Preventive Medicine Reports 19 (2020) 101142

Contents lists available at ScienceDirect

Preventive Medicine Reports

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

Review article

Calibration and validation of accelerometry using cut-points to assess physical activity in paediatric clinical groups: A systematic review

Mayara S. Bianchim[a,1], Melitta A. McNarry[a,*,1], Lillebeth Larun[b], Alan R. Barker[c,1], Craig A. Williams[c,1], Kelly A. Mackintosh[a,1]

[a] School of Sport and Exercise Sciences, Swansea University, Bay Campus, Swansea, Wales, UK
[b] Norwegian Institute of Public Health, Division of Health Services, PO Box 222 Skøyen, N-0213 Oslo, Norway
[c] Children's Health and Exercise Research Centre, University of Exeter, UK

ABSTRACT

Regular physical activity is associated with physiological and psychosocial benefits in both healthy and clinical populations. However, little is known about tailoring the analysis of physical activity using accelerometers to the specific characteristics of chronic conditions. Whilst accelerometry is broadly used to assess physical activity, recommendations on calibration in paediatric clinical groups are warranted. The aim of this systematic review was to provide a critical overview of protocols used to calibrate accelerometry in children and adolescents with clinical conditions, as well as to develop recommendations for calibration and validation of accelerometry in such populations. The search was performed between March to July 2017 using text words and subject headings in six databases. Studies had to develop moderate-to-vigorous intensity physical activity (MVPA) cut-points for paediatric clinical populations to be included. Risk of bias was assessed using a specific checklist. A total of 540,630 titles were identified, with 323 full-text articles assessed. Five studies involving 347 participants aged 9 to 15 years were included. Twenty-four MVPA cut-points were reported across seven clinical conditions, 16 of which were developed for different models of ActiGraph, seven for Actical and one for Tritrac-R3D. Statistical approaches included mixed regression, machine learning and receiver operating characteristic analyses. Disease-specific MVPA cut-points ranged from 152 to 735 counts·15 s^{-1}, with lower cut-points found for inherited muscle disease and higher cut-points associated with intellectual disabilities. The lower MVPA cut-points for diseases characterised by both ambulatory and metabolic impairments likely reflect the higher energetic demands associated with those conditions.

1. Introduction

Regular physical activity (PA) is recommended for children and adolescents to promote health and well-being (World Health Organisation, 2015), irrespective of disease status. However, PA plays a particularly potent role in youth with chronic conditions and is associated with slowing disease progression in conditions such as cerebral palsy (CP; Keawutan et al., 2017; Verschuren et al., 2016). A common issue for children and adolescents with chronic conditions is the tendency to become less physically active with age and disease progression, which can lead to deconditioning and the initiation of a vicious negative spiral involving subsequent reductions in the ability to engage in PA (Durstine et al., 2013; Torpy et al., 2018).

Careful consideration should be given when recommending PA to children and adolescents with some chronic conditions due to the enhanced nutritional, metabolic and energetic requirements associated with the condition or structural disability (West et al., 2019). Children and adolescents with chronic conditions would, therefore, benefit from a greater understanding of the dose–response relationship between PA and health benefits in order to balance this with the potential negative sequalae that could ensue (Riner and Sellhorst, 2013). However, the current recommendation that children aged 5 to 18 years should accumulate, on average, at least 60 min of moderate-to-vigorous physical activity (MVPA) per day across the week (Department of Health and Social Care, 2019) has been developed for non-clinical populations and are therefore likely to have limited applicability to clinical populations. Indeed, a specific clinical guideline would warrant a higher degree of specificity and a cautious assessment of particular risks and benefits for each condition. It is therefore imperative to account for condition-specific factors that could be associated with exercise intolerance and/or an altered physiological response to exercise/PA (Wells et al., 2019).

PA recommendations tailored for children and adolescents with clinical conditions, however, remain sparse (Morris, 2008).

Objective methods used to assess PA, such as accelerometers, are appropriate for clinical settings due to the low participant burden and relatively low cost (Trost and O'Neil, 2014). Accelerometers are capable

* Corresponding author: Fabian Way, Skewen, Swansea SA1 8EN, UK.
E-mail addresses: 890407@swansea.ac.uk (M.S. Bianchim), M.McNarry@swansea.ac.uk (M.A. McNarry), Lillebeth.Larun@fhi.no (L. Larun), A.R.Barker@exeter.ac.uk (A.R. Barker), C.A.Williams@exeter.ac.uk (C.A. Williams), K.Mackintosh@swansea.ac.uk (K.A. Mackintosh).
[1] on behalf of ActiveYouth SRC group.

https://doi.org/10.1016/j.pmedr.2020.101142
Received 25 November 2019; Received in revised form 28 May 2020; Accepted 4 June 2020
Available online 08 June 2020

of detecting patterns of PA accumulation, as well as information on PA frequency and intensity, such as sedentary time (SED), light physical activity (LPA) and MVPA (Welk, 2005). Specifically, accelerometry measures velocity over a specific period of time, which can be translated into intensities of PA by using cut-points (Welk, 2005). However, the generation of these cut-points is highly challenging, for example, even within one type of accelerometer, the MVPA cut-point in healthy youth varies from 400 to 3,600 counts·min^{-1} (Cain et al., 2013). Whilst the accurate assessment of PA levels is particularly important in chronic conditions, inaccurate cut-points can result in over- or under-estimated predictions. Additionally, it is also important to consider the limitations associated with the use of accelerometry. For example, while accelerometry can accurately assess SED, it is not able to differentiate between various sedentary activities (Hurter et al., 2018). Moreover, factors such as brand and placement are likely to have an impact on the prediction of both SED and time spent in different PA intensities (Godfrey et al., 2008).

Amongst the challenges of calibrating accelerometry are the different methods to translate (e.g., PA protocols and criterion method) and interpret (e.g., statistical approach) the accelerometer raw signals into biological and behavioural outcomes (e.g., cut-points). Indeed, a recent systematic review summarising different accelerometry calibration studies in healthy populations acknowledged the lack of cut-points that account for individual characteristics, such as demographic and physiological variations (de Almeida Mendes et al., 2018). A key limitation of generalising cut-points developed for healthy populations to clinical populations is that they will not consider the altered resting metabolic rate (RMR) and higher energy expenditure (EE) for a given activity often evident in youth with chronic conditions (Bandini et al., 1991; Epstein et al., 1989; Ramsey et al., 1992). Whilst some research has sought to calibrate accelerometry in paediatric clinical conditions (Stephens et al., 2016; Trost et al., 2015), the lack of standardisation, wide variability in protocol designs and lack of healthy matched controls limits interpretation (Logan et al., 2016). Indeed, this systematic review can contribute by providing recommendations regarding the most appropriate criterion references, types of activities and statistical analyses to calibrate and cross-validate the cut-points.

The aim of this systematic review was to provide a critical overview of the protocols used to calibrate and validate accelerometry-derived MVPA cut-points in children and adolescents with clinical conditions and identify key parameters and considerations for future research.

2. Methods

This review was performed in accordance with the Preferred Reporting items for Systematic Review and Meta-Analysis statement (Liberati et al., 2009; Moher et al., 2015) and is registered on the International Prospective Register of Systematic Review (PROSPERO registration ID: CRD42016053880).

2.1. Search methods

The search was performed between March and July of 2017 using six databases (PubMed, SPORTDiscus, ScienceDirect, Scopus, ISI Web of Knowledge, Wiley Online Library). A Population Intervention Comparison Outcome (PICO) framework was adopted to build and structure the search; a detailed description of the search protocol is available on the web-appendix. The protocol and search strategy were reviewed by an experienced librarian and a pilot was performed to ensure the suitability of the criteria and search terms. The search terms were in accordance with the 2017 Medical Subject Headings and were inserted as keywords to all the databases and platforms. The search terms were: acceleromet*; acceleromet* AND (validation OR calibration); acceleromet* AND physical activity; wearable monitors AND (calibration OR validation); physical activity AND (calibration OR validation); acceleromet* cut-points; acceleromet* cut-points; energy expenditure AND acceleromet*; and classification AND physical activity intensities. The reference lists of relevant reviews and of all the studies included therein were examined for studies matching the inclusion criteria.

2.2. Eligibility criteria

Studies published in English from the year 2000 which generated MVPA accelerometry cut-points for accelerometry in children and adolescents (5 to 18 years) with any chronic clinical condition (disease of long duration and slow progression; Goodman et al., 2013) were included. Only studies published after the year 2000 were included in order to avoid inclusion of outdated accelerometers. Non-English, non-human and unpublished studies, book chapters, theses, monographs, dissertations and abstracts were not included. Studies in adults, or calibrating for healthy populations, sedentary behaviour or wheelchair users were excluded. Thus, studies using accelerometers along with additional technologies such as a microcontroller were not included.

2.3. Data extraction and management

An EndNote X7 (Clarivate Analytics, US) database was created with potential studies, and the lead author screened all the titles and abstracts. All full-texts selected by the first author (MSB) were screened by two co-authors (MAM and KAM) according to the pre-established inclusion criteria. Supplementary information for each study was consulted when available. In the case of missing information or variables required for completion of the extraction sheet, study authors were contacted, however, no additional data was provided. Data was extracted from the included full-texts by MSB and reviewed by KAM and MAM (Table 1). Any discrepancies were discussed by the three authors until a consensus was reached.

The risk of bias was assessed independently by MSB and MAM using

Table 1
Summary of the data extracted from the included studies.

Data extraction field	Information extracted
Context and participants	The author, year and sample size of the study; participant characteristics such as age, health status, height, weight, BMI, ethnicity; and covariates measured such as self-report questionnaire data and health scales related to disease assessments were extracted.
Study design and methods used	Any information related to the accelerometer, such as accelerometer model (e.g., number of axes); accelerometer placement (e.g., wrist [dominant/non-dominant], hip, chest); accelerometer settings (e.g., epoch, sampling frequency, use of low frequency filter); and data processing decisions (e.g., wear-time criteria) were extracted. Additionally, any information related to the calibration protocol, such as protocol design (e.g., laboratory-based, field-based, daily-life protocol); duration of the protocol; adjustment of specific variables (e.g., age, body mass); performance of individual calibration; criterion measure (e.g., energy expenditure, direct observation, heart rate); resting metabolic rate assessment; statistical approach (e.g., ROC-curve analyses, linear regression, machine learning); validation method (e.g., validation, cross-validation leave-one-out, cross-validation k-fold); and assessment for agreement (e.g., Kappa, Bland-Altman) were also extracted.
Findings	The extracted outcomes were protocol design and cut-points. All the extracted protocols were classified in four categories: laboratory-based (walking or running, over-ground or on a treadmill), free-living (assessment of participant routine), daily-life (daily-life activities performed at the research site), and mixed (at least two of laboratory-based, free-living and daily-life) protocols.
Quality of the study	checklist sheet.

Table 2
Quality and risk assessment criteria according to descriptive variables and study design.

Standard	Poor	Fair	Good
1. Sample Characteristics	Study did not include any descriptive variables other than age and sex.	Study included height, weight, body mass index and variables specific to the clinical condition.	Study included height, weight, body mass index, ethnicity, resting metabolic rate, maturity stages and variables specific to the clinical condition.
2. Accelerometry Settings	Study described accelerometer model.	Study included accelerometer model, number of axes and placement position.	Study included accelerometer model, number of axes, placement, sampling frequency, epoch length and any filtering techniques.
3. Protocol Design	Calibration protocol composed by walking or treadmill test.	Calibration used a mixed protocol (daily-life activities and a treadmill test).	Mixed protocol combining daily-life activities, laboratory protocol test on a treadmill and free-living assessments.
4. Criterion	Speed or direct observation.	Heart rate or metabolic equivalent.	Energy expenditure (including resting metabolic rate estimation*).
5. Statistical Approach for Calibration	Linear regression or Individual linear regression.	ROC curve analyses.	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.
6. Statistical Approach for Validation	No validation assessment.	Leave-one-out cross-validation and agreement assessment using Bland-Altman or kappa score.	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.

ROC: receiver operating characteristic. *The criteria for a valid resting metabolic rate estimation was a minimum of 15 min of steady state, preferably adopting the formula of Weir (1949).

a specific checklist (Table 2) created according to previous recommendations for calibration protocols (Bassett et al., 2012; Freedson et al., 2005; Welk, 2005). This checklist considers six elements of the calibration protocol (sample characteristics, accelerometry settings, criterion measure, statistical approach for calibration, and statistical approach for validation) to rate studies as good, fair or poor according to the criteria described in Table 2. The inter-rater reliability was calculated using Kappa scores with 0.8 as the minimum acceptable inter-rater agreement (McHugh, 2012). Where any discrepancies arose following the risk assessment, all three authors involved in the screening and data extraction (MSB, MAM and KAM) discussed these until a consensus was reached.

A narrative synthesis of the studies was performed due to the heterogeneity of calibration protocols encountered, covering the topics of the protocol design, description of, and adjustment for, disease-specific factors, accelerometry model and settings, criterion measure and the statistical approach for generating and validating the cut-points. All cut-points in counts·min^{-1} were reintegrated to counts·15 s^{-1} epochs, which is commonly used in youth, to allow inter-study comparability.

3. Results

A total of 543,741 titles were found across all databases, with 540,630 titles remaining following the removal of duplicates. Following initial screening, 619 articles were selected by the main author for full-text assessment. In total, 614 studies were subsequently excluded, primarily due to being in a healthy population (279 studies; Fig. 1). A list of all full-text studies that were excluded can be obtained from the correspondent author. Five studies (Clanchy et al., 2011; McGarty et al., 2016; Ryan et al., 2014; Stephens et al., 2016; Trost et al., 2015), including 347, 9–15 year old, participants, with a total of 24 generated MVPA cut-points for seven clinical conditions, were included in the final synthesis. The clinical conditions were: CP, intellectual disabilities, CF, congenital heart diseases (CHD), haemophilia (HE), inherited muscle disease (IMD), juvenile idiopathic arthritis (JIA; Table 3).

The inter-rater Kappa score for risk of bias was 0.80, with authors disagreeing regarding ‘accelerometry settings’, and were resolved after MSB and MAM discussed each point, resulting in a Kappa score of 1. Most studies (n = 4) were classified as fair for sample characteristics, with only one study scoring as good. One study scored as fair, and four as good, for accelerometry settings, with three and two studies classified as fair and good, respectively, for protocol design. For criterion measure, one scored as good, three as fair and one as poor. The majority (n = 4) of the studies scored as fair for statistical approach for calibration, with only one scoring as good. Finally, regarding the statistical approach for validation, three studies scored as fair and two as poor (Table 4).

Quality of life (Varni et al., 2004), maturity status (Emmanuel and Bokor, 2017; Stephens et al., 2016) and a generic health questionnaire (Feldman et al., 1995; Huber et al., 2001) were used as co-variates. Additionally, three studies (Clanchy et al., 2011; Ryan et al., 2014; Trost et al., 2015) used the specific classification system for CP (Gross Motor Function Classification System - GMFCS). Whilst covariates were considered by most of the included studies, only one study (Stephens et al., 2016) adjusted for disease-specific factors when generating the cut-points, although no formal description was provided on the variables included in the model. None of the studies investigated whether the disease-specific factors and participant demographics impacted on the developed cut-points.

3.1. Accelerometers

Sixteen of the included MVPA cut-points were developed for different ActiGraph models (McGarty et al., 2016; Ryan et al., 2014; Stephens et al., 2016; Trost et al., 2015), seven for Actical (Stephens et al., 2016) and one for Tritrac-R3D (RT3; Table 5; (Ryan et al., 2014). This translates to 15 MVPA cut-points derived from the vertical axis (VA) (Clanchy et al., 2011; Stephens et al., 2016) and nine from the vector magnitude (VM; (McGarty et al., 2016; Ryan et al., 2014; Trost et al., 2015). Three studies utilised hip-worn accelerometry on the right side (McGarty et al., 2016; Stephens et al., 2016; Trost et al., 2015) and two studies calibrating for CP placed the accelerometer on the least affected side (Clanchy et al., 2011; Ryan et al., 2014). The sample frequency varied between 1 and 32 Hz, with one study (Clanchy et al., 2011) not specifying this information. Two studies used an epoch of 15 s (Stephens et al., 2016; Trost et al., 2015), with others using 1 s (Clanchy et al., 2011), 10 s (McGarty et al., 2016) and 60 s (Ryan et al., 2014).

3.2. Calibration protocol settings

A daily-life calibration protocol was the most commonly used (n = 3), generating 22 MVPA cut-points, with only two studies utilising a laboratory-based protocol (Clanchy et al., 2011; Ryan et al., 2014). Indirect calorimetry was the most common physiological criterion used

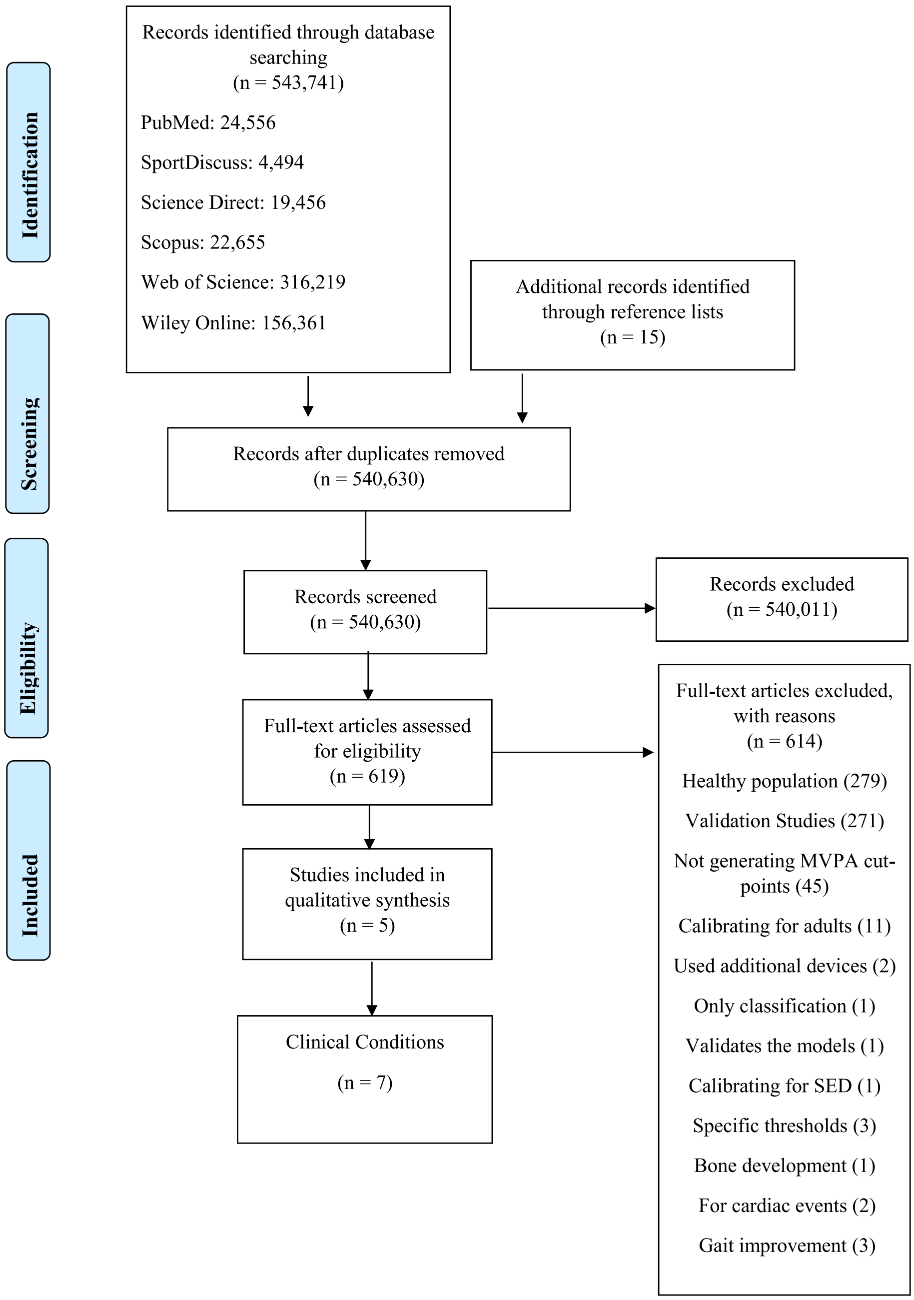

Fig. 1. PRISMA flow chart presenting the systematic literature search.

for calibration (Clanchy et al., 2011; Ryan et al., 2014; Stephens et al., 2016; Trost et al., 2015), with one study using direct observation (McGarty et al., 2016). The protocol duration varied from 35 to 240 min. Resting metabolic rate was estimated by Stephens et al. (2016) using the Weir equation, whereas Clanchy et al. (2011) and Trost et al. (2015) used the Schofield equation and Ryan et al. (2014) the Oxford equation. As McGarty et al. (2016) developed cut-points through direct observation, a RMR estimation was not required. All included studies performed a group calibration rather than individual calibrations.

3.3. Statistical approach

Fourteen MVPA cut-points were developed through mixed regression models (Stephens et al., 2016), six using machine learning (regressing trees; Trost et al., 2015), and four through Receiver Operating Characteristic (ROC) analysis (Clanchy et al., 2011; McGarty et al., 2016; Ryan et al., 2014; Stephens et al., 2016). Only one study did not perform any kind of validation (Clanchy et al., 2011), with all other validations performed using leave-one-out cross-validations. No studies utilised independent samples or a different set of activities to cross-validate. Eighteen (Clanchy et al., 2011; Ryan et al., 2014; Stephens et al., 2016; Trost et al., 2015) of the generated cut-points were validated through comparison of previously established cut-points developed for healthy populations (Evenson et al., 2006; Puyau et al., 2002; Rowlands et al., 2004; Vanhelst et al., 2010). Three studies (McGarty et al., 2016; Ryan et al., 2014; Trost et al., 2015) utilised the Kappa score for agreement assessment, whereas two studies (Clanchy et al., 2011; Stephens et al., 2016) performed ANOVA.

Table 3
Summary of included studies calibrating accelerometry in paediatric clinical groups.

Studies Author, year	Participants Sample size (n) Health status Control Group Sex (boy/girl) Age (range or mean ± SD) Height (mean ± SD) Weight (range or mean ± SD) BMI (range or mean ± SD) Ethnicity Covariates	Accelerometer Device Model Number of axes Placement Sampling frequency Filter Epoch Sampling duration Wear time	Calibration Protocol Physiological/ Observational EE estimation RMR estimation Individual calibration Protocol type Duration	Statistical Approach Calibration Validation Agreement	Outcome Cut-Points/ Equation
Trost et al., 2015	n = 51 Cerebral Palsy GMFCS I (27) GMFCS II (12) GMFCS III (12) Control: 0 28 girls 12 ± 3 years 147.0 ± 16.5 cm 46.8 ± 19.0 kg GMFCS	ActiGraph GT3X Tri-axial Right hip 30 Hz Epoch: 1 s	Physiological: VO_2 Resting VO_2: Schofield Individual calibration: no Protocol type: Mixed – daily-life and walking Duration: 120 min	Calibration: Binary DT Validation: LOOCV Agreement: Kappa and ROC	Cut-points (counts·15 s^{-1}) All levels: LPA: < 72 GMFCS I MVPA: 724 GMFSC II MVPA: 685 GMFCS III MVPA: 669
Ryan et al., 2014	n = 18 Cerebral Palsy Control: no 11.4 ± 3.2 years 147.0 ± 18.5 cm 44.6 ± 16.9 kg BMI: 20 ± 4.5 kg·m^{-2} GMFCS	RT3 Right hip Epoch: 60 s	Physiological: VO_2 RMR: Oxford equation Individual calibration: none Protocol type: laboratory Duration: 36 min	Calibration: ROC curve Validation: none Agreement: Kappa score	Cut-points (counts·min^{-1}): LPA: 52 MVPA: 689.3
Clanchy et al., 2011	n = 29 Cerebral palsy Control: no 13 girls 12.5 ± 2.0 years 156.6 ± 11.0 cm 47.7 ± 16.1 kg GMFCS	ActiGraph (7164) Uniaxial Least affected hip 10 Hz Epoch: 1 s	Physiological: VO_2 RMR: Schofield equation Individual calibration: none Protocol type: laboratory Duration: 60 min	Calibration: ROC curve Validation: none Agreement: none	Cut-points (counts·min^{-1}): LPA: 1627.3 MVPA: 2942.1 VIG: 4683.6
McGarty et al., 2016	n = 50 Validation: 36 Intellectual disabilities Control: no 37 girls 9.54 ± 1.09 years 143 ± 0.9 cm 39.33 ± 10.28 kg BMI: 19.9 ± 3.8 kg·m^{-2}	ActiGraph Wgt3X+ Tri-axial Right hip 30 Hz Epoch: 10 s	Physiological: Direct Observation Individual calibration: none Protocol type: Daily-life Duration: 45 min	Calibration: ROC Validation: LOOCV Agreement: Kappa score	Cut-points (counts·min^{-1}): VA: SED: 507 MPA: 1008–2300 VPA: 2301 MVPA: 1008 VM: SED: 1863 MPA: 2610–4214 VPA: 4215 MVPA: 2610
Stephens et al., 2016	n = 195 Control: n = 29 13 girls 13.1 ± 2.8 years 162 ± 16 cm 57.6 ± 20 kg Skinfold: 38 ± 17 Tanner stages: 30% (stages 1–2), 70% (stage 3) CHAQ: 0.15 ± 0,26 PedsQL: 83 ± 9 Cystic fibrosis (n = 32) 14 girls 12.8 ± 2.9 years 156 ± 16 cm 45 ± 14 kg Skinfold: 31 ± 13 Tanner stage: 19% (stages 1–2), 81% (stage 3) CHAQ: 0.27 ± 0.3 PedsQL: 78 ± 12 Congenital heart disease (n = 15) 5 girls 13.6 ± 3.3 years 161 ± 17 cm 54 ± 17 kg	ActiGraph (7164) and Actical Uniaxial Right hip 10 HZ / 32 Hz Epoch: 15 s	Physiological: VO_2 and HR RMR: 2 h fasting – 20 min in rest Individual calibration: no Protocol type: Mixed: laboratory and daily-life Duration: 240 min	Calibration: Mixed regression models for equation, ROC curve for cut-points. Validation: LOOCV Agreement: none	Chronic disease (combined) – ActiGraph SED: 10 LPA: 10–426 MVPA: 426–785 Chronic disease (combined) – Actical SED: 10 LPA: 17–288 MVPA: 289–570 Cystic fibrosis - ActiGraph SED: 10 LPA: 10–487 MVPA: 487–852 Cystic fibrosis - Actical SED: 5 LPA: 5–368 MVPA: 368–1025 Congenital heart disease - ActiGraph SED: 10 LPA: 10–349 MVPA: 349–785 Congenital heart disease - Actical SED: 9

(*continued on next page*)

Table 3 (*continued*)

Studies Author, year	Participants Sample size (n) Health status Control Group Sex (boy/girl) Age (range or mean ± SD) Height (mean ± SD) Weight (range or mean ± SD) BMI (range or mean ± SD) Ethnicity Covariates	Accelerometer Device Model Number of axes Placement Sampling frequency Filter Epoch Sampling duration Wear time	Calibration Protocol Physiological/ Observational EE estimation RMR estimation Individual calibration Protocol type Duration	Statistical Approach Calibration Validation Agreement	Outcome Cut-Points/ Equation
	Skinfold: 42 ± 15.5 Tanner Stage: 38% (stages 1–2), 62% (stage 3) CHAQ: 0.17 ± 0.3 PedsQL: 72 ± 12 Haemophilia (n = 28) 0 girls 12.4 ± 3.3 years 156 ± 19 cm 53 ± 20.7 kg Skinfold: 40 ± 20 Tanner Stage: 27% (stages 1–2), 73% (stage 3) CHAQ: 0.25 ± 0.4 PedsQL: 82 ± 16 Idiopathic muscular dystrophies (n = 30) 8 girls 12 ± 3.4 years 146 ± 22 cm 41 ± 14 kg Skinfold: 41 ± 18 Tanner stage: 70% (stages 1–2) 30% (stage 3) CHAQ: 0.8 ± 0.7 PedsQL: 68 ± 17 Juvenile dermatomyositis (n = 31) 20 girls 13.4 ± 2.3 years 159 ± 11 cm 52 ± 14 kg Skinfold: 48 ± 17 Tanner stage: 27% (stages 1–2), 73% (stage 3) CHAQ: 0.4 ± 0.6 PedsQL: 77 ± 15 Juvenile arthritis (n = 31) 23 girls 12.7 ± 2.6 years 154 ± 12 cm 47 ± 14 kg Skinfold: 46 ± 22 Tanner stage: 32 (stages 1–2), 68% (stage 3) CHAQ: 0.5 ± 0.5 PedQL: 72 ± 13				LPA: 9–349 MVPA: 349–633 Haemophilia - ActiGraph SED: 17 LPA: 17–432 MVPA: 432–788 Haemophilia - Actical SED: 19 LPA: 19–306 MVPA: 306–1114 Inherited muscle disease - ActiGraph SED: 37 LPA: 37–663 MVPA: 663–972 Inherited muscle disease - Actical SED: 14 LPA: 14–297 MVPA: 297–523 Juvenile dermatomyositis- ActiGraph SED: 14 LPA: 14–172 MVPA: 172–543 Juvenile dermatomyositis - Actical SED: 18 LPA: 10–166 MVPA: 166–601 Juvenile arthritis - ActiGraph SED: 25 LPA: 25–255 MVPA: 255–771 Juvenile arthritis - Actical SED: 19 LPA: 19–152 MVPA: 152–542

SD: standard deviation; BMI: body mass index; EE: energy expenditure; RMR: resting metabolic rate; GMFCS: gross motor function classification system; VO2: oxygen uptake, LOOV: leave-one-out cross-validation; ROC: receiver operating characteristic; SED: sedentary time; LPA: light physical activity; MVPA: moderate-to-vigorous physical activity; VIG: vigorous activity; CHAQ: childhood health assessment questionnaire; PedsQL: pediatric quality of life inventory.

Table 4
Checklist Risk of Bias Assessment Results.

Study	Sample Characteristics	Accelerometry Settings	Protocol Design	Criterion	Statistical Approach for Calibrations	Statistical Approach for Validations
Clanchy et al., 2011	Fair	Fair	Poor	Fair	Fair	Poor
Ryan et al., 2014	Fair	Good	Fair	Poor	Fair	Fair
Trost et al., 2015	Fair	Good	Fair	Fair	Good	Fair
McGarty et al., 2016	Fair	Good	Poor	Fair	Fair	Poor
Stephens et al., 2016	Good	Good	Fair	Good	Fair	Fair

Table 5
Summary of the accelerometer models used by the included studies.

Name / Model	Manufacturer	Weight and Size	Memory Capacity	Axis	Frequency Sampling
ActiGraph 7164 (CSA)	ActiGraph LLC Pensacola, FL	45.5 g 5.1 × 4.1 × 1.5 cm	22 days of data with 60 s epoch	Uniaxial	10 Hz
ActiGraph GT3X	ActiGraph LLC Pensacola, FL	27 g 3.8 × 3.7 × 1.8 cm	378 days using 60 s epoch	Tri-axial	30 Hz
ActiGraph wGT3X+	ActiGraph LLC Pensacola, FL	19 g 4.6 × 3.3 1.5 cm	38 days 100 Hz	Tri-axial	30–100 Hz
Actical	Mini-Mitter Sunriver, OR	17.5 g 2.8 × 2.7 × 1.0 cm	45d using 60 s epoch	Uniaxial	32 Hz
Research Tracker accelerometer (RT3)	StayHealthy, Inc; Monrovia, California	71.5 g 71 × 56 × 28 mm	30 days	Tri-axial	0.017–1 Hz

3.4. Outcome

The disease-specific MVPA cut-points ranged from 152 to 735 counts·15 s^{-1}, with 19 MVPA cut-points presented in counts·15 s^{-1}, and four presented in counts·min^{-1} (Table 6). The sensitivity of the cut-points ranged from 37 to 91%, and the specificity ranged from 85 to 97%. Cerebral palsy was the mostly widely studied clinical condition, with eight cut-points developed across three studies (Clanchy et al., 2011; Ryan et al., 2014; Trost et al., 2015). Trost et al. (2015) generated cut-points for different degrees of CP severity, with fair to excellent accuracy, demonstrating better accuracy (lower rates of misclassification, particularly for GMFCS III and for LPA classification) than Evenson et al. (2006) cut-points. In contrast, Ryan et al. (2014) and Clanchy et al. (2011) did not develop specific cut-points for different GMFCS levels or perform a leave-one-out cross validation, using specificity and sensitivity as a measure of validation. Clanchy et al. (2011) cut-points showed no significant improvement in PA classification accuracy compared to healthy population cut-points, whilst the MVPA cut-points of Ryan et al. (2014) demonstrated moderate classification agreement (Evenson et al., 2006; Rowlands et al., 2004; Vanhelst et al., 2010). Similarly, Stephens et al (2016) also applied healthy population cut-points (Evenson et al., 2006) to their participants with various chronic conditions (CF, IMD, JIA, HE and CHD), which resulted in poor-to-moderate sensitivity in PA classification. Most of the disease-specific cut-points developed were below the previously established MVPA cut-points for healthy populations (e.g., 2,020 to 8,199 counts·min^{-1}).

4. Discussion

Twenty-four MVPA cut-points were extracted from five studies across seven different paediatric clinical groups. Overall, the review revealed little consensus with regards to MVPA cut-points, due to, at least in part, the relatively low number of calibration studies and broad range of protocol designs and accelerometer settings used in the studies, thereby limiting inter-study comparisons. Nonetheless, despite this, a thorough methodological quality assessment of the included studies was performed, which contributed to a higher transparency and aided the interpretation of the outcomes. Moreover, this review presented a critical analysis of the methodological challenges faced when developing cut-points for clinical paediatric populations, providing recommendations for future studies.

4.1. Calibration protocol for paediatric clinical populations

The majority of the included studies utilised daily-life (McGarty et al., 2016) or mixed (Stephens et al., 2016; Trost et al., 2015) protocols composed of daily-life and laboratory protocols. To accommodate different disease and disability levels, Stephens et al. (2016) adjusted their laboratory-based protocol by performing two different treadmill tests based on 6-min walking test performance. Whilst the protocol can greatly impact the PA classification, the physiological criterion adopted is equally important. For example, both Trost et al. (2015) and Stephens et al. (2016) utilised indirect calorimetry as criterion, which therefore considers the higher energetic demand associated with a given activity in some chronic conditions (Walker et al., 2015). Specifically, diseases associated with chronic inflammation (e.g., CF, obesity) and musculoskeletal adaptations (e.g., CP, JIA, IMD) can reduce exercise tolerance, leading to chronic deconditioning and a higher EE demand for a given activity (Mehta, 2015).

It is well known that the majority of paediatric clinical conditions are associated with altered cardiometabolic demands (Bar-Or and Rowland, 2004). Thus, studies calibrating accelerometry for these populations should adopt EE as their criterion method. Another important consideration is that RMR changes dramatically according to maturity, disease and health parameters (McErlane et al., 2017), such as chronic inflammation and reductions in PA (Buchdahl et al., 1988; Eisenstein and Berkun, 2014). Specifically, individuals with CF often have a greater RMR, which can be explained to some extent by pulmonary impairment (Dorlochter et al., 2002) and increased cost of breathing (Bell et al., 1996; Frankenfield et al., 2017). Conversely, children with certain types of CP have a reduced RMR due to a lower energetic requirement at rest and altered body composition (e.g., reduced fat free mass and lean body mass; Bandini et al., 1995, 1991; Stallings et al., 1993). Consequently, condition-specific calibration protocols adopting EE as the criterion should measure RMR. Despite using indirect calorimetry in their protocols, some of the included studies utilised Schofield and Oxford equations (Clanchy et al., 2011) to determine RMR. Whilst such equations may provide a low-cost estimation of RMR, they are based on chronological, rather than biological, age (McMurray et al., 2015), and do not account for sex or health status. This may lead to an inaccurate estimation of RMR, and consequently of EE, in clinical populations (De Wit et al., 2010; Fuster et al., 2007). Therefore, the measurement of oxygen uptake at rest should be utilised to provide a precise estimation of RMR, and consequently enhance the accuracy of the disease-specific cut-points in youth with chronic conditions (Stephens et al., 2016).

It is also important to consider the influence of disease severity within a condition, which is likely to affect the relative energetic demand, as might differences in the treatment and medication strategies between patients (Walker et al., 2015). Indeed, Ryan et al. (2014) and Clanchy et al. (2011) did not stratify their sample by the GMFCS scale, resulting in large heterogeneity of CP-severity across participants, with some children not able to finish the protocol. In contrast, Trost et al. (2015) demonstrated that the relationship between EE and activity counts changed significantly according to GMFCS level, with children classified as level III having greater EE during locomotion when compared to levels I and II.

4.2. Statistical approach

The statistical approach chosen is highly influential in the translation of the physiological criterion into cut-points. Linear regression,

Table 6
Summary and validity of the clinical-specific moderate-to-vigorous cut-points.

Conditions (n)	Study	Reason for split	Cut-points MVPA (original)	Cut-points MVPA converted to counts·15 s^{-1}	Criterion Validity
Cerebral palsy **(7)**	Trost et al., 2015	GMFCS I / VA	535 (counts·15 s^{-1})	N/A	LOOCV – 81.1%
	Trost et al., 2015	GMFCS II / VA	333 (counts·15 s^{-1})	N/A	LOOCV – 76.7%
	Trost et al., 2015	GMFCS III/VA	200 (counts·15 s^{-1})	N/A	LOOCV – 82.9%
	Trost et al., 2015	GMFCS I / VM	724 (counts·15 s^{-1})	N/A	LOOCV – 80.5%
	Trost et al., 2015	GMFCS II / VM	685 (counts·15 s^{-1})	N/A	LOOCV – 75.6%
	Trost et al., 2015	GMFCS III / VM	669 (counts·15 s^{-1})	N/A	LOOCV – 84.2%
	Ryan et al., 2014	N/A	689.3 (counts·min^{-1})	172.3	Se – 86.7% / Sp – 91.9%
	Clanchy et al., 2011	N/A	2942 (counts·min^{-1})	735.5	Se – 91.4% / Sp – 86.2%
Intellectual disability **(2)**	McGarty et al., 2016	VA	1008 (counts·min^{-1})	252	LOOCV – 93% Se – 91% / Sp – 95%
	McGarty et al., 2016	VM	2610 (counts·min^{-1})	652	LOOCV – 87% Se – 91% / 84%
Cystic fibrosis **(2)**	Stephens et al., 2016	CF / ActiGraph 7164	487 (counts·15 s^{-1})	N/A	Se – 71% / Sp – 85%
	Stephens et al., 2016	CF / Actical 7164	368 (counts·15 s^{-1})	N/A	Se – 51% / Sp – 91%
Chronic heart disease **(2)**	Stephens et al., 2016	CHD / ActiGraph 7164	349 (counts·15 s^{-1})	N/A	Se – 42% / Sp – 85%
	Stephens et al., 2016	CHD / Actical	349 (counts·15 s^{-1})	N/A	Se – 41 / Sp – 94%
Inherited muscle disease **(2)**	Stephens et al., 2016	IMD / ActiGraph 7164	663 (counts·15 s^{-1})	N/A	Se – 81% / Sp – 90%
	Stephens et al., 2016	IMD / Actical	297 (counts·15 s^{-1})	N/A	Se – 47% / Sp – 96%
Juvenile dermatomyositis **(2)**	Stephens et al., 2016	JDM /ActiGraph 7164	172 (counts·15 s^{-1})	N/A	Se – 41% / Sp – 90%
	Stephens et al., 2016	JDM / Actical	166 (counts·15 s^{-1})	N/A	Se – 37% / Sp – 94%
Haemophilia (2)	Stephens et al., 2016	HE / Actical	306 (counts·15 s^{-1})	N/A	Se – 49% / Sp – 92%
	Stephens et al., 2016	HE / ActiGraph 7164	432 (counts·15 s^{-1})	N/A	Se – 53% / Sp – 92%
Juvenile arthritis **(2)**	Stephens et al., 2016	JIA / Actical	152 (counts·15 s^{-1})	N/A	Se – 49% / Sp – 94%
	Stephens et al., 2016	JIA / ActiGraph 7164	255 (counts·15 s^{-1})	N/A	Se – 41% / Sp – 90%
Overall (CF, JA, HE, CHD, JDM, IMD) **(2)**	Stephens et al., 2016	Overall Diseases / Actical	289 (counts·15 s^{-1})	N/A	Se – 77% / Se – 97%
	Stephens et al., 2016	Overall Diseases / ActiGraph 7164	426 (counts·15 s^{-1})	N/A	Se – 78% / Sp – 94%

MVPA: moderate-to-vigorous physical activity; GMFCS: gross motor function classification system; VA: vector axial, VM: vector magnitude; LOOCV: leave-one-out cross-validation; Se: sensitivity; Sp: specificity; CF: cystic fibrosis; CHD: congenital heart disease; IMD: inherited muscle disease; JMD: juvenile dermatomyositis; HE: haemophilia; JA: juvenile arthritis.

which was initially one of the most commonly used methods for calibration, cannot account for the non-linear relationship between PA and EE (Freedson et al., 2005; Welk, 2005). Consequently, most of the studies included in this review utilised ROC analyses to develop their cut-points. Whilst ROC is more accurate than linear regression (Welk, 2005), it is dependent on the number of participants and does not allow adjustment of disease-specific factors (Staudenmayer et al., 2009).

Alternatively, mixed regression modelling is an exploratory analysis, particularly useful due to its flexible nature that allows the inclusion of disease-specific factors (Welk, 2005). Stephens et al. (2016) utilised mixed regression modelling to control for disease-specific factors to generate predictive equations for children and adolescents with CF, HE, JIA, CHD and IMD (Aadland and Steene-Johannessen, 2012; Lopes et al., 2009), reporting that heart rate improved the model and lowered the standard error associated with the prediction. These findings agree with those in healthy populations (Altini et al., 2014), with the improvements in standard error likely to be attributable to the reduction of the inter-individual variability caused by the adjustment of physiological signals. It is noteworthy that whilst a certain degree of accuracy can be achieved with cut-points, recent PA research has moved towards using machine learning. Indeed, more complex machine learning analysis have provided a higher degree of accuracy in comparison with traditional cut-points (Bonomi et al., 2009; Staudenmayer et al., 2015, 2009; Welk, 2005). Despite this, a calibration protocol is still required even when using those techniques. Indeed, machine learning can also be used to develop cut-points, for example, Trost et al. (2015) used Binary Decision Trees to generate CP-specific cut-points. Whilst machine learning provides high accuracy, evidence suggests that considerable bias can arise from using a small sample size (Combrisson and Jerbi, 2015). Alternatively, approaches such as using different testing and training data sets, and testing algorithm performance (i.e., nested cross-validation), can provide unbiased performance estimates even with small sample sizes (Vabalas et al., 2019).

A cross-validation analysis of the cut-points evaluates the predictive models to ensure validity and avoid over-fitting, and it can be performed through different methods such as the k-fold or leave-one-out cross-validation. Specifically, considering that the developed cut-points might be biased to the sample characteristics or to the calibration protocol design, the use of an independent sample with a different set of activities for cross-validating the cut-points is recommended (Welk, 2005). Stephens et al. (2016) and Trost et al. (2015) applied a leave-one-out cross-validation, identified as the most appropriate approach when working with smaller samples (Welk et al., 2003), or to lessen the burden on the participants. It is further recommended that the disease-specific cut-points should also be validated against a healthy matched control group to ensure that potential cut-point discrepancies are a result of the pathophysiology rather than from the protocol design. Further to the cross-validation, agreement measures, such as Kappa score and Bland-Altman, indicate whether two methods can be used concomitantly or interchangeably, thereby facilitating inter-study comparisons (Bland and Altman, 1986). Alternatively, recent research has used a statistical equivalence test to measure agreement, which has been shown to be more appropriate for highlighting similarities between methods (Dixon et al., 2018; Kim et al., 2016). Particularly, the performance of agreement measures between activity counts and the criterion measures in a calibration protocol ensures that both measurements are comparable, avoiding further errors to the developed cut-points (Welk, 2005).

4.3. Outcome: cut-points

Cross-validation identified moderate to excellent accuracy for most of the disease-specific cut-points. Considerable inter-study discrepancies were found when comparisons were made between the disease-specific and previously established healthy population cut-points. For example, whilst Trost et al. (2015) found that applying cut-points developed for healthy populations (Evenson et al., 2006) to CP children resulted in poor accuracy and misclassification, Ryan et al. (2014) and Clanchy et al. (2011) demonstrated fair to moderate accuracy (Rowlands et al., 2004; Vanhelst et al., 2010). Indeed, converse to Ryan et al. (2014) and Clanchy et al. (2011), Trost et al. (2015) calibrated for each level of the GMFCS instead of performing an overall calibration, and applied machine learning techniques to generate the CP cut-points, presenting higher specificity than the cut-points developed for healthy populations. Furthermore, Stephens et al. (2016) also found that their disease-specific cut-points (CF, CHD, HE, JIA and IMD) had improved accuracy when compared with standard cut-points, thereby supporting the notion that specific cut-points are necessary for clinical populations.

Given that SED is mainly classified based on stationary activities and therefore does not consider musculoskeletal disabilities, it is unsurprising that some studies (Clanchy et al., 2011; Ryan et al., 2014; Trost et al., 2015) demonstrated fair to excellent accuracy when utilising healthy population-based SED cut-points for children with less severe CP. Despite this, poor classification of LPA may affect specific clinical populations, such as CP (Verschuren et al., 2014), who may not be able to engage in MVPA activities, and would therefore greatly benefit from a reduction in SED (Ryan et al., 2015). Specifically, considering that daily PA is a composite measure, an increase in LPA could be associated with a reduction in SED and enhancement on the total volume of PA (Bassett et al., 2017). Indeed, estimation of LPA for children with CP through standard cut-points, such as Evenson et al. (2006) and Vanhelst et al. (2010), presented poor to fair classification accuracy (Clanchy et al., 2011; Ryan et al., 2014; Trost et al., 2015). Additionally, the lack of standardisation regarding protocol design and statistical approach hinders the applicability of the cut-points, which might explain the variability found between cut-points developed for the same clinical condition. Consequently, age- and sex-matched healthy control groups are essential to elucidate whether the differences observed in the disease-specific protocol are due to the disease severity or to protocol discrepancies. However, only one study (Stephens et al., 2016) included a control group although this was only used for baseline comparisons.

4.4. Strengths and limitations

The present systematic review is associated with numerous strengths. Firstly, an experienced librarian was consulted to revise the initial protocol and a pilot search was conducted to minimise errors, leading to changes in the eligibility of participants, outcomes, risk of bias assessment and analysis. Moreover, the initial search terms were adapted following advice from the librarian. The pilot search generated a large number of studies for participants across the lifespan and health continuum, therefore, the inclusion criteria for participants were limited to only children and adolescents with clinical conditions. Nevertheless, the literature was initially screened to capture all calibration studies for healthy and clinical populations. Whilst this strategy resulted in an extensive search, it also minimised the possibility of missing studies calibrating for a clinical condition. However, this strategy is not without limitations, as it required having only one author screen all the titles and abstracts. Nonetheless, different approaches were adopted to minimise error. Specifically, an EndNote library was created, and the same search strategy was used for all databases. Whilst double data entry was not performed, a data extraction sheet was created and checked by two co-authors (KAM, MAM), and subsequently made available to all authors during the extraction process.

A qualitative data synthesis was performed due the heterogeneity of calibration protocols and the calculation of cut-point effect sizes not being possible, thereby precluding a meta-analysis from being performed. The heterogeneity of the protocols can partially be explained by the inclusion of a broad range of clinical conditions. However, whilst the comparison of numerous clinical conditions of a different nature may be questioned, the primary aim of the review was to investigate the structure of different calibration protocols and how they accounted for the pathophysiology of the respective conditions. Despite the varying nature of the conditions included, only a small range of studies calibrated accelerometry in clinical populations, which hinders further conclusions regarding the optimal protocol.

5. Conclusion

Overall, this systematic review highlights the broad range of protocol designs and accelerometer settings of studies developing MVPA cut-points for children and adolescents with clinical conditions. Research seeking to develop disease-specific paediatric cut-points should consider the pathophysiology of the disease and seek to include a measure of EE, an accurately assessed RMR and a healthy comparison group. Moreover, all cut-points developed should be cross-validated. In summary, studies calibrating accelerometry in paediatric clinical populations are urgently required to establish an optimal calibration protocol. Subsequently, the enhancement in the assessment and surveillance of PA for clinical populations could lead to the development of more informed clinically specific PA guidelines.

Authors' contributions

MSB made substantial contributions to conception, design, systematic search, data analysis and interpretation, and drafted of the manuscript. MAM and KAM made substantial contributions to conception, design, systematic search, data analysis and interpretation, manuscript writing and critically revised the manuscript for important intellectual content. LL supported the design of the search-protocol and critically revised the methodology and general content of the manuscript. AB and CW critically revised the manuscript for important intellectual content. All the authors approved the final manuscript.

Funding

This review summarises independent research funded by the Cystic Fibrosis Trust UK under its programme grant for Strategic Research Centres (grant reference No RP-PG-0108-10011). MSB is a funded PhD student by the Cystic Fibrosis Trust. The funder had no role in the conduct of the study, the writing of the manuscript, or the decision to submit it for publication.

Declaration of Competing Interest

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.

Acknowledgments

We would like to thank the librarian Philippa Price (Swansea University) for her advice on the initial protocol and input on the methods section.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pmedr.2020.101142.

References

Aadland, E., Steene-Johannessen, J., 2012. The use of individual cut points from treadmill walking to assess free-living moderate to vigorous physical activity in obese subjects by accelerometry: is it useful? BMC Med. Res. Methodol. 12, 172.
Altini, M., Penders, J., Vullers, R., Amft, O., 2014. Personalizing energy expenditure estimation using physiological signals normalization during activities of daily living. Physiol. Meas. 35, 1797–1811.
Bandini, L.G., Schoeller, D.A., Fukagawa, N.K., Wykes, L.J., Dietz, W.H., 1991. Body composition and energy expenditure in adolescents with cerebral palsy or myelodysplasia. Pediatr. Res. 29, 70–77.
Bandini, L.G., Puelzl-Quinn, H., Morelli, J.A., Fukagawa, N.K., 1995. Estimation of energy requirements in persons with severe central nervous system impairment. J. Pediatr. 126, 828–832.
Bar-Or, O., Rowland, T.W., 2004. Pediatric Exercise Medicine: From Physiologic Principles to Health Care Application. Human Kinetics.
Bassett, D.R., Rowlands, A.V., Trost, S.G., 2012. Calibration and validation of wearable monitors. Med. Sci. Sports Exerc. 44, S32–S38.
Bassett, D.R., Toth, L.P., LaMunion, S.R., Crouter, S.E., 2017. Step counting: a review of measurement considerations and health-related applications. Sports Med. 1303–1315.
Bell, S.C., Saunders, M.J., Elborn, J.S., Shale, D.J., 1996. Resting energy expenditure and oxygen cost of breathing in patients with cystic fibrosis. Thorax 51, 126–131.
Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1, 307–310.
Bonomi, A.G., Plasqui, G., Goris, A.H.C., Westerterp, K.R., 2009. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. J. Appl. Physiol. 107, 655–661.
Buchdahl, R.M., Cox, M., Fulleylove, C., Marchant, J.L., Tomkins, A.M., Brueton, M.J., Warner, J.O., 1988. Increased resting energy expenditure in cystic fibrosis. J. Appl. Physiol. 1985 (64), 1810–1816.
Cain, K.L., Sallis, J.F., Conway, T.L., Van Dyck, D., Calhoon, L., 2013. Using accelerometers in youth physical activity studies: a review of methods. J. Phys. Act Health 10, 437–450.
Clanchy, K.M., Tweedy, S.M., Boyd, R.N., Trost, S.G., 2011. Validity of accelerometry in ambulatory children and adolescents with cerebral palsy. Eur. J. Appl. Physiol. 111, 2951–2959.
Combrisson, E., Jerbi, K., 2015. Exceeding chance level by chance: the caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. J. Neurosci. Methods 250, 126–136.
de Almeida Mendes, M., da Silva, I.C.M., Ramires, V.V., Reichert, F.F., Martins, R.C., Tomasi, E., 2018. Calibration of raw accelerometer data to measure physical activity: a systematic review. Gait Posture 61, 98–110.
De Wit, B., Meyer, R., Desai, A., Macrae, D., Pathan, N., 2010. Challenge of predicting resting energy expenditure in children undergoing surgery for congenital heart disease. Pediatr. Crit. Care Med. 11, 496–501.
Department of Health and Social Care, 2019. Physical Activity Guidelines: UK Chief Medical Officers' Report. Department of Health and Social Care; London, UK. [(accessed on 02 October 2019)]. Available online: https://www.gov.uk/government/publications/physical-activity-guidelines-uk-chief-medical-officers-report.
Dixon, P.M., Saint-Maurice, P.F., Kim, Y., Hibbing, P., Bai, Y., Welk, G.J., 2018. A primer on the use of equivalence testing for evaluating measurement agreement. Med. Sci. Sports Exerc. 50, 837–845.
Dorlochter, L., Roksund, O., Helgheim, V., Rosendahl, K., Fluge, G., 2002. Resting energy expenditure and lung disease in cystic fibrosis. J. Cyst. Fibros. 1, 131–136.
Durstine, J.L., Gordon, B., Wang, Z., Luo, X., 2013. Chronic disease and the link to physical activity. J. Sport Health Sci. 2, 3–11.
Eisenstein, E.M., Berkun, Y., 2014. Diagnosis and classification of juvenile idiopathic arthritis. J. Autoimmun. 48–49, 31–33.
Emmanuel, M., Bokor, B.R., 2017. Tanner Stages.
Epstein, L.H., Wing, R.R., Cluss, P., Fernstrom, M.H., Penner, B., Perkins, K.A., Nudelman, S., Marks, B., Valoski, 1989. Resting metabolic rate in lean and obese children: relationship to child and parent weight and percent-overweight change. Am. J. Clinical Nutrition 49, 331–336.
Evenson, K.R., Catellier, D.J., Gill, K., Ondrak, K.S., McMurray, R.G., 2006. Calibration of two objective measures of physical activity for children, J. Sports Sci. Dec 2006, p. 1557.
Feldman, B.M., Ayling-Campos, A., Luy, L., Stevens, D., Silverman, E.D., Laxer, R.M., 1995. Measuring disability in juvenile dermatomyositis: validity of the childhood health assessment questionnaire. J. Rheumatol. 22, 326–331.
Frankenfield, D.C., Ashcraft, C.M., Drasher, T.L., Reid, E.K., Vender, R.L., 2017. Characteristics of resting metabolic rate in critically Ill, mechanically ventilated adults with cystic fibrosis. JPEN J. Parenteral Enteral Nutrition 41, 601–606.
Freedson, P., Pober, D., Janz, K.F., 2005. Calibration of accelerometer output for children. Med. Sci. Sports Exerc. 37, S523–S530.
Fuster, C.O., Fuster, G.O., Galindo, A.D., Galo, A.P., Verdugo, J.M., Lozano, F.M., 2007. Analysis of energy expenditure in adults with cystic fibrosis: comparison of indirect calorimetry and prediction equations. Arch. Bronconeumol. 43, 366–372.
Godfrey, A., Conway, R., Meagher, D., G, O.L., 2008. Direct measurement of human movement by accelerometry. Med. Eng. Phys. 30, 1364–1386.
Goodman, R.A., Posner, S.F., Huang, E.S., Parekh, A.K., Koh, H.K., 2013. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. Prev. Chronic Dis. 10, E66.
Huber, A.M., Hicks, J.E., Lachenbruch, P.A., Perez, M.D., Zemel, L.S., Rennebohm, R.M., Wallace, C.A., Lindsley, C.B., Passo, M.H., et al., 2001. Validation of the Childhood Health Assessment Questionnaire in the juvenile idiopathic myopathies. Juvenile Dermatomyositis Disease Activity Collaborative Study Group. J. Rheumatol. 28, 1106–1111.
Hurter, L., Fairclough, S.J., Knowles, Z.R., Porcellato, L.A., Cooper-Ryan, A.M., Boddy, L. M., 2018. Establishing Raw Acceleration Thresholds to Classify Sedentary and Stationary Behaviour in Children. Children (Basel) 5.
Keawutan, P., Bell, K.L., Oftedal, S., Ware, R.S., Stevenson, R.D., Davies, P.S.W., Boyd, R.,

2017. Longitudinal Physical Activity and Sedentary Behaviour in Preschool-Aged Children with Cerebral Palsy Across All Functional Levels. Developmental medicine and child neurology, 59.
Kim, Y., Crouter, S.E., Lee, J.M., Dixon, P.M., Gaesser, G.A., Welk, G.J., 2016. Comparisons of prediction equations for estimating energy expenditure in youth. J. Sci. Med. Sport 19, 35–40.
Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gotzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., et al., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 6, e1000100.
Logan, G.R., Duncan, S., Harris, N.K., Hinckson, E.A., Schofield, G., 2016. Adolescent physical activity levels: discrepancies with accelerometer data analysis. J. Sports Sci. 34, 2047–2053.
Lopes, Vitor Pires, Magalhães, Pedro, Bragada, José, Vasques, Catarina, 2009. Actigraph calibration in obese/overweight and type 2 diabetes mellitus middle-aged to old adult patients. J. Phys. Activity Health 6 (s1), S133–S140. https://doi.org/10.1123/jpah.6.s1.s133.
McErlane, F., Carrasco, R., Kearsley-Fleet, L., Baildam, E.M., Wedderburn, L.R., Foster, H. E., Ioannou, Y., Chieng, S.E.A., Davidson, J.E., et al., 2017. Growth patterns in early juvenile idiopathic arthritis: Results from the Childhood Arthritis Prospective Study (CAPS). Semin Arthritis Rheum.
McGarty, A.M., Penpraze, V., Melville, C.A., 2016. Calibration and Cross-Validation of the ActiGraph wGT3X+ Accelerometer for the Estimation of Physical Activity Intensity in Children with Intellectual Disabilities. PLoS ONE 11, e0164928.
McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 276–282.
McMurray, R.G., Butte, N.F., Crouter, S.E., Trost, S.G., Pfeiffer, K.A., Bassett, D.R., Puyau, M.R., Berrigan, D., Watson, K.B., et al., 2015. Exploring metrics to express energy expenditure of physical activity in youth. PLoS ONE 10, e0130869.
Mehta, N.M., 2015. Energy expenditure: how much does it matter in infant and pediatric chronic disorders? Pediatr. Res. 77, 168–172.
Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A., 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Systematic Rev. 4, 1.
Morris, P.J., 2008. Physical activity recommendations for children and adolescents with chronic disease. Curr. Sports Med. Rep. 7, 353–358.
Puyau, M.R., Adolph, A.L., Vohra, F.A., Butte, N.F., 2002. Validation and calibration of physical activity monitors in children. Obes. Res. 10, 150–157.
Ramsey, B.W., Farrell, P.M., Pencharz, P., 1992. Nutritional assessment and management in cystic fibrosis: a consensus report. The Consensus Committee. Am. J. Clin. Nutr. 55, 108–116.
Riner, W.F., Sellhorst, S.H., 2013. Physical activity and exercise in children with chronic health conditions. J. Sport Health Sci. 2, 12–20.
Rowlands, A.V., Thomas, P.W.M., Eston, R.G., Topping, R., 2004. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. Med. Sci. Sports Exerc. 36, 518–524.
Ryan, J.M., Hensey, O., McLoughlin, B., Lyons, A., Gormley, J., 2015. Associations of sedentary behaviour, physical activity, blood pressure and anthropometric measures with cardiorespiratory fitness in children with cerebral palsy. PLoS ONE.
Ryan, J., Walsh, M., Gormley, J., 2014. Ability of RT3 accelerometer cut points to detect physical activity intensity in ambulatory children with cerebral palsy. Adapted Phys. Activity Quarterly 31, 310–324.
Stallings, V.A., Charney, E.B., Davies, J.C., Cronk, C.E., 1993. Nutrition-related growth failure of children with quadriplegic cerebral palsy. Dev. Med. Child Neurol. 35, 126–138.
Staudenmayer, J., Pober, D., Crouter, S., Bassett, D., Freedson, P., 2009. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J. Appl. Physiol. 107, 1300–1307.
Staudenmayer, J., He, S., Hickey, A., Sasaki, J., Freedson, P., 2015. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. J. Appl. Physiol. 1985 (119), 396–403.
Stephens, S., Pullenayegum, E., Schneiderman, J., McCrindle, B., Abad, A., Ignas, D., Takken, T., Beyene, J., Biggar, D., et al., 2016. Validation of accelerometer prediction equations in children with chronic disease. Pediatric Exercise Sci. 28, 117–132.
Torpy, J.M., Campbell, A., Glass, R.M., 2018. Chronic diseases of children. JAMA 303, 682–782.
Trost, S.G., O'Neil, M., 2014. Clinical use of objective measures of physical activity. Br. J. Sports Med. 48, 178–181.
Trost, S.G., Fragala-Pinkham, M., Lennon, N., O'Neil, M.E., 2015. Decision trees for detection of activity intensity in youth with cerebral palsy. Med. Sci. Sports Exerc. 48, 958–966.
Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PLoS ONE.
Vanhelst, Jeremy, Béghin, Laurent, Rasoamanana, Patrick, Theunynck, Denis, Meskini, Touffik, Iliescu, Catalina, Duhamel, Alain, Turck, Dominique, Gottrand, Frédéric, 2010. Calibration of the RT3 accelerometer for various patterns of physical activity in children and adolescents. J. Sports Sci. 28 (4), 381–387. https://doi.org/10.1080/02640410903508821.
Varni, J.W., Burwinkle, T.M., Szer, I.S., 2004. The PedsQL multidimensional fatigue scale in pediatric rheumatology: reliability and validity. J. Rheumatol. 31, 2494–2500.
Verschuren, O., Darrah, J., Novak, I., Ketelaar, M., Wiart, L., 2014. Health-enhancing physical activity in children with cerebral palsy: more of the same is not enough. Phys. Ther. 94, 297–305.
Verschuren, O., Peterson, M.D., Balemans, A.C., Hurvitz, E.A., 2016. Exercise and physical activity recommendations for people with cerebral palsy. Dev. Med. Child Neurol. 58, 798–808.
Walker, J.L., Bell, K.L., Stevenson, R.D., Weir, K.A., Boyd, R.N., Davies, P.S., 2015. Differences in body composition according to functional ability in preschool-aged children with cerebral palsy. Clin. Nutr. 34, 140–145.
Weir, J.B., 1949. New methods for calculating metabolic rate with special reference to protein metabolism. J. Physiol. 109, 1–9. https://doi.org/10.1113/jphysiol.1949.sp004363.
Welk, G.J., 2005. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med. Sci. Sports Exerc. 37, S501–S511.
Welk, G.J., Almeida, J., Morss, G., 2003. Laboratory calibration and validation of the Biotrainer and Actitrac activity monitors. Med. Sci. Sports Exerc. 35, 1057–1064.
Wells, S.L.W., Laura, B., Jane, E.S., Jessica, E.C., Samantha, S., Gillian, W., Shilpa, D., Greg, D., 2019. Physical activity for children with chronic disease; a narrative review and practical applications. BMC Pediatrics 19, 1–18.
West, S.L., Banks, L., Schneiderman, J.E., Caterini, J.E., Stephens, S., White, G., Dogra, S., Wells, G.D., 2019. Physical activity for children with chronic disease; a narrative review and practical applications. BMC Pediatrics 19, 12.
World Health Organisation, 2015. Physical activity and young people. World Health Organisation.