16. Sadeh A, Sharkey KM, Carskadon MA. Activity-based sleep-wake identification: an empirical test of methodological issues. Sleep. 1994;17(3):201–6.
17. Cole RJ, Kripke DF, Gruen W, et al. Automatic sleep/wake identification from wrist activity. Sleep. 1992;15:461–9.
18. Tudor-Locke C, Barreira TV, Schuna JM, et al. Fully automated waist-worn accelerometer algorithm for detecting children's sleep-period time separate from 24-h physical activity or sedentary behaviors. Appl Physiol Nutr Metab. 2014;39:53–7.
19. Barreira TV, Schuna JM, Mire EF, et al. Identifying children's nocturnal sleep using 24-h waist accelerometry. Med Sci Sports Exerc. 2015. doi:10.1249/MSS.0000000000000486.
20. Crouter SE, Horton M, Bassett DR. Use of a two-regression model for estimating energy expenditure in children. Med Sci Sports Exerc. 2012;44:1177–85.
21. Trost SG, Ward DS, Moorehead SM, et al. Validity of the computer science and applications (CSA) activity monitor in children. Med Sci Sports Exerc. 1998;30:629–33.
22. Hildebrand M, Van Hees VT, Hansen BH, et al. Age-group comparability of raw accelerometer output from wrist- and hip-worn monitors. Med Sci Sports Exerc. 2014;46(9):1816–24. doi:10.1249/MSS.0000000000000289.
23. Fairclough SJ, Noonan R, Rowlands AV, et al. Wear compliance and activity in children wearing wrist and hip-mounted accelerometers. Med Sci Sports Exerc. 2016;48(2):245–53. doi:10.1249/MSS.0000000000000771.
24. Staudenmayer J, He S, Hickey A, et al. Methods to estimate aspects of physical activity and sedentary behavior from high frequency wrist accelerometer measurements. J Appl Physiol. 2015. doi:10.1152/japplphysiol.00026.2015.
25. Wijndaele K, Westgate K, Stephens SK, et al. Utilization and harmonization of adult accelerometry data. Med Sci Sports Exerc. 2015;47(10):2129–39.
26. Aguilar-Farias N, Brown WJ, Peeters GM. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. J Sci Med Sport. 2014;17:293–6. doi:10.1016/j.jsams.2013.07.002.
27. Santos-Lozano A, Santín-Medeiros F, Cardon G, et al. Actigraph GT3X: validation and determination of physical activity intensity cut points. Int J Sports Med. 2013;34:975–82.
28. Ellis K, Kerr J, Godbole S, et al. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. Physiol Meas. 2014;35:2191–203.
29. Kim Y, Lee JM, Peters BP, et al. Examination of different accelerometer cut-points for assessing sedentary behaviors in children. PLoS One. 2014;9:1–8.
30. Keadle SK, Shiroma EJ, Freedson PS, et al. Impact of accelerometer data processing decisions on the sample size, wear-time and physical activity level of a large cohort study. BMC Public Health. 2014;14(1):1210.
31. Cain KL, Sallis JF, Conway TL, et al. Using accelerometers in youth physical activity studies: a review of methods. J Phys Act Health. 2013;10:437–50.
32. Booth A, Clarke M, Dooley G, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. Syst Rev. 2012;1:2.
33. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Reprinted from Annals of Internal Medicine). Phys Ther. 2009;89:873–80.
34. Johansson E, Ekelund U, Nero H, et al. Calibration and cross-validation of a wrist-worn ActiGraph in young preschoolers. Pediatr Obes. 2015;10(1):1–6. doi:10.1111/j.2047-6310.2013.00213.x.
35. Costa S, Barber SE, Cameron N, et al. Calibration and validation of the ActiGraph GT3X+ in 2–3 year olds. J Sci Med Sport. 2013;17:617–22. doi:10.1016/j.jsams.2013.11.005.
36. Tudor-Locke C, Barreira TV, Schuna JM, et al. Improving wear-time compliance with a 24-h waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). Int J Behav Nutr Phys Act. 2015;12(1):1–9.
37. Crouter SE, Flynn JI, Bassett DR. Estimating physical activity in youth using a wrist accelerometer. Med Sci Sports Exerc. 2015;47:944–7.
38. Romanzini M, Petroski EL, Ohara D, et al. Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. Eur J Sport Sci. 2014;14(1):91–9. doi:10.1080/17461391.2012.732614.
39. Hjorth MF, Chaput JP, Damsgaard CT, et al. Measure of sleep and physical activity by a single accelerometer: can a waist-worn ActiGraph adequately measure sleep in children? Sleep Biol Rhythms. 2012;10:328–35.
40. Aadland E, Ylvisåker E. Reliability of the ActiGraph GT3X+ accelerometer in adults under free-living conditions. PLoS One. 2015;10:e0134606. doi:10.1371/journal.pone.0134606.

41. Ozemek C, Kirschner MM, Wilkerson BS, et al. Intermonitor reliability of the GT3X + accelerometer at hip, wrist and ankle sites during activities of daily living. Physiol Meas. 2014;35:129–38.
42. Ellis K, Kerr J, Godbole S, et al. Hip and wrist accelerometer algorithms for free-living behavior classification. Med Sci Sports Exerc. 2016;48(5):933–40. doi:10.1249/MSS.0000000000000840.
43. Stec MJ, Rawson ES. Estimation of resistance exercise energy expenditure using triaxial accelerometry. J Strength Cond Res. 2012;26:1413–22.
44. Tudor-Locke C, Barreira TV, Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. Med Sci Sports Exerc. 2014;47(4):839.
45. Choi L, Ward SC, Schnelle JF, et al. Assessment of wear/non-wear-time classification algorithms for triaxial accelerometer. Med Sci Sports Exerc. 2012;44:2009–16.
46. Brønd C, Arvidsson D. Sampling frequency affects the processing of ActiGraph raw acceleration data to activity counts. J Appl Physiol. 2016;120(3):362–9. doi:10.1152/japplphysiol.00628.2015.
47. Toftager M, Kristensen PL, Oliver M, et al. Accelerometer data reduction in adolescents: effects on sample retention and bias. Int J Behav Nutr Phys Act. 2013;10:140.
48. Freedson PS, John D. Comment on "Estimating activity and sedentary behavior from an accelerometer on the hip and wrist". Med Sci Sports Exerc. 2013;45:962–3.
49. Aadland E, Ylvisåker E. Reliability of objectively measured sedentary time and physical activity in adults. PLoS One. 2015;10:e0133296. doi:10.1371/journal.pone.0133296.
50. Donaldson SC, Montoye AHK, Tuttle MS, et al. Variability of objectively measured sedentary behavior. Med Sci Sports Exerc. 2016;48(4):755–61. doi:10.1249/MSS.0000000000000828.
51. Lyden K, Kozey Keadle SL, Staudenmayer JW, et al. Validity of two wearable monitors to estimate breaks from sedentary time. Med Sci Sports Exerc. 2012;44:2243–52.
52. Ried-Larsen M, Brønd JC, Brage S, et al. Mechanical and free living comparisons of four generations of the ActiGraph activity monitor. Int J Behav Nutr Phys Act. 2012;9:1–10. doi:10.1186/1479-5868-9-113.
53. Cain KL, Conway TL, Adams MA, et al. Comparison of older and newer generations of ActiGraph accelerometers with the normal filter and the low frequency extension. Int J Behav Nutr Phys Act. 2013;10:51.
54. Cellini N, Buman MP, McDevitt EA, et al. Direct comparison of two actigraphy devices with polysomnographically recorded naps in healthy young adults. Chronobiol Int. 2013;30:691–8.
55. Wanner M, Martin BW, Meier F, et al. Effects of filter choice in GT3X accelerometer assessments of free-living activity. Med Sci Sports Exerc. 2013;45:170–7.
56. Barreira TV, Brouillette RM, Foil HC, et al. Comparison of older adults steps/day using NL-1000 pedometer and two GTX+ accelerometer filters. J Aging Phys Act. 2012;21:402–16.
57. Jimmy G, Seiler R, Mäder U. Development and validation of GT3X accelerometer cut-off points in 5- to 9-year-old children based on indirect calorimetry measurements. Schweizerische Zeitschrift fur Sport und Sport. 2013;61:37–43.
58. Kahan D, Nicaise V, Reuben K. Convergent validity of four accelerometer cutpoints with direct observation of preschool children's outdoor physical activity. Res Q Exerc Sport. 2013;84:59–67. doi:10.1080/02701367.2013.762294.
59. Aibar A, Bois JE, Zaragoza J, et al. Do epoch lengths affect adolescents' compliance with physical activity guidelines? J Sports Med Phys Fit. 2014;54:255–63.
60. Peeters G, van Gellecum Y, Ryde G, et al. Is the pain of activity log-books worth the gain in precision when distinguishing wear and non-wear-time for tri-axial accelerometers? J Sci Med Sport. 2013;16:515–9. doi:10.1016/j.jsams.2012.12.002.
61. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008;40:181–8.
62. Choi L, Liu Z, Matthews E, et al. Validation of accelerometer wear and nonwear-time classification algorithm. Med Sci Sports Exerc. 2012;43:357–64.
63. Katzmarzyk PT, Barreira TV, Broyles ST, et al. The International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE): design and methods. BMC Public Health. 2013;13:900–13.
64. Janssen X, Cliff DP, Reilly JJ, et al. Predictive validity and classification accuracy of ActiGraph energy expenditure equations and cut-points in young children. PLoS One. 2013;8(11):e79124–9.
65. Sirard JR, Trost SG, Pfeiffer KA, et al. Calibration and evaluation of an objective measure of physical activity in preschool children. J Phys Act Health. 2005;2:345–57.
66. Reilly JJ, Coyle J, Kelly L, et al. An objective method for measurement of sedentary behavior in 3- to 4-year olds. Obes Res. 2003;11:1155–8.
67. Pate RR, Almeida MJ, McIver KL, et al. Validation and calibration of an accelerometer in preschool children. Obesity (Silver Spring). 2006;14:2000–6.
68. Puyau MR, Adolph AL, Vohra FA, et al. Validation and calibration of physical activity monitors in children. Obes Res. 2002;10:150–7.
69. Van Cauwenberghe E, Gubbels J, De Bourdeaudhuij I, et al. Feasibility and validity of accelerometer measurements to assess physical activity in toddlers. Int J Behav Nutr Phys Act. 2011;8:67.
70. Pulakka A, Cheung YB, Ashorn U, et al. Feasibility and validity of the ActiGraph GT3X accelerometer in measuring physical activity of Malawian toddlers. Acta Paediatr Int J Paediatr. 2013;102:1192–8.
71. Butte NF, Wong WW, Lee JS, et al. Prediction of energy expenditure and physical activity in preschoolers. Med Sci Sports Exerc. 2014;46:1216–26.
72. Zhu Z, Chen P, Zhuang J. Intensity classification accuracy of accelerometer-measured physical activities in Chinese children and youth. Res Q Exerc Sport. 2013;84:S4–11. doi:10.1080/02701367.2013.850919.
73. Vanhelst J, Béghin L, Turck D, et al. New validated thresholds for various intensities of physical activity in adolescents using the ActiGraph accelerometer. Int J Rehabil Res. 2011;34:175–7.
74. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. Med Sci Sports Exerc. 2005;37:523–30.
75. Mattocks C, Leary S, Ness A, et al. Calibration of an accelerometer during free-living activities in children. Int J Pediatr Obes. 2007;2:218–26.
76. Peterson NE, Sirard JR, Kulbok PA, et al. Validation of accelerometer thresholds and inclinometry for measurement of sedentary behavior in young adult university students. Res Nurs Health. 2015;38:492–8. doi:10.1002/nur.21694.
77. Rowlands AV, Rennie K, Kozarski R, et al. Children's physical activity assessed with wrist- and hip-worn accelerometers. Med Sci Sports Exerc. 2014;2006:2308–16.
78. Aittasalo M, Vähä-Ypyä H, Vasankari T, et al. 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 Sci Med Rehabil. 2015;7:18. doi:10.1186/s13102-015-0010-0.
79. Kozey-Keadle S, Libertine A, Lyden K, et al. Validation of wearable monitors for assessing sedentary behavior. Med Sci Sports Exerc. 2011;43:1561–7.

80. Vähä-Ypyä H, Vasankari T, Husu P, et al. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. Clin Physiol Funct Imaging. 2015;35(1):64–70. doi:10.1111/cpf.12127.
81. Zakeri IF, Adolph AL, Puyau MR, et al. Cross-sectional time series and multivariate adaptive regression splines models using accelerometry and heart rate predict energy expenditure of preschoolers. J Nutr. 2013;143:114–22.
82. Zhu Z, Chen P, Zhuang J. Predicting Chinese children and youth's energy expenditure using ActiGraph accelerometers: a calibration and cross-validation study. Res Q Exerc Sport. 2013;84:S56–63. doi:10.1080/02701367.2013.850989.
83. Meredith-Jones K, Williams S, Galland B, et al. 24 h accelerometry: impact of sleep-screening methods on estimates of sedentary behaviour and physical activity while awake. J Sports Sci. 2015;414:1–7. doi:10.1080/02640414.2015.1068438.
84. Rosenberger ME, Buman MP, Haskell WL, et al. Twenty-four hours of sleep, sedentary behavior, and physical activity with nine wearable devices. Med Sci Sports Exerc. 2016;48(3):457–65. doi:10.1249/MSS.0000000000000778.
85. Kaplan RF, Wang Y, Loparo KA, et al. Performance evaluation of an automated single-channel sleep-wake detection algorithm. Nat Sci Sleep. 2014;6:113–22. doi:10.2147/NSS.S71159.
86. Slater JA, Botsis T, Walsh J, et al. Assessing sleep using hip and wrist actigraphy. Sleep Biol Rhythms. 2015;13(2):172–8. doi:10.1111/sbr.12103.
87. Zinkhan M, Berger K, Hense S, et al. Agreement of different methods for assessing sleep characteristics: a comparison of two actigraphs, wrist and hip placement, and self-report with polysomnography. Sleep Med. 2014;15(9):1107–14. doi:10.1016/j.sleep.2014.04.015.
88. Trost SG, Mciver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc. 2005;37:531–43.
89. Tryon WW, Williams R. Fully proportional actigraphy: a new instrument. Behav Res Methods Instru Comput. 1996;28:392–403.
90. Cavagna GA, Franzetti P. The determinants of the step frequency in walking in humans. J Physiol. 1986;373:235–42.
91. Cavagna GA, Willems PA, Franzetti P, et al. The two power limits conditioning step frequency in human running. J Physiol. 1991;437:95–108.
92. John D, Miller R, Kozey-Keadle S, et al. Biomechanical examination of the "plateau phenomenon" in ActiGraph vertical activity counts. Physiol Meas. 2012;33:219–30.
93. Robusto KM, Trost SG. Comparison of three generations of ActiGraph^{TM} activity monitors in children and adolescents. J Sports Sci. 2012;30:1429–35.
94. Grydeland M, Hansen BH, Ried-Larsen M, et al. Comparison of three generations of ActiGraph activity monitors under free-living conditions: do they provide comparable assessments of overall physical activity in 9-year old children? BMC Sports Sci Med Rehabil. 2014;6:26.
95. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. Med Sci Sports Exerc. 2004;36:1259–66.
96. Ridgers ND, Salmon J, Ridley K, et al. Agreement between activPAL and ActiGraph for assessing children's sedentary time. Int J Behav Nutr Phys Act. 2012;9:15.
97. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors in the United States, 2003-2004. Am J Epidemiol. 2008;167:875–81.
98. Metzger JS, Catellier DJ, Evenson KR, et al. Patterns of objectively measured physical activity in the United States. Med Sci Sports Exerc. 2008;40:630–8.
99. Davis MG, Fox KR. Physical activity patterns assessed by accelerometry in older people. Eur J Appl Physiol. 2007;100:581–9.
100. Andersen LB, Harro M, Sardinha LB, et al. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). Yearb Sport Med. 2006;368:299–304.
101. Van Cauwenberghe E, Labarque V, Trost SG, et al. Calibration and comparison of accelerometer cut points in preschool children. Int J Pediatr Obes. 2011;6:e582–9.
102. Grydeland M, Bergh IH, Bjelland M, et al. Correlates of weight status among Norwegian 11-year-olds: the HEIA study. BMC Public Health. 2012;12:1053.
103. Matthews C. Calibration for accelerometer output for adults. Med Sci Sports Exerc. 2005;S512:S512–22.
104. Pruitt LA, Glynn NW, King AC, et al. Use of accelerometry to measure physical activity in older adults at risk for mobility disability. J Aging Phys Act. 2008;16:416–34.
105. Zisko N, Carlsen T, Salvesen Ø, et al. New relative intensity ambulatory accelerometer thresholds for elderly men and women: the Generation 100 study. BMC Geriatr. 2015;15:97.
106. Crouter SE, Churilla JR, Bassett DR. Estimating energy expenditure using accelerometers. Eur J Appl Physiol. 2006;98:601–12.
107. Crouter SE, Kuffel E, Haas JD, et al. Refined two-regression model for the ActiGraph accelerometer. Med Sci Sports Exerc. 2010;42:1029–37.
108. Liu S, Gao RX, Freedson PS. Computational methods for estimating energy expenditure in human physical activities. Med Sci Sports Exerc. 2012;44:2138–46.
109. Schmitz KH, Treuth M, Hannan P, et al. Predicting energy expenditure from accelerometry counts in adolescents girls. Med Sci Sports Exerc. 2005;37:155–61.
110. Mâsse LC, Fuemmeler BF, Anderson CB, et al. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. Med Sci Sports Exerc. 2005;37(11 Suppl):S544–54.

View publication stats