Calibration and validation of physical behaviour cut-points using wrist-worn ActiGraphs for children and adolescents: A systematic review

Journal of Science and Medicine in Sport xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Journal of Science and Medicine in Sport

journal homepage: www.elsevier.com/locate/jsams

Review

Calibration and validation of physical behaviour cut-points using wrist-worn ActiGraphs for children and adolescents: A systematic review

Kelly Clanchy [a,b,*], Matthew Stanfield [c], Esther Smits [d], Jenna Liimatainen [d], Carrie Ritchie [d]

[a] School of Health Sciences and Social Work, Griffith University, Gold Coast Campus, Australia
[b] Menzies Health Institute of Queensland, Griffith University, Gold Coast Campus, Australia
[c] School of Human Movement and Nutrition Sciences, University of Queensland, Australia
[d] RECOVER Injury Research Centre and National Health and Medical Research Council (NHMRC) Centre for Research Excellence: Better Health Outcomes After Compensable Injury, University of Queensland, Australia

ARTICLE INFO

Article history:
Received 17 May 2023
Received in revised form 2 November 2023
Accepted 14 November 2023
Available online xxxx

Keywords:
Paediatrics
Measurement
Fitness tracker
Physical exertion
Sedentary

ABSTRACT

Objectives: To review cut-points calibrated and independently validated from wrist-worn ActiGraph accelerometers to measure moderate to vigorous physical activity (MVPA) and time spent sedentary (SED) in children and adolescents.
Design: Systematic literature review.
Methods: Five databases were searched for relevant cut-point calibration and independent validation studies relating to wrist worn ActiGraphs in children and adolescents from inception through 30 April 2022. Extracted data included: country of publication; study name; population; device model; wear location; sampling frequency; epoch length; activity protocol; criterion method and definitions used to classify PA intensity; statistical methods for calibration; statistical methods for validation/cross-validation; and MVPA and SED outcome.
Results: Fourteen calibration studies and seven independent validation studies were identified. Calibrated cut-points for MVPA vector magnitude counts ranged from 7065 to 9204 counts per minute (cpm) and 63.5 to 201 milli-gravitational units (*mg*). For SED, calibrated cut-points ranged from <2556 cpm to 4350 cpm and 30.8 to 48.1 *mg*. Classification accuracy values determined by independent validation studies varied, with kappa values ranging from 0.31 to 0.60 and area under the curve statistics ranging from 0.51 to 0.84 for MVPA and kappa values ranging from 0.31 to 0.44 and area under the curve statistics ranging from 0.70 to 0.85 for SED.
Conclusions: The results of this systematic literature review support the use of the Crouter and colleagues cut-points for the measurement of MVPA and SED for children and adolescents aged 6–12 years. Further work is required to independently validate cut-points developed in younger children and older adolescents.

Practical implications

- Wrist-worn accelerometers are one of the best options for accurately assessing physical activity in children due to their small, unobtrusive, and light-weight design, ability to objectively collect short bouts of activity data in real time, and evidence of increased wear compliance. ActiGraph accelerometers (ActiGraph LLC, Pensacola, FL, USA) are the most commonly used wearable, research-grade accelerometer.
- Cut-points provide a straightforward method to convert accelerometer data time spent in different physical activity intensities, which allows meaningful comparison with the physical activity guidelines. Studies relating to cut point development can be divided into two groups: (1) "calibration studies" in which cut-points are determined based on the classification of ActiGraph counts using criterion activities or measures representing pre-established activity intensities; or (2) "independent validation studies" in which previously calibrated cut-points are validated in an independent paediatric population. Independent validation is required to reduce the likelihood of bias due to study characteristics, populations, or protocol design.
- For children and adolescents aged 6–12 years, the Crouter and colleagues cut-points (MVPA = VM ≥ 610 counts/5 s; SED = VM <100 counts/5 s) should be used to classify moderate-to-vigorous physical activity and time spent sedentary. The ActiGraph accelerometer should be placed on the dominant wrist, using a sampling frequency of 30 Hz and 1-second epochs.

1. Introduction

Physical activity (PA) is defined 'as any bodily movement produced by skeletal muscle contracture that results in caloric expenditure' and

* Corresponding author.
E-mail address: k.clanchy@griffith.edu.au (K. Clanchy).

https://doi.org/10.1016/j.jsams.2023.11.008

includes sport, exercise, active hobbies (including play), walking, cycling, and some activities of daily living.[1] For infants and toddlers, active play leads to improved balance, movement, and coordination.[2] Improvements in bone health, weight, cognitive function, cardiorespiratory and muscular fitness, and mental health are additional benefits for children and adolescents.[3,4] International PA guidelines recommend that children accumulate at least 60 min of moderate-to-vigorous intensity PA (MVPA) per day and limit the amount of time spent sedentary (SED).[5–7] These guidelines do not include a specific recommendation for light intensity physical activity.

Quantifying MVPA for children is challenging since it is likely to be achieved through participation in activities that are often unscheduled and sporadic such as ball games, swimming, dancing, skipping, and running. Accelerometers are a widely accepted method of objectively measuring free-living PA and sedentary time for children.[8,9] Accelerometers are small, unobtrusive, and light-weight, and can objectively collect activity data in real time, including short activity bouts that are difficult to capture with questionnaires, interviews, or diaries.[8–11]

Historically, most accelerometer data were derived from hip-worn devices.[12] However, wear compliance of hip-located devices is poor for children and adolescents.[13–16] Perceived discomfort,[15] embarrassment,[14] and forgetting to put the monitor back on after taking it off at night were reported reasons for non-compliance.[15,17] Children state that wrist-worn devices are acceptable[13,18] and compliance with 24-hour wear protocols is optimised by enabling measurement of the full spectrum of physical behaviours (PA, sedentary behaviour and sleep).[19]

ActiGraph accelerometers (ActiGraph LLC, Pensacola, FL, USA) are the most commonly used wearable, triaxial research-grade accelerometer.[9,20–22] Triaxial ActiGraph devices measure acceleration in three different axes (transverse, sagittal, and anterio-posterior) with raw accelerometer signals recorded at a set frequency. These signals are converted to counts per minutes (cpm) for discrete periods or epochs (e.g., 1 min) via ActiLife software (ActiGraph LLC, Pensacola, FL, USA). The process and/or algorithms used to convert raw accelerometer data into counts differ between devices as each device makes use of advances in hardware and software technology, with much of this process held proprietary.[23] Whilst various methods have been used to generate activity metrics from accelerometer data (e.g., cut-points, machine learning methods, or MATLAB or R software), the majority of studies use 'cut-points' to categorise PA intensity and SED from cpm.[24] Cut-points provide a straightforward method to convert cpm to PA intensity levels and SED, thereby facilitating interpretation of these data in line with recommendations and guidelines.[25]

Currently, there is no consensus for optimal cut-points to estimate PA intensity and SED from wrist-worn ActiGraph accelerometers in children. Compared with accelerations from hip-worn devices, accelerations from wrist-worn devices are disproportionately greater for movements using the arms including various sports, computer gaming and homework, therefore hip-worn ActiGraph cut points cannot be used for wrist-worn accelerometer data.[13,22] To effectively use cut-points, an understanding of the methodological criteria used to calibrate and validate the cut-points is required.[22,24]

The aim of this study was to review studies that calibrated and independently validated ActiGraph data to generate cut-points to identify MVPA and SED using wrist-worn ActiGraph accelerometers in children and adolescents. The outcomes of this study aim to inform decision-making processes for the selection and administration of wrist-worn ActiGraph accelerometers to measure MVPA and SED, and consequently assess compliance to the physical activity guidelines, for children and adolescents.

2. Methods

This systematic review was conducted and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[26] guidelines. Protocol registration was with the Prospective Register of Systematic Reviews (PROSPERO registration number: CRD42020148543).

2.1. Search strategy

Five electronic databases: Medline, Embase, Web of Science, PsycInfo, and CINAHL were searched for studies from inception through 30 April 2022. The search strategy was designed by the authors with the assistance of a professional librarian. Full details of the search terms used for each of the databases are shown in Supplementary Material A.

2.2. Eligibility criteria

A three-stage review process of the identified records was undertaken. Eligibility criteria for each step were applied independently to each identified record by two of four researchers (CR, ES, MCS, or JL). Disagreements were resolved by discussion between the two researchers, and if necessary, a third reviewer made the final decision.

2.2.1. Stage 1

The search strategy was applied to identify all studies that met the inclusion criteria. All studies were uploaded to Covidence, a web-based collaboration software platform (Veritas Health Innovation, Melbourne, Australia) and the following filters were applied at the title/abstract review step: child, children, adolescent, youth, paediatric, juvenile, teen, girl, boy, school, puberty, pubertal, childhood, infant, schoolchildren, toddler, young person, young people, and teenager.

2.2.2. Stage 2

All titles and abstracts identified at Stage 1 were screened for eligibility, followed by full-text review when eligibility was uncertain. Records, excluding protocol papers, reviews, theses and conference abstracts, were included that: (1) measured PA using a wrist-worn ActiGraph device; (2) converted PA data retrieved from the ActiGraph to MVPA and SED using specified cut-points; (3) included participants who were ≤18 years, or if adult populations were included, analysed MVPA and SED data separately for children and adults; (4) included participants who were ambulatory (e.g., did not use mobility aids such as wheelchair, walking frame, rollator, or a walking stick/cane); and (5) were published in English.

2.2.3. Stage 3

Studies were divided into two groups: (1) "calibration studies" in which cut-points are determined based on the classification of ActiGraph counts using criterion activities or measures representing pre-established activity intensities; or (2) "independent validation studies" in which previously calibrated cut-points are validated in an independent paediatric population. Studies were excluded that: (1) did not undertake calibration or validation assessments; (2) did not calibrate or validate ActiGraph counts using cut-point methods, e.g., used machine learning classification models; or (3) used MVPA and SED cut-point data from ActiGraphs to validate another device or instrument.

To ensure no additional relevant papers were missed, a reference check was completed of the included papers.

2.3. Data extraction

The following data were extracted from the Stage 3 studies: (1) country of publication; (2) study name, if data were from a larger study; (3) population (age, sex, BMI, health condition); (4) device model; (5) wear location (dominant or non-dominant wrist); (6) sampling frequency; (7) epoch length; (8) activity protocol; (9) criterion method and definitions used to classify PA intensity; (10) statistical methods for calibration; (11) statistical methods for validation/cross-validation; and (12) MVPA and SED outcome. Strength of agreement categorisation for kappa and area under the curve statistics (AUC) was undertaken as per accepted standards.[27,28]

Table 1
Criteria[a] for risk of bias assessment of calibration and independent validation studies.

	Study type	Good	Fair	Poor
Sample characteristics	Calibration Independent validation	Study included height, weight, body mass index, ethnicity, resting metabolic rate, maturity stages and variables specific to the clinical condition.	Study included height, weight, body mass index and variables specific to the clinical condition.	Study did not include any descriptive variables other than age and sex.
Accelerometry settings	Calibration Independent validation	Study included accelerometer model, number of axes, placement, sampling frequency, epoch length and any filtering techniques.	Study included accelerometer model, number of axes and placement position.	Study described accelerometer model.
Protocol design	Calibration Independent validation	The calibration study used a mixed protocol combining daily-life activities, laboratory protocol test on a treadmill and free-living assessments. The validation study used a mixed protocol or free-living only.	The study used a mixed protocol (daily-life activities and a treadmill test). For sedentary only studies, the calibration used a mixed protocol (daily-life sedentary activities and walking/running).	The study protocol included a walking or treadmill test.
Criterion method	Calibration	Energy expenditure (including resting metabolic rate).	Heart rate or metabolic equivalent.	Speed or direct observation.
Criterion method	Independent validation	Energy expenditure (including resting metabolic rate).	Heart rate or metabolic equivalent, or direct observation using a validated tool and video recorded.	Speed or direct observation with no video record.
Statistical approach to calibration	Calibration	Machine learning techniques, hierarchical models or multilevel modelling, adjusting for factors related to participants characteristics and to the pathophysiology of the clinical condition to develop the cut-point.	ROC curve analyses.	Linear regression or Individual linear regression.
Statistical approach to cross-validation	Calibration	K-fold cross-validation using different samples and activities. Agreement assessment using Bland–Altman or Kapa score, and estimates the intraclass correlation coefficient, and/or limits of agreement.	Leave-one-out cross-validation and agreement assessment using Bland–Altman or kappa score.	No validation assessment.
Comparison to calibration study	Independent validation	Population characteristics and accelerometer processing settings similar to original study.	Some similarities and some differences to original study in terms of participant characteristics and accelerometry processing settings.	Population characteristics and accelerometer processing settings different to original study.
Statistical approach to independent validation	Independent validation	Independent sample statistical methods (e.g., regression models) or Bland–Altman methods.	ROC curve analysis (Accuracy statistics) or kappa score.	Assessment of percentage agreement or correlation only.

[a] Criteria selected based on Bianchim and colleagues.[24]

2.4. Methodological quality

The methodological quality of each of the included studies was independently evaluated by two of four researchers (CR, ES, MCS, JL). Table 1 presents the criteria for risk of bias assessment of the calibration and independent validation studies.

For calibration studies, the criteria suggested by Bianchim and colleagues[24] were used to assess the risk of bias related to six elements important for accelerometer calibration protocols.

Criteria for the risk of bias assessment of independent validation of accelerometer cut-points do not exist. As a result, the following elements were amended from the calibration criteria[24] for quality assessment of independent validation studies: a solely free-living protocol was deemed a good protocol design; direct observation using a validated tool and video-recording was deemed to be a fair criterion method; the statistical approach to calibration was amended to assess the statistical approach to independent validation; the statistical approach to cross-validation was not relevant and removed; and a sixth element was added to assess the independent validation study's comparison to the calibration study (including participant characteristics, and accelerometer settings and signal processing criteria).[29]

Each element was rated as good, fair, or poor.

3. Results

3.1. Search results

The results of the search are presented in Fig. 1. Following the removal of duplicates, the search identified 33855 records. Of these, 8074 records were identified that included children and adolescents as participants. Following abstract/title exclusions, 3404 records underwent full-text review. Over 50 % (n = 1914) of full-text records reviewed were excluded at Stage 2 because the ActiGraph device was not worn on the wrist. One hundred and ninety-seven papers were identified that converted wrist-worn ActiGraph counts to MVPA or SED using cut-points in children and adolescents. Full-text review of the Stage 2 studies identified 14 calibration studies[12,30–42] and seven independent validation studies.[43–49]

3.2. Calibration studies

Table 2 presents a summary of included studies that calibrated ActiGraph PA counts to MVPA and/or SED cut-points (n = 14). Nine studies calibrated cut-points for MVPA and SED activities[12,30–33,35–38]; 4 studies calibrated cut-points for SED only[39–42]; and one study calibrated cut-points for MVPA only.[34]

Four studies included toddlers (1–3 years)[33,36–38]; six included pre-school aged children (4–5 years)[31,33,35,37–39]; ten included pre-adolescent children (6–12 years)[12,30–32,34,38–42]; and three included adolescent children (13 years and over).[30,32,42] Eight studies included children that spanned across two or more age groups.[30–33,37–39,42] Only one study included children with a disability, cystic fibrosis.[30]

One study placed the accelerometer on the dominant wrist,[32] nine used the non-dominant wrist location,[12,33–38,40,42] and four placed devices on both wrists.[30,31,39,41] All studies used a triaxial ActiGraph, however, different sample frequencies (hertz (Hz)) and epoch lengths were used for each study. Device sampling frequency varied between 30 Hz,[32,35–37,42] 60 Hz[34,40] and 100 Hz.[30,31,33,38,39,41] Epoch length varied between 1-second[31,32,34,40–42] and 30-second[33] with 5-second epochs used most frequently.[12,30,33,35,36,39] Most of the protocols used to assess PA included specific age-appropriate lifestyle and recreational physical and sedentary activities[12,30–32,34–36,38–42] and/or free play/free living activity.[32,33,35–37,40,41] Sedentary activities included watching a video, colouring/writing, and using handheld devices; whilst physical activities included sports including tennis and football; outdoor recreational

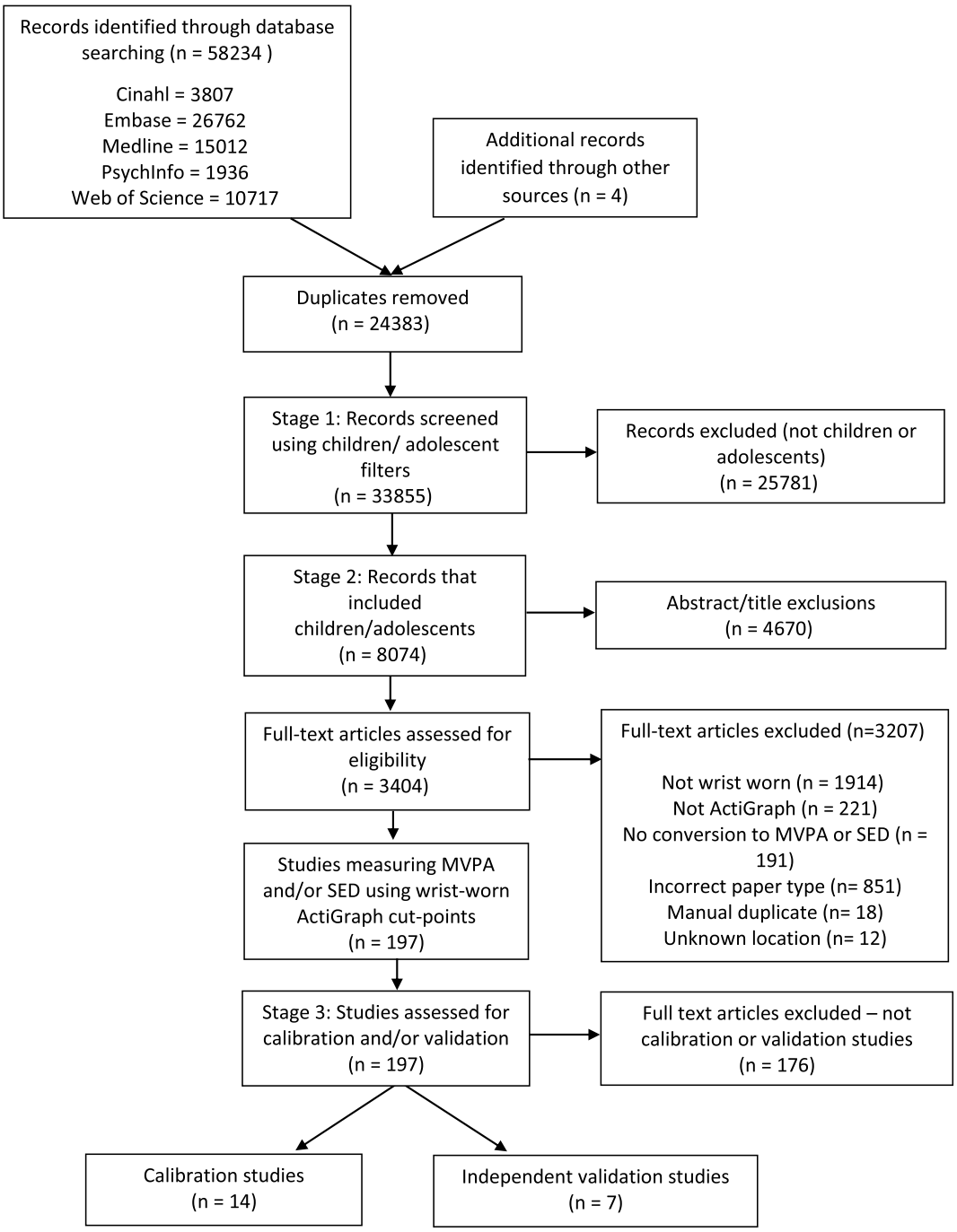

Fig. 1. PRIMSA flow chart for searches from inception through to 30th April, 2022.

activities including walking, skipping, obstacle courses, and dancing; and conventional exercise including treadmill walking, aerobic and cardiovascular exercise.

Eight studies used direct observation as a criterion measure[12,31,33,35,36,38,39,42]; three studies used indirect calorimetry[30,32,34]; and one used heart rate (i.e., percentage of heart rate reserve [HRR]).[12] Two studies calibrated SED cut-points with an activPal monitor as a criterion measure.[40,41] Studies using indirect calorimetry or heart rate specified metabolic equivalent of task (MET) values to classify MVPA, where one MET is generally considered the amount of energy used whilst sitting quietly.[50] Two studies classified MVPA as ≥3.0 METs[32,34]; one study defined MVPA as ≥4.0 METs[30]; and one study defined MVPA as ≥50 % HRR.[12] Five studies used a pre-determined task classification to identify activities as MVPA.[31,33,35,36,38] Three studies classified SED as <1.5 METs.[30,32,42] Nine studies used pre-determined task classification to identify activities as SED.[12,31,33,35,36,38–41] One study evaluated three machine learning methods to define MVPA and SED.[37]

The resultant cut-points for MVPA VM cpm ranged from 7065[37] to 9204[33] and from 63.5 *mg*[30] to 201 *mg*.[34] There was considerable variability in the SED cut-points determined from ≤2556 cpm[37] to ≤4350 cpm[33] and from <30.8 *mg*[30] to ≤48.1 *mg*.[41]

3.3. Independent validation studies

Table 3 presents a summary of included studies that independently validated previously developed MVPA and SED cut-points (n = 7). Four studies independently validated cut-points for MVPA and SED activities[43–45,47]; one study independently validated cut-points for SED

Table 2
Summary of included calibration studies for cut-points classifying time spent sedentary and light- and moderate-to-vigorous physical activity participation in children and adolescents using wrist-worn accelerometers.

Study Author & citation Year Country	Participants • Sample size (n) • Mean age (range) • Sex: % male • Clinical condition	Accelerometer • Device • Sampling frequency (Hz) • Epoch (s) • Wrist placement	Criterion • Criterion method • PA intensity definitions	SED and PA protocol	Statistical approach Calibration Cross-validation Agreement	Outcome (cut-points)		
						SED	LPA	MVPA
Bianchim et al.,[30] 2022 Wales	• n = 63 • 11 (7–17) years • 51 % • Cystic Fibrosis (CF) (n = 35) and typically developing (n = 28)	• GT9X Link • 100 Hz • 5 s • Both	• Indirect calorimetry • SED <1.5 METs • MVPA >4.0 METs	3 lab sessions: 1. RMR (20 min), 2. daily-life activity protocol (6 activities for 3–10 min each): watching a video, colouring/writing, playing on a handheld device, games - tennis, football, skipping, walking, stairs, 3. treadmill exercise test.	• ROC analysis • Leave one out cross-validation • Bland–Altman plot	CF Non-dominant ENMO <38.4 *mg* Dominant ENMO <55.5 *mg* Typically developing Non-dominant <30.8 *mg* Dominant <51.4 *mg*	Not measured	CF Non-dominant ENMO >60.2 *mg* Dominant ENMO >63.0 *mg* Typically developing Non-dominant 65.9 *mg* Dominant > 63.5 *mg*
Chandler et al.,[12] 2016 USA	• n = 25 (calibration); n = 20 (cross validation) • 9 (8–12) years • 49 % • Typically developing	• GT3X+ • NR • 5 s • Non-dominant	• DO (SOPLAY) and heart rate monitor • SED = resting & enrichment activities • LPA = 13.5 % HRR • MVPA >50 % HRR	4 activities (each 10 min): resting, enrichment (arts & crafts), walking, PACER (progressive aerobic cardiovascular endurance) + 3 activities - playing on a playground, splash pad, swimming in a pool.	• ROC analysis for cut-points and regression analyses to create prediction equations • Separate participants • Descriptive	VM < 305 counts/5 s	VM = 306–817 counts/5 s	VM ≥ 818 counts/5 s
Crotti et al.,[31] 2020 England	• n = 22 (calibration); n = 10 (cross validation) • 6.4 (5–7) years • 47 % • NR	• GT9X Link • 100 Hz • 1 s • Both	• DO • SED, LPA & MVPA = tasks nominated by intensity	10 tasks in 60 min. SED: lying whilst watching TV, sitting whilst colouring, sitting and play with a tablet and playing with LEGO. LPA: passive standing. MPA: walking briskly together, throwing and catching. VPA: running, obstacle course run and hopping.	• ROC analysis; Cohen's kappa and equivalency analysis for agreement with criterion method • Separate participant • Cohen's kappa, Bland Altman plots and accuracy statistics	Non-dominant ENMO <36 *mg* Dominant ENMO <39 *mg*	Non-dominant ENMO 36–189 *mg* Dominant ENMO 39–181 *mg*	Non-dominant ENMO ≥189 *mg* Dominant ENMO ≥181 *mg*
Crouter et al.,[32] 2015 USA	• n = 178 (calibration); n = 40 (cross validation) • 12 (8–15) years • 53 % • Able to exercise	• GT3X or GT3X+ • 30 Hz • 1 s • Dominant	• Indirect calorimetry • SED <1.5 METs • LPA 1.5–2.99 METs • MPA > 3.00 METs	Calibration - 2 days: 1 RMR (30 min); 2: lifestyle and sporting activities 8 min each (e.g., reading, sweeping, playing video games, slow track walking, brisk track walking). Validation free play in lab with games and recreational activities available (2 h).	• ROC analysis for cut-points and regression analyses to create prediction equations • Separate participants • One-way ANOVA, mean bias and 95 % prediction intervals	VM < 100 counts/5 s	VM = 101–609 counts/5 s	VM ≥ 610 counts/5 s
Dobell et al.,[33] 2019 UK	• n = 62 • 3.5 (3–4) years • 55 % • Apparently healthy	• GT3X • 100 Hz • 5 s, 10s, 15 s and 30s • Non-dominant	• DO (OSRAC-P) • SED = stationary/motionless • LPA = stationary torso with movement of limbs and slow/easy movement • MVPA = movements classified as moderate and fast	Free play at preschool (1 h).	• ROC analysis • None • Spearman's rank correlation	Epochs 5 s: VM ≤ 3456 cpm 10 s: VM ≤ 4116 cpm 15 s: VM ≤ 4096 cpm 30 s: VM ≤ 4350 cpm	Epochs 5 s: VM = 3457–9203 cpm 10 s: VM = 4117–7793 cpm 15 s: VM = 4097–7743 cpm 30 s: VM = 4351–7711 cpm	Epochs 5 s: VM ≥ 9204 cpm 10 s: VM ≥ 7794 cpm 15 s: VM ≥ 7744 cpm 30 s: VM ≥ 7712 cpm

(continued on next page)

Table 2 (*continued*)

Study Author & citation Year Country	Participants • Sample size (n) • Mean age (range) • Sex: % male • Clinical condition	Accelerometer • Device • Sampling frequency (Hz) • Epoch (s) • Wrist placement	Criterion • Criterion method • PA intensity definitions	SED and PA protocol	Statistical approach Calibration Cross-validation Agreement	Outcome (cut-points) SED	LPA	MVPA
Hildebrand et al.,[34] 2014 Norway	• n = 29 • 8.9 (7–11) years • 53 % • Apparently healthy	• GT3X+ • 60 Hz • 1 s • Non-dominant	• Indirect Calorimetry • MVPA ≥3METs	Laboratory (each 5 min except 10 min for lying down): lying down, sitting, standing, circuit, slow walk, fast walk, step, run.	• Linear regression analyses • Leave one out cross-validation • None	Not measured	Not measured	ENMO ≥201 *mg*
Johansson et al.,[36] (JohanssonA) 2014 Sweden	• n = 26 (calibration); n = 12 (validation) • 2.2 (1–3) years • NR • Ambulatory	• GT3X+ • 30 Hz • 5 s • Non-dominant	• DO (CARS) • SED = minor hand or arm movement • LPA = moving arms, walking a slow pace • High intensity = brisk walking and running	Activities (each 5 min): watching a cartoon, drawing, running an obstacle course; plus 15 min outdoor free play.	• ROC analysis • Separate participants • Spearman's rank correlation	VA ≤ 89 counts/5 s VM ≤ 221 counts/5 s	VA = 90–439 counts/5 s VM = 222–729 counts/5 s	High Intensity VA ≥ 440 counts/5 s VM ≥ 730 counts/5 s
Johansson et al.,[35] (JohanssonB) 2016 Sweden	• n = 30 • 4.0 (4–5) years • 46 % • Able to participate in PA	• GT3X+ • 30 Hz • 5 s • Non-dominant	• DO (CARS) • SED = minor hand or arm movement • LPA = moving arms, walking a slow pace • MVPA = walking briskly and running	Activities (each 4–6 min): watching a cartoon, drawing, dancing to music; plus 15 min outdoor free play.	• ROC analysis • Leave one out cross-validation • Kappa	VA ≤ 178 counts/5 s VM ≤ 328 counts/5 s	VA = 179–870 counts/5 s VM = 329–1392 counts/5 s	VA ≥ 871 counts/5 s VM ≥ 1393 counts/5 s
Li et al.,[37] 2020 USA	• n = 34 • 4 (3–5) years • 41 % • Healthy	• WGT3X-BT • 30 Hz • 15 s • Non-dominant wrist and hip	Three machine learning models • SED: Hip VM < 820 cpm • LPA: Hip VM = 820–3908 cpm • MVPA: Hip VM > 3908 cpm	Free living at day care programme (3 days, 5 h per day)	• Machine learning methods - two supervised using hip derived cut-points as the reference (ROC analysis and ordinal logistic regression model); one unsupervised using the underlying structure of the wrist counts (k-means cluster) • None • Kappa	Supervised 1 VM ≤ 3406 cpm Supervised 2 VM ≤ 5837 cpm Unsupervised VM ≤ 2556 cpm	Supervised 1 VM = 3407–5690 cpm Supervised 2 VM = 5938–14,020 cpm Unsupervised VM = 2557–7067 cpm	Supervised 1 VM ≥ 5691 cpm Supervised 2 VM ≥ 14, 021 cpm Unsupervised VM ≥ 7068 cpm
Trost et al.,[38] 2017 Australia	• n = 11 • 4.8 (3–6) years • 45 % • NR	• GT3X+ • 100 Hz • 15 s • Non-dominant	• DO (CARS) • SED = TV, reading, tablet, quiet play • LPA = art, treasure hunt, clean up • MVPA = self-paced walking and running	2 laboratory sessions of 6 activities (each 4–5 min). 1: Watching TV, reading, standing to do art, walking, play a game with instructor, completing obstacle course. 2: sitting to play computer game, sitting to play with toys,	• ROC analysis • Leave one out cross-validation • Kappa	VA ≤ 349 counts/15 s VM ≤ 625 counts/15 s	Not reported	VA ≥ 1284 counts/15 s VM ≥ 2103 counts/15 s

				treasure hunt, cleaning toys, cycling, running.				
Sedentary only studies								
Chandler et al.,[39] 2018. USA	• n = 100 (calibration); n = 67 (validation) • 8.0 (5–11) years • 58 % • Apparently healthy	• GT3X+ • 100 Hz • 5 s • Both	• DO • Tasks nominated by intensity	45–60 min total with 5 min stations. SED: reading books, sorting cards, cutting and pasting, playing board games, eating snacks, playing with tablets, watching TV, and writing. LPA: walking.	• ROC analysis • Separate participants • Descriptive	Non-dominant • Axis 1 ≤ 203 counts/5 s • VM ≤ 397 counts/5 s Dominant • Axis 1 ≤ 229 counts/5 s • VM ≤ 428 counts/5 s	–	–
Hildebrand et al.,[40] 2017 Norway	• n = 27 • 8.9 (7–11) years • 53 % • Apparently healthy	• GT3X+ • 60 Hz • 1 s • Non-dominant	• activPal • SED = lying and sitting	• SED: 10 min lying down, 5 min sitting. • Activity from light to vigorous intensity (each 5 min). • Free living: 24 h.	• ROC analysis • 10-fold cross validation • Bland Altmann and limits of agreement	ENMO ≤35.6 *mg*	–	–
Hurter et al.,[41] 2018 UK	• n = 27 (calibration); n = 21 (validation) • 10.2 (9–10) years • 37 % • Apparently healthy	• GT9X • 100 Hz • 1 s • Both	• activPal • SED = lying and sitting	Seven stations, 5 min. • SED: resting, TV, tablet, seated activities (playing lego, writing). • Non-sedentary: standing playing mobile phone game, walking at own pace. Free living: 2 days	• ROC analysis • Separate participants • Bland Altmann and limits of agreement	Non-dominant ENMO ≤48.1 *mg* Dominant ENMO ≤55.6 *mg*	–	–
Kim et al.,[42] 2014 USA	• n = 49 • 10.1 (7–13) years • 40.8 % • Apparently healthy	• GT3X+ • 30 Hz • 1 s • Non-dominant	• DO • SED = 1.0 < MET ≤1.5 • LPA = 1.5 < MET ≤3.0 • MPA = 3.0 < MET ≤6.0 • VPA 6.0 < MET	12 from 24 activities, each 5 min with 1 min rest. • SED: sitting, reading, TV, computer, video game. • LPA: slow walking, playing catch, loading and uploading boxes. • MPA: games, chores, sports. • VPA: jogging, stair climbing, cycling at vigorous pace.	• ROC analysis • None • None	VA ≤ 1756 cpm VM ≤ 3958 cpm	–	–

CARS, Children's Activity Rating Scale; CF, cystic fibrosis; cpm, counts per minute; DO, direct observation; ENMO, Euclidean Norm Minus One; HRR, heart rate reserve; LPA, light physical activity; MET, metabolic equivalent; *mg*, milli-gravitational units; MPA, moderate intensity physical activity; MVPA, moderate to Vigorous Activity; NR, not reported; OSRAC-P, Observation System for Recording Activity in Children-Pre-school; RMR, resting metabolic rate; s, second; SED, time spent sedentary; SOPLAY, System for Observing Play and Leisure Activity in Youths; VA, Vertical Axis; VM, Vector Magnitude; VPA, vigorous intensity physical activity.

Table 3
Summary of included independent validation studies for cut-points classifying moderate-to-vigorous physical activity participation in children and adolescents using wrist-worn accelerometers.

Study Author & citation Year Country	Participants • Sample size (n) • Mean age (range) • Sex: % male • Clinical condition	Accelerometer • Device • Sampling Frequency (Hz) • Epoch (s) • Wrist Placement	Criterion • Criterion method • PA definition	PA protocol	Statistical approach	Calibration reference and validation result		
						SED	LPA	MVPA
Ahmadi et al.,[43] 2022 Australia	• n = 10 • 4 (3–5) years • NR • NR	• GT3X+ • 100 Hz • 15 s • Non-dominant	• DO (CARS, video) • SED = stationary/motionless • LPA = stationary/movement of limbs or trunk; slow/easy movement • MVPA = self-paced walking and running	Free play at home, community park or green space (20 min)	• Weighted Kappa statistics • Heat map confusion matrices • 95 % equivalence testing	*Crotti et al.*[31] K = 0.44 Sed recognition accuracy good (86.3 %) *Johansson et al.*[36] K = 0.35 Sed recognition accuracy modest (62.4 %) *Li et al.*[37] K = 0.31 Sed recognition accuracy modest (64.5 %)	*Crotti et al.*[31] K = 0.44 LPA recognition accuracy modest (55.1 %) *Johansson et al.*[36] K = 0.35 MVPA recognition accuracy modest (60.4 %) *Li et al.*[37] K = 0.31 LPA recognition accuracy poor (44.7 %)	*Crotti et al.*[31] K = 0.44 MVPA recognition accuracy poor (48.9 %) *Johansson et al.*[36] K = 0.35 MVPA recognition accuracy poor (40.4 %) *Li et al.*[37] K = 0.31 MVPA recognition accuracy poor (44.1 %)
Altenburg et al.,[44] 2021 Netherlands	• n = 63 • 2.7 (preschool) years • 47 % • NR	• GT3X • 100 Hz • 30s • Dominant	• DO (OSRAC-P) • SED = stationary/motionless and stationary/movement of limbs or trunk • LPA = slow/easy movement • MVPA = movements classified as moderate and fast	Free-living at preschool (60 min)	• ROC analysis • Bias in time estimates, average time in minutes and 95 % confidence interval	*Dobell et al.*[33] AUC (95 % CI) = 0.71 (0.70–0.72) *Johansson et al.*[35] AUC (95 % CI) = 0.72 (0.71–0.73)	*Dobell et al.*[33] AUC (95 % CI) = 0.54 (0.53–0.56) *Johansson et al.*[35] AUC (95 % CI) = 0.53 (0.51–0.54)	*Dobell et al.*[33] AUC (95 % CI) = 0.59 (0.57–0.61) *Johansson et al.*[35] AUC (95 % CI) = 0.50 (0.49–0.52)
Duncan et al.,[45] 2020. NR	• n = 28 • 9.4 (8–11) years • 50 % • Typically developing	• GT3X • 100 Hz • 5 s • Non-dominant	• Indirect calorimetry • SED < 1.5 METs • LPA 1.5–2.99 METs • MPA ≥ 3.00 METs	9 activities (each 5 min): lying supine, playing with lego whilst seated, slow and medium paced walking, medium paced running, throwing and catching, playing football, cycling	• ROC analysis • McNemar's tests for paired proportions to examine sensitivity and specificity differences • De Long test to examine ROC differences at each epoch	*Chandler et al.*[12] VA AUC = 0.73 VM AUC = 0.71	*Chandler et al.*[12] VA AUC = 0.62 VM AUC = 0.61	Without cycling *Chandler et al.*[12] VA AUC = 0.65 VM AUC = 0.67
Hislop et al.,[46] 2016 Scotland	• n = 32 • 4.2 (3–5) years • 66 % • Typically developing	• GT3X+ • NR • 5 s • Non-dominant	• DO (CARS, video) • SED = stationary/motionless • LPA = stationary/movement of limbs or trunk; slow/easy movement • MVPA = walking briskly and running	Free-living preschool (60 min)	• Correlational analysis • Bland–Altman methods	NR	NR	*Johansson et al.*[36] Mean difference (limits of agreement) = −9.3 (−20.0, 1.5) minutes
Kang et al.,[47] 2019 South Korea	• n = 43 • 9.7 (8–12) years • 58 % • Able to participate in PA	• GT3X+ • NR • 5 s • Both[a]	• Indirect Calorimetry • SED <1.5 METs • LPA = 1.5–2.99 METs • MVPA >3 METs	4 groups of activities. 1. Sedentary (29 min): supine, sitting in chair, playing video game, watching TV; 2.	• Cohen's kappa • Sensitivity, specificity and ROC analysis • Mean absolute percentage error	*Chandler et al.*[12] Dominant: AUC (95 % CI) = 0.74 (0.71–0.76); K = 0.43 Non-dominant: AUC	*Chandler et al.*[12] Dominant: AUC (95 % CI) = 0.63 (0.61–0.65); K = 0.15 Non-dominant: AUC	Without cycling *Chandler et al.*[12] Right: AUC (95 % CI) = 0.70 (0.67–0.72); K = 0.40

				free-living activity (6 min): mopping, hand-weight exercise; 3. cycling (6 min): stationary cycling; 4. locomotor activity (15 min): treadmill - increasing speed	(MAPE) and repeated measures ANOVA with Bonferroni post-hoc corrections • Intra-class correlation estimates and 95 % confidence intervals • Cronbach's Alpha	(95 % CI) = 0.70 (0.67–0.72); K = 0.35 *Crouter et al.*[32] Dominant: AUC (95 % CI) = 0.76 (0.74–0.78); K = 0.48 Non-dominant: AUC (95 % CI) = 0.76 (0.73–0.78); K = 0.46	(95 % CI) = 0.58 (0.56–0.60); K = 0.06 *Crouter et al.*[32] Dominant: AUC (95 % CI) = 0.62 (0.58–0.62); K = 0.18 Non-dominant: AUC (95 % CI) = 0.60 (0.58–0.62); K = 0.17	Left: AUC (95 % CI) = 0.59 (0.56–0.62); K = 0.60 *Crouter et al.*[32] Right: AUC (95 % CI) = 0.84 (0.82–0.86); K = 0.60 Left: AUC (95 % CI) = 0.78 (0.75–0.80); K = 0.60
Sacko et al.,[48] 2019 USA	• n = 42 • 8.0 (primary school) years • 50 % • Apparently healthy	• GT3X+ • 100 Hz • 1 s • Both	• Indirect Calorimetry • SED & LPA: NR • MVPA ≥4 METs	Three sessions (each 9 min) of 5 kicks, 5 throws and 5 strikes performed in block fashions for 6, 12 or 30 s intervals	• One-sample *t*-tests. • Bland–Altman analysis. • Mean error and 95 % prediction intervals	NR	PA (i.e., light, moderate, vigorous) *Chandler et al.*[12] • Dominant: NR • Non-dominant males: $\chi 2$ (2, N = 22) = 5.45, p < .05 • Females: NR *Crouter et al.*[32] • Dominant total sample: $\chi 2$ (2, N = 42) = 20.77, p < .01; • Males: $\chi 2$ (2, N = 22) = 19.00, p < .01; and • Females: $\chi 2$ (2, N = 20) = 9.82, p < .01 • Non-dominant: NR	
Sedentary only van Loo et al.,[49] 2017 Australia	• n = 25 (5–8 years); n = 32 (9–12 years) • X • 49 % • Healthy	• GT3X+ • 100 Hz • 5 s, 60 s • Both	• DO (video)	Sedentary (lying down - 10 min) plus 5 min of each activity. • SED: TV, handheld game, writing, colouring, computer game • LPA: getting ready for school, standing, slow walking, dance • MVPA: brisk walk, sports, running, locomotor course	• Equivalence testing • Bland–Altman analysis • Sensitivity, specificity and ROC analysis • Paired sample *t*-tests	5–8 years (VM) Crouter et al.[32] AUC (95 % CI) = 0.77 (0.77–0.78). Chandler et al.[12] AUC (95 % CI) = 0.78 (0.78–0.79). Kim et al.[42] AUC (95 % CI) = 0.85 (0.84–0.86). 9–12 years (VM) Crouter et al.[32] AUC (95 % CI) = 0.73 (0.73–0.74). Chandler et al.[12] AUC (95 % CI) = 0.72 (0.72–0.73). Kim et al.[42] AUC (95 % CI) = 0.82 (0.81–0.83).	–	–

AUC, area under the curve; CARS, Children's Activity Rating Scale; CI, confidence interval; cpm, counts per minute; DO, direct observation; ENMO, Euclidean Norm Minus One; HRR, heart rate reserve; LPA, low intensity physical activity; MET, metabolic equivalent; *mg*, milli-gravitational units; MVPA, moderate to Vigorous Activity; NR, not reported; OSRAC-P, Observation System for Recording Activity in Children-Pre-school; PA, physical activity; RMR, resting metabolic rate; s, second; SED, time spent sedentary; SOPLAY, System for Observing Play and Leisure Activity in Youths; VA, Vertical Axis; VM, Vector Magnitude.

[a] Kang[47] reports dominant/non-dominant hand for SED and LPA and right/left hand (without an indication of dominance) for MVPA.

only[49]; and two studies independently validated cut-points for MVPA[46,48] only.

Two studies undertook validation of MVPA and SED cut-points calibrated from Chandler and colleagues.[12,45,47] Single studies evaluated the validity of MVPA and SED cut-points from Crotti and colleagues,[31,43] Crouter and colleagues,[32,47] Dobell and colleagues,[33,44] Johansson and colleagues (JohanssonA),[36,43] Johansson and colleagues (JohannsonB),[35,44] and Li and colleagues.[37,43]

The MVPA cut-points from JohannsonA[36,46]; Chandler and colleagues[12,48] and Crouter and colleagues[32,48] and the SED cut-points from Crouter and colleagues[32,49]; Chandler and colleagues[12,49] and Kim and colleagues[42,49] were validated in single studies.

No studies were identified that independently validated the MVPA or SED cut-points from Bianchim and colleagues,[30] Chandler and colleagues,[39] Hildebrand and colleagues,[34,40] Hurter and colleagues[41] and Trost and colleagues.[38]

All four studies that evaluated the validity of the Chandler cut-points[12] included pre-adolescent children (6–12 years),[45,47–49] and one concurrently validated the cut-points for 5–8 year old children.[49] Three studies placed devices on both wrists,[47–49] and one on the non-dominant wrist.[45] Similar to the Chandler calibration study,[12] three studies used 5-second epochs,[45,47,49] whilst one used 1-second.[48] One study used both a 5-second epoch and a 60-second epoch.[49] The protocols used to assess PA included specific age-appropriate lifestyle and recreational activities[45,47–49] and/or free play.[47] Three studies used indirect calorimetry to assess energy expenditure[45,47,48] and one study used direct observation as a criterion measure.[49] Poor agreement was shown for two independent populations,[45,47] though kappa statistics showed moderate agreement for left hand wear.[47] Good agreement was shown for SED classification with three independent populations.[45,47,49]

Three studies also evaluated the validity of the Crouter cut-points.[32,47–49] These studies included pre-adolescent children (6–12 years),[47–49] and one study validated the cut-points for 5–8 year old children.[49] All studies placed devices on both wrists.[47–49] Similar to the Crouter calibration study,[32] one study used a 1-second epochs,[48] whilst the remaining two studies used 5-second epochs,[47,49] with one also using a 60-second epoch.[49] The protocols used to assess PA included specific age-appropriate lifestyle and recreational activities,[47–49] with no included free play periods. Two studies used indirect calorimetry to assess energy expenditure[47,48] and one study used direct observation as a criterion measure.[49] For MVPA, AUC statistics showed good agreement and kappa statistics indicated moderate agreement for right-hand wear.[47] Good agreement was also shown for SED classification in two independent populations.[47,49]

Both studies that undertook validation of the JohanssonA[36] cut-points included toddlers (3 years) and pre-school aged children (4–5 years), and both used the non-dominant wrist location.[43,46] The study protocols varied between specific age-appropriate lifestyle and recreational activities[43] and free play.[46] Epoch length varied between 5-second[46] and 15-second.[43] Both protocols used direct observation of free-living activities to assess energy expenditure.[43,46] The results showed a significant difference between calibration cut-points and MVPA for the independent sample.[43,46] For SED, kappa statistics showed fair agreement and heat map confusion matrices showed modest recognition accuracy.[43]

One study also evaluated the validity of MVPA cut-points from Crotti[31] and Li.[37,43] The study protocol included toddlers (3 years) and pre-school aged children (4–5 years), who wore the monitor on their non-dominant wrist, whilst completing specific age-appropriate lifestyle and recreational activities. Moderate agreement for Crotti and fair agreement for Li were reported for both MVPA and SED, respectively.[43] For SED, heat map confusion matrices showed good (Crotti) and modest (Li) recognition accuracy, whilst poor classification accuracy for both these cut-points was reported for MVPA.[43]

One study evaluated the validity of the Dobell[33] and JohanssonB[35] cut-points in pre-school age children using the dominant wrist location and a 30-second epoch during free-living play.[44] Direct observation of free-living activities was used to evaluate energy expenditure. Poor agreement was shown for both the Dobell and JohanssonB MVPA cut-points, and good agreement for both sets of SED cut-points.[44]

One study[49] independently validated the Kim[42] SED cut-points in children aged 5–8 years and 9–12 years. Monitors were worn on both wrists using 5-second and 60-second epochs whilst participants completed specific age-appropriate lifestyle and recreational activities.

Table 4
Quality assessment of calibration and independent validation studies for classifying moderate-to-vigorous physical activity participation in children and adolescents using wrist-worn accelerometers.

Calibration studies	Sample characteristics	Accelerometer settings	Protocol design	PA criterion	Statistical approach for calibration	Statistical approach for cross validation
Bianchim et al.[30]	Good	Good	Fair	Good	Fair	Fair
Chandler et al.[12]	Fair	Good	Fair	Fair	Fair	Fair
Chandler et al.[39]	Poor	Good	Fair	Poor	Fair	Fair
Crotti et al.[31]	Fair	Good	Good	Poor	Fair	Good
Crouter et al.[32]	Good	Good	Good	Good	Fair	Good
Dobell et al.[33]	Fair	Good	Fair	Poor	Fair	Poor
Hildebrand et al.[34]	Good	Good	Fair	Fair	Poor	Fair
Hildebrand et al.[40]	Fair	Good	Good	Poor	Fair	Fair
Hurter et al.[41]	Good	Fair	Good	Poor	Fair	Good
Johansson et al.[36]	Fair	Good	Good	Poor	Fair	Fair
Johansson et al.[35]	Fair	Good	Fair	Poor	Fair	Fair
Kim et al.[42]	Fair	Good	Fair	Poor	Fair	Poor
Li et al.[37]	Fair	Good	Fair	Poor	Good	Poor
Trost et al.[38]	Fair	Good	Fair	Poor	Good	Fair
Independent validation studies	**Sample characteristics**	**Accelerometer settings**	**Protocol design**	**PA criterion**	**Comparison to calibration study**	**Statistical approach to independent validation**
Ahmadi et al.[43]	Poor	Good	Good	Fair	Poor	Fair
Altenburg et al.[44]	Fair	Good	Good	Poor	Fair	Fair
Duncan et al.[45]	Good	Good	Fair	Good	Good	Fair
Hislop et al.[46]	Fair	Fair	Good	Fair	Good	Good
Kang et al.[47]	Good	Fair	Fair	Good	Good	Fair
Sacko et al.[48]	Good	Good	Poor	Good	Fair	Poor
Van Loo et al.[49]	Good	Good	Fair	Fair	Good	Good

Direct observation of free-living activities was used as a criterion measure and excellent agreement was reported in both age groups.[49]

3.4. Quality assessment

The results of the quality assessment of all included studies are presented in Table 4.

One calibration study received a rating of good for five of the six elements.[32] All but one[41] calibration study received a rating of good for accelerometer settings. Whereas, only five calibration studies were rated as good for the protocol design,[31,32,36,40,41] and four for sample characteristics.[30,32,34,41] The ratings for PA criterion measure for the calibration studies were mixed with two studies rated good,[30,32] two rated fair[12,34] and ten rated poor.[31,33,35–42] Two calibration studies scored good for statistical approach for calibration,[37,38] with the majority scored as fair (n = 11[12,30–33,35,36,39–42]). Three calibration studies were rated as good for statistical approach to cross-validation[31,32,41]; eight scored fair[12,30,34–36,38–40]; and three studies did not include cross-validation and were rated poor.[33,37,42]

Quality assessment of the independent validation studies showed that four studies scored good for sample characteristics,[45,47–49] two scored fair[44,46] and one poor.[43] Description of accelerometer settings was rated as good for five studies[43–45,48,49] and fair for two.[46,47] Protocol design was mixed with three studies scored as good,[43,44,46] three fair[45,47,49] and one poor.[48] Ratings for PA criterion were mixed with three studies scored as good,[45,47,48] three fair[43,46,49] and one poor.[44] The comparison of the validation population and accelerometer settings with the calibration study was rated as good for four studies,[45–47,49] fair for two,[44,48] and poor for one.[43] The statistical approach to independent validation was rated as good for two studies,[46,49] fair for four,[43–45,47] and poor for one.[48]

4. Discussion

Wrist-worn accelerometers are proposed to be a valid and acceptable objective measure of PA in children.[13] Increased compliance to monitor wear is attributed to reduced participant discomfort, embarrassment, and burden.[14,15,17] There is currently no clear consensus regarding the estimation of MVPA and SED from ActiGraph cut-point counts from wrist-worn devices for children and adolescents. This systematic review highlights the wide variability in wrist-worn ActiGraph MVPA and SED cut-points, and the inconsistent validity of these cut-points in children. For children and adolescents aged 6–12 years, our results support the use of the MVPA and SED cut-points calibrated by Crouter and colleagues,[32] using an ActiGraph accelerometer placed on the wrist of the dominant hand. This recommendation is based on the good methodological quality of the original calibration study,[32] and independent validation of the cut-points from three studies that showed moderate to good agreement.[47–49] However, it should be noted that variations in accelerometer settings, including monitor placement, sample frequency, and epoch length, may influence the concordance between studies.[22] Therefore, it is recommended that the accelerometer settings of the original calibration study are used[22] i.e., accelerometer placed on the dominant wrist, with an epoch length of 1-second and a sampling frequency of 30 Hz. For toddlers, pre-school aged children and older adolescents, recommendation of specific cut-points is difficult due to an absence of validation in independent populations[30,34,38] or poor statistical agreement when validated in an independent sample.[12,44]

Without independent validation, it is not known whether the calibration results are specific to individual study characteristics, populations, or protocol design.[51,52] Overall, most included calibration studies provided sufficient information for replication regarding the accelerometer model, number of axes, placement, sampling frequency, epoch-length and filtering techniques.[12,30–37,39,40,42–45,48,49] Yet only seven studies were identified that undertook independent validation of calibrated cut-points.[43–49] The classification accuracy determined by these independent validation studies varied, with kappa values ranging from fair to moderate (0.31–0.60)[27] and AUC statistics ranging from poor to good (0.51–0.84)[28] for MVPA and kappa values ranging from 0.31 to 0.44[43] and area under the curve statistics ranging from 0.70[47] to 0.85[49] for SED.

To generate behaviourally valid and population-specific accelerometer cut-points, the protocols administered need to be straight-forward, cost-effective, replicable, and should include an age-appropriate range of typical intermittent and continuous locomotor and free-play activities.[52,53] All calibration and independent validation studies included in this review included age-appropriate lifestyle and recreational activities[12,30–49] and seven studies included free play.[33,37,41,43,44,46,47] The methodological quality of these studies was limited by the lack of objective measure of energy expenditure, with only two calibration[30,32] and three independent validation studies[45,47,48] including indirect calorimetry as objective measure of energy expenditure. The use of an objective biological criterion measure, such as indirect calorimetry, provides quantifiable metrics to compare to accelerometer counts in calibration studies.[51] However, the use of indirect calorimetry is challenging in children particularly during free-living protocols where bulky equipment is likely to influence natural activity. Consequently, direct observation protocols have been adopted that use validated observational tools and video recording to quantify PA. Whilst direct observation methods also provide insight into PA related factors,[9] these methods can be limited by the time required to train observers, code data, and the potential reactivity of the children being observed.[9]

Cross-validation of cut-points in an independent sample of children reduces biases related to sample characteristics and/or calibration protocol design.[54] Cross-validation of the cut-points was undertaken in 11 from 14 studies,[12,30–32,34–36,38,39,41] however only four of these studies included separate participants for cross-validation (e.g., Chandler, Crotti, Crouter, JohannsonA). Only two of the seven independent validation studies were rated as good for the statistical approach undertaken.[46,49] Statistical approaches that include K-fold cross-validation in different samples and activities, and agreement assessments including Bland–Altman or Kappa scores, intraclass correlation coefficient, and/or limits of agreement are recommended.[24,55]

Combining the methods of all studies accepted in the current review, the accelerometer was placed most frequently on the non-dominant wrist (twelve from 21 studies),[12,33–38,40,42,43,45,46] with a device sampling frequency of 100 Hz (11 from 21 studies),[30,31,33,38,39,41,44,45,48,49] and 5-second epochs (ten from 21 studies).[12,30,33,35,36,39,45–47,49] BrØnd and Arvidsson[56] determined that altering the sampling frequency of hip-worn devices influences the number of activity counts measured, with a higher sampling frequency resulting in an increased number of counts. However, the impact of higher sampling frequencies on classification accuracy for wrist-wear is unclear. Results of this review indicate only small differences in classification accuracy for the use of the Crouter[32] MVPA cut-point, despite differences in sampling frequencies from 30 Hz[32] to 100 Hz.[48] Furthermore, as movement patterns of children are typically characterised by short bouts, short epoch lengths (e.g., 1–15-second) should be used to ensure all PAs are captured.[22] Some researchers have suggested the use of non-dominant wrist device wear[22] to minimise activity counts from sedentary activities such as drawing, colouring and playing with mobile devices.[21] However, in this review, the highest classification accuracy was derived from cut-points calibrated and independently validated from a device placed on the dominant wrist.[32] There is a paucity of studies that include monitor placement on both the left and right sides of the body (seven from a total 21 studies[30,31,39,41,47–49]). Studies that calibrated separate cut-points for dominant and non-dominant hands demonstrated small differences in the cut points derived[30,31,39,41] (e.g., 60.2 *mg* for the non-dominant hand and 63.0 *mg* for the dominant hand[30]). Only one study conducted a direct comparison of the classification accuracy derived from dominant and non-dominant wrist-wear with increased classification accuracy for dominant wrist-wear.[47]

The lack of consensus regarding the collection and analysis of accelerometer data has had a significant negative impact on the ability to make comparisons between different devices and studies, specifically determining time spent in MVPA and SED, thereby enabling comparisons to established PA guidelines.[20,57–60] This has a carry-on effect for driving the public health agenda relating to PA in children and adolescents, and the ability to evaluate the efficacy of programmes designed to promote PA participation.

This study is the first to systematically review MVPA and SED cut-points calibrated and validated from wrist-worn ActiGraph accelerometers in children and adolescents. However, limitations exist. First, only one study was identified that included children with a disability. This may be due to the exclusion criteria of walking aides for ambulation.[30] Generalising cut-points developed for non-disabled populations to clinical populations is problematic because of potentially altered resting metabolic rate and energy expenditure for youth with chronic conditions.[61,62] Second, this review intentionally excluded studies that determined PA intensity using methods beyond a count-based cut points approach (e.g., machine learning methods without conversion to cut-points or processing of raw *mg* units). There is some evidence that machine learning models for wrist accelerometer data may produce more accurate predictions of PA intensity than cut-point based methods.[38,63] However, implementation of these algorithms requires access to programming skills and specific software that might be difficult and costly to access in research or clinical settings, therefore limiting availability to health researchers and healthcare professionals.[64] Thirdly, the review focussed on studies that utilised ActiGraph accelerometers only and the studies included in the review utilised different ActiGraph models (e.g., GT3X, GT3X+, GT9X Link). To our knowledge, there are no data available about the cross-model comparability of wrist-worn ActiGraph monitors in children. VM counts and time spent in PA intensities based on cut-points were comparable between wrist-worn GT3X+ and GT9X models in adults,[65] and between hip-worn GT3X-BT and GT9X models in children.[66] Whilst further research is needed to confirm the comparability of data from wrist-worn ActiGraph models in children, it appears that these outcomes may be comparable across different triaxial ActiGraph models. Finally, consistent with the physical activity guidelines for children and adolescents, we specifically reviewed studies that calibrated and validated cut-points for MVPA and/or SED intensity activity. If the included studies, reported or validated cut-points for light-intensity physical activity these were included in the results, however light-intensity physical activity was not a focus of this review.

5. Conclusion

To ensure the good health of children, interventions that promote community-based MVPA and discourage SED are vital. For younger populations, the use of wrist-worn accelerometers to assess MVPA and SED is promoted as an objective and feasible method, with accelerometer cut-points being the most widely used method for accelerometer data analysis. Whilst additional work is required to improve the methodological rigour of protocols used to calibrate and independently validate cut-points, the results of this review support the use of the Crouter and colleagues[26] cut-points for children and adolescents aged 6–12 years with a minimum device sampling frequency of 30 Hz and 1-second epochs, and the device placed on the dominant wrist. Further research is required to independently validate cut-points developed in younger children (<6 years) and older adolescents (>12 years).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsams.2023.11.008.

Funding information

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Confirmation of ethical compliance

The study contained in this paper did not involve human or animal subjects therefore no ethical clearance was required. The systematic review was conducted and reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Protocol registration was with the Prospective Register of Systematic Reviews (PROSPERO; registration number: CRD42020148543).

CRediT authorship contribution statement

- Kelly Clanchy: formal analysis; data curation; writing original draft; writing review and editing; visualisation
- Matthew Stanfield: conceptualisation; methodology; formal analysis; data curation; writing original draft; writing review and editing; visualisation
- Esther Smits: conceptualisation; methodology; formal analysis; data curation; writing original draft; writing review and editing; visualisation
- Jenna Liimatainen: conceptualisation; methodology; formal analysis; data curation; writing original draft; writing review and editing; visualisation
- Carrie Ritchie: conceptualisation; methodology; formal analysis; data curation; writing original draft; writing review and editing; visualisation; supervision; project administration

Declaration of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

There are no acknowledgments for this work. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. Casperson CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985;100:126-131.
2. Timmons BW, Naylor P-J, Pfeiffer KA. Physical activity for pre-school children - how much and how? *Appl Physiol Nutr Metab* 2007;32(S2E). doi:10.1139/H07-112.
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):1-16. doi:10.1186/1479-5868-7-40.
4. Sothern M, Loftin M, Suskind R et al. The health benefits of physical activity in children and adolescents: implications for chronic disease prevention. *Eur J Pediatr* 1999;158:271-274. doi:10.1007/s004310051070.
5. Department of Health and Aged Care. Physical Activity and Exercise Guidelines for All Australians [Author], 2021.
6. World Health Organization. WHO Guidelines on Physical Activity and Sedentary Behaviour [Author], 2020.
7. U.S. Department of Health and Human Services. Physical Activity Guidelines for Americans [Author], 2018.
8. Lettink A, Altenburg TM, Arts J et al. Systematic review of accelerometer-based methods for 24-h physical behavior assessment in young children (0-5 years old). *Int J Behav Nutr Phys Act* 2002;19(1):1-116. doi:10.1186/s12966-022-01296-y.
9. Trost SG. Measurement of physical activity in children and adolescents. *Am J Lifestyle Med* 2007;1(4):299-314. doi:10.1177/1559827607301686.
10. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med* 2001;31:439-454. doi:10.2165/00007256-200131060-00004.
11. Bonomi AG, Goris AH, Yin B et al. Detection of type, duration, and intensity of physical activity using an accelerometer. *Med Sci Sports Exerc* 2009;41(9):1770-1777. doi:10.1249/MSS.0b013e3181a24536.
12. Chandler JL, Brazendale K, Beets MW et al. Classification of physical activity intensities using a wrist-worn accelerometer in 8-12-year-old children. *Pediatr Obes* 2016;11(2):120-127. doi:10.1111/ijpo.12033.
13. 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):245-253. doi:10.1249/MSS.0000000000000771.

14. Scott JJ, Rowlands AV, Cliff DP et al. Comparability and feasibility of wrist- and hip-worn accelerometers in free-living adolescents. *J Sci Med Sport* 2017;20(12):1101-1106. doi:10.1016/j.jsams.2017.04.017.
15. Schaefer SE, Van Loan M, German B. A feasibility study of wearable activity monitors for pre-adolescent school-age children. *Prev Chronic Dis* 2014;22(11):E85. doi:10.5888/pcd11.130262.
16. Troiano RP, McClain JJ, Brychta RJ et al. Evolution of accelerometer methods for physical activity research. *Br J Sports Med* 2014;48:1019-1023. doi:10.1136/bjsports-2014-093546.
17. Phillips LR, Parfitt G, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport* 2013;16(2):124-128. doi:10.1016/j.jsams.2012.05.013.
18. Bagot KS, Matthews SA, Mason M et al. Current, future and potential use of mobile and wearable technologies and social media data in the ABCD study to increase understanding of contributors to child health. *Dev Cogn Neurosci* 2018;32:121-129. doi:10.1016/j.dcn.2018.03.008.
19. Rowlands AV, Cliff DP, Fairclough SJ et al. Moving forward with backward compatibility: translating wrist accelerometer data. *Med Sci Sports Exerc* 2016;48(11):2142-2149. doi:10.1249/MSS.0000000000001015.
20. Bornstein DB, Beets MW, Byun W et al. Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *J Sci Med Sport* 2011;14:6. doi:10.1016/j.jsams.2011.05.007.
21. Gao Z, Liu W, McDonough DJ et al. The dilemma of analyzing physical activity and sedentary behavior with wrist accelerometer data: challenges and opportunities. *J Clin Med* 2021;10(24):5951. doi:10.3390/jcm10245951.
22. Migueles JH, Cadenas-Sanchez C, Ekelund U et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Med* 2017;47(9):1821-1845. doi:10.1007/s40279-017-0716-0.
23. Neishabouri A, Nguyen J, Samuelsson J et al. Quantification of acceleration as activity counts in ActiGraph wearable. *Sci Rep* 2022;12(11958). doi:10.1038/s41598-022-16003-x.
24. Bianchim MS, McNarry MA, Larun L et al. Calibration and validation of accelerometry using cut-points to assess physical activity in paediatric clinical groups: a systematic review. *Prev Med Rep* 2020;19(101142). doi:10.1016/j.pmedr.2020.101142.
25. Rhudy MB, Dreisbach SB, Moran MD et al. Cut points of the Actigraph GT9X for moderate and vigorous intensity physical activity at four different wear locations. *J Sports Sci* 2020;38(5):503-510. doi:10.1080/02640414.2019.1707956.
26. Page MJ, McKenzie JE, Bossuyt PM et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;134:178-189. doi:10.1136/bmj.n71.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159-174.
28. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978;8:283-298. doi:10.1016/s0001-2998(78)80014-2.
29. Bassett DR, Rowlands AV, Trost SG. Calibration and validation of wearable monitors. *Med Sci Sports Exerc* 2012;44:S32-S38. doi:10.1249/MSS.0b013e3182399cf7.
30. Bianchim MS, McNarry MA, Evans R et al. Calibration and cross-validation of accelerometry in children and adolescents with cystic fibrosis. *Meas Phys Educ Exerc Sci* 2023;27(1):51-59. doi:10.1080/1091367X.2022.2065919.
31. Crotti M, Foweather L, Rudd JR et al. Development of raw acceleration cut-points for wrist and hip accelerometers to assess sedentary behaviour and physical activity in 5-7-year-old children. *J Sports Sci* 2020;38(9):1036-1045. doi:10.1080/02640414.2020.1740469.
32. Crouter SE, Flynn JI, Bassett DR. Estimating physical activity in youth using a wrist accelerometer. *Med Sci Sports Exerc* 2015;47(5):944-951. doi:10.1249/MSS.0000000000000502.
33. Dobell AP, Eyre ELJ, Tallis J et al. Examining accelerometer validity for estimating physical activity in pre-schoolers during free-living activity. *Scand J Med Sci Sports* 2019;29(10):1618-1628. doi:10.1111/sms.13496.
34. Hildebrand M, van Hees VT, Hermann Hansen B et al. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc* 2014;46(9):1816-1824. doi:10.1249/MSS.0000000000000289.
35. Johansson E, Larisch L-M, Marcus C et al. Calibration and validation of a wrist- and hip-worn actigraph accelerometer in 4-year-old children. *PloS One* 2016;11(9):e0162436. doi:10.1371/journal.pone.0162436.
36. 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.
37. Li S, Howard JT, Sosa ET et al. Calibrating wrist-worn accelerometers for physical activity assessment in preschoolers: machine learning approaches. *JMIR Form Res* 2020;4(8):e16727-e16738. doi:10.2196/16727.
38. Trost S, Cliff DP, Ahmadi MN et al. Sensor-enabled activity class recognition in preschoolers: hip versus wrist data. *Med Sci Sports Exerc* 2017;50(3):634-641. doi:10.1249/MSS.0000000000001460.
39. Chandler J, Beets M, Saint-Maurice P et al. Wrist-based accelerometer cut-points to identify sedentary time in 5-11-year-old children. *Children* 2018;5(137). doi:10.3390/children5100137.
40. Hildebrand M, Hansen BH, van Hees VT et al. Evaluation of raw acceleration sedentary thresholds in children and adults. *Scand J Med Sci Sports* 2017;27(12):1814-1823. doi:10.1111/sms.12795.
41. Hurter L, Fairclough SJ, Knowles ZR et al. Establishing raw acceleration thresholds to classify sedentary and stationary behaviour in children. *Children* 2018;5(172):1-18. doi:10.3390/children5120172.
42. Kim Y, Lee J-M, Peters BP et al. Examination of different accelerometer cut-points for assessing sedentary behaviours in children. *PloS One* 2014;9(4):e90630. doi:10.1371/journal.pone.0090630.
43. Ahmadi MN, Trost SG. Device-based measurement of physical activity in preschoolers: comparison of machine learning and cut point methods. *PloS One* 2022;17(4):e0266970. doi:10.1371/journal.pone.0266970.
44. Altenburg TM, de Vries L, op den Buijsch R et al. Cross-validation of cut-points in preschool children using different accelerometer placements and data axes. *J Sports Sci* 2022;40(4):379-385. doi:10.1080/02640414.2021.1994726.
45. Duncan MJ, Eyre ELJ, Cox V et al. Cross-validation of Actigraph derived accelerometer cut-points for assessment of sedentary behaviour and physical activity in children aged 8-11 years. *Acta Paediatr* 2020;109(9):1825-1830. doi:10.1111/apa.15189.
46. Hislop J, Palmer N, Anand P et al. Validity of wrist worn accelerometers and comparability between hip and wrist placement sites in estimating physical activity behaviour in preschool children. *Physiol Meas* 2016;37(10):1701-1714. doi:10.1088/0967-3334/37/10/1701.
47. Kang S, Kim Y, Byun W et al. Comparison of a wearable tracker with actigraph for classifying physical activity intensity and heart rate in children. *Int J Environ Res Public Health* 2019;16(15):2663. doi:10.3390/ijerph16152663.
48. Sacko R, McIver K, Brazendale K et al. Comparison of indirect calorimetry- and accelerometry-based energy expenditure during children's discrete skill performance. *Res Q Exerc Sport* 2019;90(4):629-640. doi:10.1080/02701367.2019.1642440.
49. VAN Loo CM, Okely AD, Batterham MJ et al. Wrist accelerometer cut points for classifying sedentary behavior in children. *Med Sci Sports Exerc* 2017;49(4):813-822. doi:10.1249/MSS.0000000000001158.
50. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clin Cardiol* 1990;13(8):555-565. doi:10.1002/clc.4960130809.
51. Kim Y, Beets MW, Welk GJ. Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but…: a systematic review. *J Sci Med Sport* 2012;15(4):311-321. doi:10.1016/j.jsams.2011.12.001.
52. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc* 2005;37:S501-S511. doi:10.1249/01.mss.0000185660.38335.de.
53. Mackintosh KA, Fairclough SJ, Stratton G et al. A calibration protocol for population-specific accelerometer cut-points in children. *PloS One* 2012. doi:10.1371/journal.pone.0036919.
54. Welk GJ, Almeida J, Morss G. Laboratory calibration and validation of the Biotrainer and Actitrac activity monitors. *Med Sci Sports Exerc* 2003;35:1057-1064. doi:10.1249/01.MSS.0000069525.56078.22.
55. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: measures of agreement. *Perspect Clin Res* 2017;8(4):187-191. doi:10.4103/picr.PICR_123_17.
56. Brønd C, Arvidsson D. Sampling frequency affects the processing of ActiGraph raw acceleration data to activity counts. *J Appl Physiol* 2016;120(3):362-369. doi:10.1152/japplphysiol.00628.2015.
57. Cliff DP, Okely AD. Comparison of two sets of accelerometer cut-off points for calculating moderate-to-vigorous physical activity in young children. *J Phys Act Health* 2007;4:510-514. doi:10.1123/jpah.4.4.510.
58. Kim Y, Beets MW, Pate RR et al. The effect of reintegrating Actigraph accelerometer counts in preschool children: comparison using different epoch lengths. *J Sci Med Sport* 2013;16:2. doi:10.1016/j.jsams.2012.05.015.
59. Tucker P. The physical activity levels of preschool-aged children: a systematic review. *Early Child Res Q* 2008;23:4. doi:10.1016/j.ecresq.2008.08.005.
60. Guinhouya CB, Lemdani M, Vilhelm C et al. Actigraph-defined moderate-to-vigorous physical activity cut-off points among children: statistical and biobehavioural relevance. *Acta Paediatr* 2009;98:708-714. doi:10.1111/j.1651-2227.2008.01187.x.
61. Bandini LG, Schoeller DA, Fukagawa NK et al. Body composition and energy expenditure in adolescents with cerebral palsy or myelodysplasia. *Pediatr Res* 1991;29:70-77. doi:10.1203/00006450-199101000-00014.
62. Ramsey BW, Farrell PM, Pencharz P. Nutritional assessment and management in cystic fibrosis: a consensus report. The Consensus Committee. *Am J Clin Nutr* 1992;55:108-116. doi:10.1093/ajcn/55.1.108.
63. Trost SG, Wong WK, Pfeiffer KA et al. Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc* 2012;44:1801-1809. doi:10.1249/MSS.0b013e318258ac11.
64. Trost SG, Brookes DSK, Ahmadi MN. Evaluation of wrist accelerometer cut-points for classifying physical activity intensity in youth. *Front Digit Health* 2022;4:884307. doi:10.3389/fdgth.2022.884307.
65. Clevenger KA, Pfeiffer KA, Montoye AHK. Cross-generational comparability of hip- and wrist-worn ActiGraph GT3X+, wGT3X-BT, and GT9X accelerometers during free-living in adults. *J Sports Sci* 2020;38(24):2794-2802. doi:10.1080/02640414.2020.1801320.
66. Clevenger KA, Pfeiffer KA, Montoye AHK. Cross-generational comparability of raw and count-based metrics from ActiGraph GT9X and wGT3X-BT accelerometers during free-living in youth. *Meas Phys Educ Exerc Sci* 2020;24(3):194-204. doi:10.1080/1091367X.2020.1773827.