Calibration and validation of accelerometry to measure physical activity in adult clinical groups: A systematic review

Mayara S Bianchim¹⁎, Melitta A. McNarry³, Lillebeth Larun⁴, Kelly A. Mackintosh⁵, on behalf of ActiveYouth SRC group, Applied Sports Science Technology, Medicine Research Centre

¹ School of Sport and Exercise Sciences, Swansea University, Bay Campus, Fabian Way, SA1 8EN Swansea, Wales, United Kingdom
² Norwegian Institute of Public Health, Division of Health Services, PO Box 222, Skøyen N-0213, Oslo, Norway

A R T I C L E   I N F O

Keywords:
ActiGraph
Disease-specific
Cut-points
Medical conditions
Motion

A B S T R A C T

A growing body of research calibrating and validating accelerometers to classify physical activity intensities has led to a range of cut-points. However, the applicability of current calibration protocols to clinical populations remains to be addressed. The aim of this review was to evaluate the accuracy of the methods for calibrating and validating of accelerometers to estimate physical activity intensity thresholds for clinical populations. Six databases were searched between March and July to 2017 using text words and subject headings. Studies developing moderate-to-vigorous intensity physical activity cut-points for adult clinical populations were included. The risk of bias was assessed using the health measurement instruments and a specific checklist for calibration studies. A total of 543,741 titles were found and 323 articles were selected for full-text assessment, with 11 meeting the inclusion criteria. Twenty-three different methods for calibration were identified using different models of ActiGraph and Actical accelerometers. Disease-specific cut-points ranged from 591 to 2717 counts·min⁻¹ and were identified for two main groups of clinical conditions: neuromusculoskeletal disorders and metabolic diseases. The heterogeneity in the available clinical protocols hinders the applicability and comparison of the developed cut-points. As such, a mixed protocol containing a controlled laboratory exercise test and activities of daily-life is suggested. It is recommended that this be combined with a statistical approach that allows for adjustments according to disease severity or the use of machine learning models. Finally, this review highlights the generalisation of cut-points developed on healthy populations to clinical populations is inappropriate.

1. Background

Physical activity (PA) is defined as any bodily movement that requires an energy expenditure above resting (Caspersen et al., 1985). Regular PA has been associated with the prevention and treatment of a range of diseases, such as cardiovascular disease (Li et al., 2013), type II diabetes (Colberg et al., 2010), osteoporosis (McMillan et al., 2017; Senderovich et al., 2017) and breast cancer (Goncalves et al., 2014). However, 31% of adults are inactive, making physical inactivity a major international public health concern (Hallal et al., 2012).

Although accelerometers are capable of measuring raw acceleration at high sampling frequencies, the majority of studies rely on cut-points to classify PA intensities. Consequently, a growing body of calibration studies has led to a range of cut-points to classify PA intensities in adults (Freedson et al., 1998; Troiano et al., 2008), with little consensus as to the optimal cut-points or their applicability to populations other than those in which they were developed. Indeed, inter-study comparisons and cut-point generalisability are limited by a lack of standardisation of methodologies. Specifically, considerable variation in calibration protocols has arisen due, at least in part, to the progression from uni- to tri-axial accelerometry, the growing range of accelerometer models available and the broad range of configuration options (e.g., epoch, frequency). Furthermore, inter-study discrepancies in moderate-to-vigorous physical activity (MVPA) cut-points may also be attributable to variations in the criterion measures adopted and to the specific calibration protocol utilised; calibration protocols may range from a laboratory-based treadmill or walking protocol (Freedson et al., 2011) to a field-based protocol (Payey et al., 2017), or a combination of both (Midorikawa et al., 2017). Finally, the statistical approach used to translate activity counts into thresholds aligned with the criterion...
Table 1
Summary of the extracted from the included studies.

| Data 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, health scales related to disease assessments |
| **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 anchoring (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) |
| **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 rating for performing calibration for accelerometry in clinical adult population |

varies considerably between studies, with little evidence currently available regarding the comparability of different statistical methods.

A key question that remains to be addressed is the applicability of current calibration protocols to clinical populations. Specifically, physiological and biomechanical differences, common in many chronic conditions such as Chronic Obstructive Pulmonary Disease and Parkinson’s Disease (PD), may result in a higher cost of breathing or daily living activities and altered resting metabolic rate (RMR) demands (Bell et al., 1996; Goldstein et al., 1987; Levi et al., 1990; Psota and Chen, 2013; Sandroff et al., 2014a; Serra et al., 2016). Subsequently, cut-points developed for healthy populations are unlikely to appropriately reflect the activity levels of those with such diseases (McGinley et al., 2015; Serra et al., 2017) and population-specific cut-points, accounting for condition-specific energy expenditure (EE), are warranted. For example, applying cut-points developed on healthy populations was shown to be inappropriate for some clinical conditions, such as chronic stroke (Serra et al., 2017) and type II diabetes (McGinley et al., 2015). However, whilst accelerometry seems to be valid for some clinical conditions (Clarke et al., 2017), the development of population-specific cut-points was shown to improve the accuracy of the PA measurement in multiple sclerosis (MS) and in obese populations (Valenti et al., 2014). Given this lack of consensus, a synthesis of currently available cut-points, and calibration protocols, in clinical populations could afford valuable information for future clinical physical activity research.

Therefore, the aim of this systematic review was to describe current protocols utilised for the calibration of accelerometry to estimate MVPA thresholds for adult clinical populations. Secondly, the purpose was to provide recommendations for future studies seeking to calibrate accelerometers for clinical conditions in adults.

2. Methods

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

3. Search methods

The search was performed between March and July 2017 using six databases (PubMed, SPORTDiscus, ScienceDirect, Scopus, ISI Web of Knowledge, Wiley Online Library). Further details regarding the full search can be found on the Web-Appendix. The protocol was revised by the first author (MSB) and cross-checked by two co-authors (KAM and MAM). Further details of the data extraction are presented in Table 1. The risk of bias in individual studies was assessed by two authors (MSB and KAM), independently, using a checklist that was specifically tailored for calibration of accelerometry protocols (Table 2) based on previous literature (Bassett et al., 2012; Freedson et al., 2005; Lyden et al., 2014; Welk, 2005). This checklist rates studies as good, fair or poor for six elements of the calibration protocol (sample characteristics, accelerometry settings, criterion, statistical approach for calibration, and statistical approach for validation). Studies scoring poor for all the sections were excluded in order to prevent potentially biased and skewed results (Kane et al., 2017). Inter-rater reliability was determined by using Kappa scores, with 0.8 considered the minimum acceptable inter-rater agreement (McHugh, 2012). Following the risk assessment, all three authors discussed any discrepancies until a
were considered by nine studies, clinical conditions were identified. Descriptive characteristics of the study samples are provided in Table 2. A total of 488 participants aged 24–73 years being included in this review. Descriptive characteristics of the study samples are provided in Table 2. Twenty-three studies identified disease-specific cut-points for different conditions. For the final synthesis, the six clinical conditions were stratified into either metabolic (n = 4; obesity, type II diabetes mellitus) or neuromusculoskeletal diseases (n = 7; MS, PD, down syndrome, chronic stroke) (Fig. 1).

Initially the reviewers achieved an inter-rater kappa score of 0.716 for the risk of bias assessment, with the criteria utilised to define RMR the main reasons for disagreement. Subsequently, MSB and MAM resolved discrepancies by discussing each point which resulted in a kappa score of 1 and the criteria to define RMR being specified in the checklist. The majority of the studies had high scores for sample characteristics and accelerometer settings (Table 4), with 5 studies classified as good, five as fair and two as poor for both criteria. Similar results were not found for protocol design, with 10 studies scoring as poor and one as fair. For physiological criterion, 9 studies were classified as fair, 1 as good and 1 as poor. Only two studies scored as good for statistical approach for calibration, with the majority classified as poor (n = 5) and fair (n = 4). Almost all studies were poor (n = 8) for statistical approach for validation, with only 3 studies classified as fair.

Indirect calorimetry was the most common method (n = 10) used to estimate the physiological criterion (e.g. EE, METs or VO2). Covariates were considered by nine studies, five of which utilised disease-specific assessments (e.g., Multiple Sclerosis Walking Scale). Among the studies including covariates, four either included disease-related factors in the analysis or investigated whether the inclusion of those variables would improve the model adopted for calibration. Four studies also included demographic factors in the analysis. Two studies investigated the relationship of the covariates through correlations with accelerometer derived counts·min⁻¹.

### 6. Results

A total of 543,741 titles were identified from all databases, with 540,630 titles remaining after the removal of duplicates. Subsequently, the main author applied the eligibility criteria to all 540,630 titles and abstracts, which resulted in 619 articles remaining for full-text assessment. In total, 608 studies were excluded, primarily due to the inclusion of abstracts, which resulted in 619 articles remaining for full-text assessment.

### 7. Accelerometers

Thresholds were developed for 6 different accelerometers (Table 5); the majority were different models of ActiGraph (n = 9) (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Agiovlasitis et al., 2012; Motl et al., 2009; Nero et al., 2015; Sandroff et al., 2014a, 2014b; Weikert et al., 2011; Lopes et al., 2009) with the others using Actical (Giffuni et al., 2012; Serra et al., 2017). Seventeen of the MVPA cut-points were developed using a uni-axial accelerometer and six using a tri-axial accelerometer. The hip was the most common placement, adopted by nine studies to develop 22 MVPA cut-points. Nine of the MVPA cut-points were developed with the accelerometer placed on the right hip (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Giffuni et al., 2012; Lopes et al., 2009; Motl et al., 2009), seven on non-paretic hip (Sandroff et al., 2014a,b), one on non-paretic hip (Serra et al., 2017), one on both hips (Aadland and Steene-Johannessen, 2012), two on the left hip (Aadland and Steene-Johannessen, 2012). One study placed the accelerometer on the right wrist (Agiovlasitis et al., 2012) and one did not specify the side (Nero et al., 2015). Reported sampling frequency varied from 10 Hz (Weikert et al., 2011) to 30 Hz (Nero et al., 2015; Sandroff et al., 2014a, 2014b), although eight studies did not report the sampling frequency used (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Agiovlasitis et al., 2012; Giffuni et al., 2012; Lopes et al., 2009; Motl et al., 2009; Sandroff et al., 2014b; Serra et al., 2017). Furthermore, only four studies described how they filtered the accelerometer data, with three (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Nero et al., 2015) using the standard filtering provided by the accelerometer software and one (Sandroff et al., 2014b) applying the low-filtering extension provided by ActiLife.
| Studies         | Participants | Accelerometer | Calibration Protocol | Statistical Approach | Outcome Cut-Points |
|----------------|--------------|---------------|----------------------|----------------------|-------------------|
| **Author, year** | **Sample size (n)** | **Device Model** | **Physiological/EE estimation** | **Validation** | **Validation Agreement** |
| **Motl et al., 2009** | n = 48 | 40 females and 8 males | Actigraph | Physiological: VO$_2$, RMR: 3.5 ml.kg$^{-1}$ | Calibration: linear regression |
| Multiple Sclerosis (n = 24) | | Uniaxial | Individual calibration: no | Validation: none | Agreement: none |
| Control (n = 24) | | Right hip | Protocol type: laboratory | | |
| Age (range or mean ± SD) 43.5 ± 12.2 years | | Epoch: 30 s | Duration: 30 min | | |
| Height (mean ± SD) 167.0 ± 11.6 cm | | | | | |
| Weight (range or mean ± SD) 76.7 ± 19.2 kg | | | | | |
| BMI (range or mean ± SD) Demographic scale | | | | | |
| Ethnicity | | | | | |
| Covariates | | | | | |
| **Lopes et al., 2009** | n = 26 | 15 females and 11 males | Actigraph | Physiological: VO$_2$ and HR | Calibration: Hierarchal Model for equation and ROC for cut-points |
| Overweight/obese/ Type 2 Diabetes Mellitus | | Right hip | RMR: 15 min rest | Validation: cross-validation for the regression Agreement: concordance correlation coefficient |
| Control = no | | Epoch: 60 s | Individual calibration: none | | |
| Calibration group (n: 14): Male: 168.07 ± 5.18 cm Female: 151.49 ± 8.54 cm | | | Protocol type: laboratory | | |
| Female: 80.32 ± 7.21 kg Female: 77.05 ± 21.03 kg | | | Duration: 30 min | | |
| 31 ± 5.17 kg·m$^{-2}$ Obese: 57.1% | | | | | |
| Overweight: 42.9% Caucasians | | | | | |
| HBA1c: 7.2 ± 1.8% Insulin: 9.6 ± 4.41 mg.dL$^{-1}$ | | | | | |
| HOMA-IR: 1.59 ± 0.71 | | | | | |
| Validiation group (n = 12): Male: 162.63 ± 3.54 cm Female: 155.1 ± 7.99 cm | | | | | |
| Male: 75.9 ± 16.03 kg Female: 72.19 ± 17.58 kg | | | | | |
| 29.33 ± 4.85 kg·m$^{-2}$ Obese: 41.7% | | | | | |
| Overweight: 58.3% Caucasians | | | | | |
| HBA1c: 7.34 ± 1.61% Insulin: 9.25 ± 4.47 mg.dL$^{-1}$ | | | | | |
| HOMA-IR: 1.53 ± 0.72 | | | | | |
| **Weikert et al., 2011** | n = 24 | 20 females and 4 males | Actigraph (7164) | Physiological: VO$_2$ | Calibration: linear regression |
| Multiple Sclerosis | | Uniaxial | Individual calibration: none | Validation: none | Agreement: none |
| Group with gait disability (n = 10) | | Waist nondominant hip 10 Hz | Protocol type: laboratory | | |
| Group without gait disability (n = 14) | | Epoch: 1 s | Duration: 16 min | | |
| Control: no | | | | | |
| 42 ± 11.7 years | | | | | |
| 20 Caucasian | | | | | |
| 18 graduated from college | | | | | |
| Patient-Determined Disease Steps – 1 (0-4) Multiple Sclerosis Walking Scale | | | | | |
| **Aadland and Andersen, 2012** | n = 42 | 31 females and 11 males | Actigraph (GT1M) | Physiological: VO$_2$ and HR | Calibration: linear regression, linear mixed model and 1. ROC curve with high sensitivity and specificity and 2. ROC with high accuracy. Validation: cross-validation Agreement: none |
| Obesity | | Uniaxial | RMR: 1 h fast – 10 min in rest | | |
| Control: no | | Right hip | Individual calibration: yes | | |
| 43.2 ± 9.2 years | | Normal Filtering Epoch: 10 s | Protocol type: laboratory | | |
| 172.2 ± 9.1 cm | | | Duration: 40 min | | |
| 118.2 ± 18.2 kg | | | | | |
| 39.8 ± 5.7 kg·m$^{-2}$ | | | | | |
| Waist circumference: 127.6 ± 13.2 cm | | | | | |
| (continued on next page)
Table 3 (continued)

| Studies | Participants | Accelerometer | Calibration Protocol | Statistical Approach | Outcome |
|---------|--------------|---------------|----------------------|----------------------|---------|
| Aadland and Steene-Johannessen, 2012 | n = 42 | Actigraph (GT1M) | Physiological: VO₂ and HR | Calibration: Linear regression (individual calibration) and mixed model (group calibration). | Cut-points (counts·min⁻¹): |
| | Control: no | Uniaxial | RMR: 1 h fast – 10 min in rest | Validation: none | Right hip: |
| | 43.2 ± 9.2 years | Right hip and left hip (n: 22) | Individual calibration: yes | Agreement: Bland-Altman | MPVA: 1,078 Left hip: |
| | 172.2 ± 9.1 cm | Normal Filtering | Protocol type: laboratory | | MPVA: 1,095 |
| | 118.2 ± 18.2 kg | Epoch: 10 s | Duration: 40 min | | |
| | 39.8 ± 5.7 kg·m⁻² | Waist circumference: | | | |
| | | | | | |
| | 127.6 ± 13.2 cm | Waist circumference: | | | |
| | | | | | |
| Agiovlasitis et al., 2012 | n = 38 | Actigraph (7164) | Physiological: VO₂ | Calibration: multilevel modelling. | Cut-points (counts·min⁻¹): |
| | 21 females and 27 males | Uniaxial | RMR: 3 h fast – 6 min rest in sitting position. | Validation: none | Control: self-paced walking: |
| | Control | Right wrist | Individual calibration: none | Agreement: Bland-Altman | 2.758 ± 1.373 |
| | 26.3 ± 5.2 years | Epoch: 30 s | Protocol type: laboratory | | 0.5 m/s: 714 ± 279 |
| | 171.1 ± 8.2 cm | | Duration: 30 min | | 0.75 m/s: 1,036 ± 420; |
| | 73.4 ± 22.6 kg | | | | 1 m/s: 1,992 ± 669 |
| | 24.9 ± 7.4 kg·m⁻² | | | | 1.25 m/s: 2,743 ± 1,140 |
| | n = 17 | | | | 1.5 m/s: 3,185 ± 1,140 |
| | Down Syndrome | | | | |
| | 9 females | | | | |
| | 24.7 ± 6.9 years | | | | |
| | 154 ± 79 cm | | | | |
| | 76.9 ± 16.8 kg | | | | |
| | 52.6 ± 7.7 kg·m⁻² | | | | |
| Giffuni et al., 2012 | n = 29 | Actical | Physiological: VO₂ | Calibration: Linear regression. | Cut-points (counts·min⁻¹): |
| | 17 females and 12 males | Uniaxial | RMR: 2 min rest in sitting position. | Validation: none | Obese: |
| | Obese / overweight | Midline of the right | Individual calibration: yes | Agreement: none | 3 METs: 1,923 |
| | 31.9 ± 9 years | tight | Protocol type: laboratory | | 6 METs: 6,032 |
| | 169.1 ± 8.3 cm | Epoch: 60 s | Duration: 45 min | | Control: |
| | 100.8 ± 23.3 kg | | | | 3 METs: 1,137 |
| | 35.2 ± 7.6 kg·m⁻² | | | | 6 METs: 4,525 |
| | VO₂: 29.1 ± 11.5 ml·kg⁻¹·min⁻¹ | | | | |
| | n = 25 | | | | |
| | Control | | | | |
| | 13 males | | | | |
| | 26.1 ± 9.4 years | | | | |
| | 174.3 ± 8.7 cm | | | | |
| | 70 ± 10 kg | | | | |
| | 23 ± 2.2 kg·m⁻² | | | | |
| | VO₂: 40.8 ± 10.2 ml·kg⁻¹·min⁻¹ | | | | |
| | n = 25 | | | | |
| Sandroff et al., 2012 | n = 86 | Actigraph (7164, GT3X) | Physiological: VO₂ | Calibration: linear regression. | Cut-points (counts·min⁻¹): |
| | 76 females and 10 males | Uniaxial and triaxial | Individual calibration: yes | Validation: none | Multiple Sclerosis: MVPA: |
| | Control | Non-dominant hip | Protocol type: laboratory | Agreement: none | 1.723 ± 0.792 |
| | 46.5 ± 10 years | 30 Hz | Duration: 20 min | | Multiple Sclerosis: MVPA: 2,017 ± 801 |
| | 168.5 ± 8.9 cm | Epoch: 15 s | | | GT3X: |
| | 75.4 ± 16.2 kg | | | | Multiple Sclerosis: MVPA: 1,584 ± 697 |
| | n = 43 | | | | Control: 1,950 ± 852 |
| | Multiple Sclerosis | | | | |
| | 47.2 ± 9.1 years | | | | |
| | 168.2 ± 8.3 cm | | | | |
| | 75.7 ± 19.4 kg | | | | |
| | Demographic and exercise history questionnaires | | | | |

(continued on next page)
software. A wide variety of epoch lengths were used to develop the MVPA cut-points, with five studies using 60 s epochs (Giffuni et al., 2012; Lopes et al., 2009; Sandroff et al., 2014b; Serra et al., 2017; Wei kert et al., 2011), followed by one using 10 s (Aadland and Steene-Johannessen, 2012), two studies using 15 s (Nero et al., 2015; Sandroff et al., 2014a, 2014b) and two using 30 s epochs (Agiovlasitis et al., 2012; Motl et al., 2009). The epoch length was extracted from MVPA cut-point unit (i.e. counts·min^-1), counts per 15 s) when not specified in the methodology (Nero et al., 2015).

8. Calibration protocol settings

Laboratory-based protocols were utilised in 10 studies (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Agiovlasitis et al., 2012; Motl et al., 2009; Nero et al., 2015; Sandroff et al., 2014a, 2014b; Wei kert et al., 2012; Giffuni et al., 2012; Lopes et al., 2009), with only one study (Lopes et al., 2009) applying a mixed protocol. Indirect calorimetry was performed by 10 of the studies (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Agiovlasitis et al., 2012; Giffuni et al., 2012; Lopes et al., 2009; Motl et al., 2009; Sandroff et al., 2014a,b; Serra et al., 2017; Weikert et al., 2011), with one study (Nero et al., 2015; Serra et al., 2017) using both indirect calorimetry and HR and another using speed (Nero et al., 2015) and the duration of the protocol varied from 9 to 60 min. Indirect calorimetry was utilised as the physiological criterion by the majority of studies (n = 10). Specifically, six studies derived Metabolic Equivalents of Task (MET) from oxygen uptake (VO₂), whereas four studies used the VO₂ itself to determine the relationship with accelerometry counts. Four studies performed an individual calibration (Aadland and Steene-Johannessen, 2012; Giffuni et al., 2012; Sandroff et al., 2014a,b), five performed a group calibration (Motl et al., 2009; Nero et al., 2015; Serra et al., 2017; Weikert et al., 2011; Lopes et al., 2009) and one study performed both (Aadland and Steene-Johannessen, 2012).

9. Statistical approach

Linear regression was the most common technique employed to generate eight MVPA cut-points in adult clinical populations (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Motl et al., 2009; Sandroff et al., 2014a,b; Serra et al., 2017; Weikert et al., 2011).
Fig. 1. PRISMA flow presenting the systematic literature search.

Records identified through database searching (n = 543,741)
- PubMed: 24,556
- SportDiscuss: 4,494
- Science Direct: 19,456
- Scopus: 22,655
- Web of Science: 316,219
- Wiley Online: 156,361

Additional records identified through other sources (references) (n = 15)

Records after duplicates removed (n = 540,630)

Records screened (n = 540,630)

Full-text articles assessed for eligibility (n = 619)

Studies included in qualitative synthesis (n = 11)

Metabolic Disease (n = 4)
Neuromusculoskeletal Diseases (n = 7)

Records excluded (n = 540,011)

Full-text articles excluded, with reasons (n = 608)
- 279 Healthy population
- 301 Validation Studies
- 15 only validation studies
- 1 calibration to detect postures
- 5 calibrating for children
- 2 used additional devices
  - 1 only classification
  - 1 Calibrating for SED
  - 3 Specific thresholds
  - 1 bone development
  - 2 For cardiac events
  - 3 Gait improvement
2011; Weikert et al., 2012), followed by hierarchical modelling, generating four MVPA cut-points (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Agiovlasitis et al., 2012; Lopes et al., 2009), and receiver operating characteristic (ROC) analysis, developing five MVPA cut-points (Aadland and Anderssen, 2012; Nero et al., 2015; Giffuni et al., 2012). Thus, one study (Aadland and Anderssen, 2012) applied two different ROC models; the first model prioritized higher sensitivity (true positives/total positives) and specificity (true negatives/total negatives), whilst the second model used overall accuracy (true positives and true negatives/total positives and negatives). Ten studies (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Agiovlasitis et al., 2012; Giffuni et al., 2012; Lopes et al., 2009; Motl et al., 2009; Sandroff et al., 2014a,b; Weikert et al., 2011) did not perform any kind of validation and one performed a leave-one-out cross-validation (Nero et al., 2015). Furthermore, most of the studies did not perform any agreement assessment (n = 8) (Aadland and Anderssen, 2012; Aadland and Steene-Johannessen, 2012; Lopes et al., 2009; Motl et al., 2009; Sandroff et al., 2014a,b; Serra et al., 2017; Weikert et al., 2011); one study performed Bland-Altman (Agiovlasitis et al., 2012) and one calculated the Kappa Score (Nero et al., 2015).

11. Discussion

In total, 11 studies generating 23 MVPA cut-points in clinical conditions revealed a broad range of MVPA cut-points. Key recommendations for future studies are to include a variety of free-living activities that are applicable to the specific disease-population, of various intensities, and to ensure that a robust measure of EE and precise estimation of RMR are included to account for disease-related alterations.

12. Calibration protocol for clinical populations

Numerous factors should be considered in the development of a calibration protocol for clinical populations, including the inclusion of participant demographics and disease-related factors. Another key consideration in the development of cut-points for clinical populations is the addition of a physiological criterion to the calibration protocol, particularly related to energetic cost. Specifically, some conditions might be associated with an alteration in the daily total EE. This variation is likely to occur due to many factors, including impaired bio-mechanics (e.g., neuromusculoskeletal disorders), higher energetic cost of breathing (e.g., respiratory conditions) and disease severity and treatments (e.g., medications; (Bell et al., 1996; Psota and Chen, 2013; Montaurier et al., 2007; Nawata et al., 2004; Serra et al., 2015; Wens et al., 2014). Alternatively, METs can be used to estimate EE; Serra et al. (Serra et al., 2017) found METs to be the strongest predictor of activity counts, despite explaining only 65% of

Table 4
Checklist risk of bias assessment results.

| Study                | Sample Characteristics | Accelerometry Settings | Protocol Design | Criterion | Statistical Approach for Calibrations | Statistical Approach for Validations |
|---------------------|------------------------|------------------------|-----------------|-----------|--------------------------------------|-------------------------------------|
| Motl et al., 2009   | Fair                   | Fair                   | Poor            | Fair      | Poor                                 | Poor                                |
| Weikert et al., 2011| Poor                   | Fair                   | Poor            | Fair      | Poor                                 | Poor                                |
| Sandroff et al., 2012| Good                  | Good                   | Poor            | Fair      | Poor                                 | Poor                                |
| Sandroff et al., 2014b| Fair                 | Good                   | Poor            | Poor      | Fair                                 | Poor                                |
| Lopes et al., 2009  | Good                   | Poor                   | Poor            | Fair      | Fair                                 | Poor                                |
| Giffuni et al., 2012| Fair                   | Fair                   | Poor            | Fair      | Poor                                 | Poor                                |
| Aadland and Anderssen, 2012| Good     | Good                   | Poor            | Fair      | Good                                 | Fair                                |
| Aadland and Steene-Johannessen, 2012| Good | Good                   | Poor            | Fair      | Fair                                 | Poor                                |
| Agiovlasitis et al., 2012| Fair   | Fair                   | Poor            | Good      | Good                                 | Poor                                |
| Nero et al., 2015   | Fair                   | Good                   | Poor            | Poor      | Fair                                 | Fair                                |
| Serra et al., 2017  | Good                   | Fair                   | Fair            | Poor      | Poor                                 | Poor                                |

Table 5
Summary of Accelerometer Models Calibrated in the Included Studies.

| Name/Model     | Manufacturer               | Dimensions (Weight and Size) | Memory Capacity          | Axis       | Frequency Sampling  |
|----------------|----------------------------|------------------------------|--------------------------|------------|--------------------|
| ActiGraph 7164 (CSA) | ActiGraph LLC Pensacola, FL | 45.5g 5.1 x 4.1 x 1.5 cm | 22 days of data with 60 s epoch | Uniaxial   | 10 Hz              |
| GT1M ActiGraph  | ActiGraph LLC Pensacola, FL | 27g 3.8 x 3.7 x 1.8 cm     | 378 days using 60 s epoch | Biaxial    | 30 Hz              |
| ActiGraph GT3X  | ActiGraph LLC Pensacola, FL | 27g 3.8 x 3.7 x 1.8 cm     | 378 days using 60 s epoch | Triaxial   | 30 Hz              |
| ActiGraph GT3X+ | ActiGraph LLC Pensacola, FL | 19g 4.6 x 3.3 1.5 cm       | 38 days using 100 Hz     | Triaxial   | 30–100 Hz          |
| ActiGraph wGT3X+ | ActiGraph LLC Pensacola, FL | 19g 4.6 x 3.3 1.5 cm       | 38 days 100 Hz           | Triaxial   | 30–100 Hz          |
| Actical        | Mini-Mitter Sunriver, OR   | 17.5g 2.8 x 2.7 x 1.0 cm   | 45d using 60 s epoch     | Uniaxial   | 32 Hz              |
| Disease (n*)                           | Study                        | Reason for split                | Cut-points MVPA (original) | Cut-points MVPA converted to counts·min⁻¹ | Criterion Validity |
|---------------------------------------|------------------------------|---------------------------------|----------------------------|-------------------------------------------|-------------------|
| Multiple Sclerosis (7)                | Motl et al., 2009            | N/A                             | 6460 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Weikert et al., 2011         | No Gait-disability Group        | 2717 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Weikert et al., 2011         | Overall Group (gait and non-gait-disability) | 2371 (counts·min⁻¹)   | N/A                                      | N/A               |
|                                       | Weikert et al., 2011         | Gait-disability Group           | 1886 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Sandroff et al., 2012        | ActiGraph 7164                  | 1723 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Sandroff et al., 2012        | ActiGraph GT3X                  | 1584 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Sandroff et al., 2014b       | Overall Group (gait and non-gait-disability) | 1745 (counts·min⁻¹)    | N/A                                      | N/A               |
| Overweight/obesity/Type 2 Diabetes Mellitus (10) | Sandroff et al., 2014b       | Gait-disability Group           | 1185 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Lopes et al., 2009           | N/A                             | 2400 (counts·min⁻¹)       | N/A                                      | Concordance Correlation Coefficient: 0.8 |
|                                       | Giffin et al., 2012          | N/A                             | 4012 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Aadland and Andersen, 2012   | ROC 1                           | 1646 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Aadland and Andersen, 2012   | ROC 2                           | 1310 (counts·min⁻¹)       | N/A                                      | N/A               |
|                                       | Aadland and Sjønning-Johannesen, 2012 | Individual Calibration/Linear Regression | 1151 (counts·min⁻¹) | N/A                                      | Bland-Altman/LOA  |
|                                       | Aadland and Sjønning-Johannesen, 2012 | Linear Regression/Left Hip      | 1095 (counts·min⁻¹)       | N/A                                      | Bland-Altman/LOA  |
| Down Syndrome (1)                     | Nero et al., 2015            | N/A                             | 720 (counts·min⁻¹)        | N/A                                      | N/A               |
|                                       | Nero et al., 2015            | MIX REG/Left Hip                | 685 (counts·min⁻¹)        | N/A                                      | Bland-Altman/LOA  |
|                                       | Nero et al., 2015            | MIX REG/Right Hip               | 612 (counts·min⁻¹)        | N/A                                      | N/A               |
| Parkinson disease (2)                 | Nero et al., 2015            | N/A                             | 730 (counts·15 s⁻¹)       | 2880                                     | Cross-validation: 74%–64% of agreement; Kappa Score: 0.79 for y axis and kappa score: 0.69 for VM. |
| Chronic Stroke (1)                    | Serra et al., 2017           | N/A                             | 851 (counts·15 s⁻¹)       | 3404                                     | N/A               |

ROC: receiver operating characteristic, ROC 1: Roc with best sensitivity and specificity, ROC 2: ROC with better accuracy definition, OLR: Ordinary Linear Regression, MIX REG: Linear Mixed Model Regression, LOA: limits of agreement.

*Converted when not available.
the accelerometer activity counts. Thus, most of the included studies derived MET values from a measure of oxygen uptake, which arguably would encompass any possible alteration in energetic cost arising from the disease. However, careful consideration must be given when using METs due to the controversial nature of this method and its failure to represent clinical subgroups (McMurray et al., 2014). Indeed, the standard MET value of 3.5 ml·kg⁻¹·min⁻¹ was developed based on healthy populations and therefore does not reflect pathological, biomechanical, metabolic and respiratory adaptations which are common in many clinical conditions (Byrne et al., 2005).

13. Accelerometer setting and analysis description

Whilst hip was the most popular choice among the included studies (Aadland and Andersen, 2012; Aadland and Steene-Johannessen, 2012; Giffuni et al., 2012; Lopes et al., 2009; Motl et al., 2009; Sandroff et al., 2014a,b; Serra et al., 2017), the best location for monitor placement in clinical populations is unclear. Indeed, comparisons of hip- and wrist-generated thresholds demonstrated great variability which may be explained by biomechanical differences related to dominance (Aadland and Steene-Johannessen, 2012) or functional adaptations due to clinical conditions (Lerner et al., 2014; Ling et al., 2012). For example, in PD, freezing of gait can lead to a rapid trembling in the legs, which would be more efficiently measured by an accelerometer placed on the lower limb (Suzuki et al., 2017). Similarly, other conditions affecting the gait biomechanics might benefit from hip or lower limb placements, as demonstrated under a free-living protocol for chronic stroke and MS patients (Rand et al., 2009; Sparaco et al., 2018).

The choice of accelerometer settings and signal processing should be described in the calibration protocol to allow comparability between studies and generalisability of the developed cut-points (Brond and Arvidsson, 2016). Nonetheless, five of the included studies did not report the sampling frequency and filtering methods used (Agiovlasitis et al., 2012; Giffuni et al., 2012; Lopes et al., 2009; Motl et al., 2009; Serra et al., 2017). The most popular choice of epoch was 60 s (Giffuni et al., 2012; Lopes et al., 2009; Sandroff et al., 2014a,b; Serra et al., 2017; Weikert et al., 2011), with the majority of studies presenting MVPA in counts·min⁻¹. Alternatively, the choice of 1 or 5 s epochs is appropriate to capture short bursts of activities and could be a suitable choice for free-living protocols or for analyses utilising pattern recognition (Gabriel et al., 2010; Staudenmayer et al., 2009a). Whilst counts·min⁻¹ are commonly used, the units are somewhat arbitrary and lack direct practical meaning and transparency due to their proprietary nature (Sievanen and Kujala, 2017). Indeed, the brand-specific units limit inter-study comparisons. In contrast, the use of raw acceleration signals allow more complex analyses and, consequently, higher prediction accuracy (Montoye et al., 2018).

14. Protocol design

The calibration protocols were classified into four categories: laboratory-based protocols that involved walking or running on a treadmill; free-living protocols that assessed participants during their daily routines; daily-life protocols that involved daily-life activities in the laboratory and mixed protocols that utilised more than one of the previously described protocols. A free-living protocol is widely considered the most appropriate for calibration as it determines the relationship between EE and PA in an ecologically valid manner (Freedson et al., 2005; Mackintosh et al., 2012). Despite that, almost all the studies in the neuromusculoskeletal disease group (Agiovlasitis et al., 2012; Motl et al., 2009; Nero et al., 2015; Sandroff et al., 2014a, 2014b; Weikert et al., 2011) utilised over-ground walking protocols. Likewise, almost all of the studies in metabolic disease populations (Aadland and Andersen, 2012; Aadland and Steene-Johannessen, 2012; Giffuni et al., 2012; Lopes et al., 2009) performed treadmill walking protocols, with only one study encompassing jogging. A limitation of such walking protocols is that they are unlikely to provide a fair classification of activities beyond those of locomotion (Crouter et al., 2006). In addition, studies suggest that individuals with chronic stroke and PD are more prone to adopt a different strategy to increase gait speed when walking on the treadmill (Lamontagne et al., 2016; Warlop et al., 2018). During treadmill ambulation, the lack of visual cues and a moving floor results in a cautious gait, with individuals adopting slower speeds and increased stride length compared to over-ground walking (Lamontagne et al., 2016; Warlop et al., 2018). As such, the use of treadmill to calibrate for such populations may result in a misrepresentation of their gait during daily-life and should be considered with caution. Alternatively, a free-living protocol would be the ideal framework to provide a more ecologically valid measure of PA (Welk, 2005) in clinical populations.

15. Statistical approach

The statistical approach adopted to translate activity counts and EE into cut-points could substantially impact the derived thresholds. For example, whilst linear regression has been most widely used (Aadland and Andersen, 2012; Aadland and Steene-Johannessen, 2012; Motl et al., 2009; Sandroff et al., 2014a,b; Serra et al., 2017; Weikert et al., 2011), it assumes that the relationship between activity counts and metabolic data (i.e. VO₂ METs) is linear. To address this issue, recent calibration studies have incorporated more flexible statistical methods, such as ROC analysis, hierarchical models, and machine learning (Bassett et al., 2000; Crouter et al., 2011; Freedson et al., 2005; Montoye et al., 2017). However, in the context of clinical populations, it is pertinent to note that ROC analysis does not allow adjustment for clinical factors and may therefore not be an optimal approach.

Machine learning and pattern recognition have been identified as the optimal methods for classifying PA (Bonomi et al., 2009; Staudenmayer et al., 2015; Staudenmayer et al., 2009b; Welk, 2005). A recent systematic review highlighted the high predictive accuracy of laboratory-calibrated protocols using machine learning models ( Farrahi et al., 2019), with hidden markov models (Pober et al., 2006), decision trees (Mathie et al., 2004) and artificial neural networks (Staudenmayer et al., 2009b) the most common models used to estimate PA from raw accelerometer signals. Indeed, the use of such models improved PA prediction, overcoming the inherent limitations of using static epoch lengths (Montoye et al., 2018). Whilst promising, the use of machine learning to estimate PA from raw accelerations is still in the early phases of development. Specifically, the reproducibility of machine learning approaches in free-living settings requires further investigation ( Kerr et al., 2016). Additionally, machine learning models often require considerably sized training data sets, particularly deep learning, which might be a challenge when using indirect calorimetry ( Mannini and Sabatini, 2010). Future studies calibrating accelerometer for clinical populations should consider using machine learning in order to achieve higher prediction accuracy and promote advancements in the field. However, it is noteworthy that even complex statistical approaches such as pattern recognition would still require an optimised calibration protocol in order to ensure high prediction accuracy. In addition, other statistical approaches should also be considered, such as probability analysis which has been employed to translate activity counts into PA behavioural data in mental illness patients ( Chapman, 2017).

Cross-validation establishes the validity of the developed cut-points and verifies that the thresholds are applicable across any participant of similar age and health status to the sample it was generated from. Whilst it is recommended that a cross-validation should be conducted utilising an independent sample and different activities (Welk, 2005), the use of a leave-one-out-approach can also be considered. For example, Nero et al., 2015 used a leave-one-out approach to cross-calibrate the specific PD cut-points. Additionally, a measure of agreement should be performed in addition to a cross-validation (Lopes et al., 2009), and the cross-validation should be applied after developing the
Thresholds and not as a robustness check prior to the analysis (Aadland and Andersen, 2012). Future studies should continue to cross-validate the disease-specific thresholds to ensure their reliability and validity across different protocols and clinical stages.

16. Outcome: MVPA cut-points

Disease-specific cut-points are essential in understanding and promoting PA in clinical populations. The majority of the MVPA cut-points developed for clinical populations were different to those previously developed for healthy adults (Freedson et al., 1998); disease-specific MVPA cut-points varied greatly, from 612 counts·min\(^{-1}\) to 6460 counts·min\(^{-1}\), even within the same condition. Indeed, Serra et al. (2017) developed Actical MVPA cut-points for stroke patients that were equivalent to LPA cut-points for general population. This large variability can be attributed to the occurrence of gait impairment at advanced stages of the disease, in addition to differences in treatments and medications. However, a control group is warranted in order to ascertain whether any variation in the cut-points is caused by the pathophysiology of the disease or differences in the calibration protocol. Indeed, whilst a control group is highly beneficial in the interpretation of the findings of each study, cut-points previously established for general populations could be used when necessary to investigate whether the use of the disease-specific cut-points enhances the predictive accuracy (Janssen et al., 2015; Trost et al., 2015).

17. Strengths and limitations

It is important to acknowledge that the search protocol was developed with a subject-specific librarian, following a rigorous iterative process. Specifically, initial pilot searches were conducted to assess the feasibility of the initial criteria and search terms. Revisions were subsequently made to the outcomes, risk of bias assessment and final analyses. Moreover, extensive screening was performed by the first author to capture all calibration studies, irrespective of healthy or clinical status, to ensure that no clinical calibration studies were missed. Whilst this review is associated with numerous strengths, there are, nonetheless, limitations. Firstly, only studies generating MVPA cut-points were included; whilst cut-points are still widely used in PA research, major limitations associated with this practice should be acknowledged. The large variability of intensity-related cut-points also occurs among general population (Reilly et al., 2008), causing what Trost (2007) described as the ‘cut-point conundrum’. This discrepancy is multifactorial, arising in part from the lack of standardization of calibration protocols and broad range of statistical approaches applied to reduce accelerometer data to cut-points. It is also important to acknowledge the high risk of bias encountered within the included studies which limits our conclusions. Finally, it is noteworthy that the present recommendations were based on a relatively small range of clinical conditions, further demonstrating the need for more population-specific calibration protocols.

18. Conclusion

This systematic review highlights the large variability in MVPA cut-points developed for clinical populations. Indeed, a lack of standardization in the protocol design, as well as the statistical approach, makes it impossible to compare disease-specific cut-points to those generated for healthy populations. To ensure ecological validity, future calibration protocols should incorporate a large variety of free-living activities, of various intensities, instead of protocols composed predominantly of walking. Moreover, future research should ensure a robust measure of EE is adopted as the criterion measure for accelerometry, as well as a precise estimation of RMR. Studies incorporating a control group and utilising cross-validation of the developed clinical thresholds are warranted. Finally, whilst standardization is necessary, it is highly recommended that future studies consider the pathophysiology of the disease when designing the protocol.

Declarations

Ethics Approval and consent to participate: Not applicable.
Consent for publication: Not applicable.
Availability of data and material: Not applicable.

Authors’ contributions

MSB made substantial contributions to conception, design, systematic search, data analysis and interpretation, and drafted of the manuscript. MM and KM made substantial contributions to conception, design, systematic search, data analysis and interpretation, manuscript writing and critically revised the manuscript for important intellectual content. LL made substantial contribution to the conception of the initial protocol, development of methodology and input revising the manuscript critically. 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.2019.101001.

References

Aadland, E., Andersen, S.A., 2012. Treadmill calibration of the actigraph GT1M in young-to-middle-aged obese-to-severely obese subjects. J. Obes. 2012, 8.
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. Method 12, 172.
Agiovlastis, S., Mol, R.W., Foley, J.T., Fernhall, B., 2012. Prediction of energy expenditure from wrist accelerometer in people with and without down syndrome. Adapt. Phys. Activity Q. 29, 179–190.
Alavand, A.O., Merghani, T.H., Ballal, M.A., 2013. Resting metabolic rate in obese diabetic and obese non-diabetic subjects and its relation to glycaemic control. BMC Res. Notes 6, 382.
Bassett, Jr., D.R., Ainsworth, B.E., Swartz, A.M., Strath, S.J., O’Brien, W.L., King, G.A., 2000. Validity of four motion sensors in measuring moderate intensity physical activity. Med. Sci. Sports Exerc. 32, 5471–5480.
Bassett, D.R., Bowland, A.V., Trost, S.G., 2012. Calibration and validation of wearable monitors. Med. Sci. Sports Exerc. 44, S32–S38.
Bell, S.C., Saunders, M.J., Elborn, J.S., Shah, D.J., 1996. Resting energy expenditure and oxygen cost of breathing in patients with cystic fibrosis. Thorax 51, 126–131.
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.
Broidi, J.C., Arvidsson, D., 2016. Sampling frequency affects the processing of actigraph raw acceleration data to activity counts. J. Appl. Physiol. 1985 (120), 362–369.
Byrne, N.M., Hills, A.P., Hunter, G.R., Weinsier, R.L., Schutz, Y., 2005. Metabolic equivalent: one size does not fit all. J. Appl. Physiol. (Bethesda, Md. 1985) 99, 1112–1119.
Casperson, C.J., Powell, K.E., Christensen, G.M., 1985. Physiological activity, exercise, and physical dimensions and distinctions for health-related research. Public Health Rep. 100, 126–131.
Chapman, J.J., 2017. Quantification of Free-living Activity Patterns Using Accelerometry in Adults with Mental Illness. Nature Publishing Group, Macmillan Building, 4 Crinan St, London N1 9XW, England.
Clarke, C.L., Grighton, J.L., Goodbrand, J.A., McMurdo, M.E.T., Witham, M.D., 2017. Validation of the AX3 triaxial accelerometer in older functionally impaired people. Aging Clin. Exp. Res. 29 (3), 451–457. https://doi.org/10.1007/s40520-016-0604-9.
Colberg, S.R., Sigal, R.J., Fernholt, R., Regensteiner, J.G., Blässer, J.B., Rubin, R.R., Chasan-Taber, L., Albright, A.L., Braun, B., 2010. Exercise and Type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement: part II: exercise and diabetes management. Diabetes Care 147–176.
Crouter, S.E., Hurria, J.R., Bassett Jr., D.R., 2011. Use of a two-regression model for estimating energy expenditure in children. Med. Sci. Sports Exerc. 44, 1177–1185.
Fares, S., Miller, M.D., Masters, S., Crotty, M., 2008. Measuring energy expenditure in children. Eur. J. Appl. Physiol. 98, 601–612.
Crouter, S.E., Horton, M., Bassett Jr., D.R., 2011. Use of a two-regression model for estimating energy expenditure in children. Med. Sci. Sports Exerc. 44, 1177–1185.
Freedson, P.S., Melanson, E., Sirard, J., 1998. Calibration of the Computer Science and Applications, Inc. accelerometer. Med. Sci. Sports Exerc. 30, 7525–5330.
Freedson, P.S., Lyden, K., Koozy-Keadle, S., Staudenmayer, J., 2011. Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample. J. Appl. Physiol. 1985 (111), 1812–1818.
Freedson, P.S., Melanson, E., Sirard, J., 1998. Calibration of the Computer Science and Applications, Inc. accelerometer. Med. Sci. Sports Exerc. 30, 777–781.
Gabriel, K.P., McClain, J.J., Schmid, K.K., Storti, L.K., High, R.R., Underwood, D.A., Kuller, L.H., Krika, A.M., 2010. Issues in accelerometer methodology: the role of epoch length on estimates of physical activity and relationships with health outcomes in overweight, post-menopausal women. Int. J. Behav. Nutr. Phys. Act 7, 53.
Gajewski, M., Rudzikiewicz, P., Malincki, S., 2017. The body as an energetic hybrid? New perspectives for chronic disease treatment? Reumatologia 55, 94–105.
Giffuni, J., McMurry, R.G., Schwartz, T., Berry, D., 2012. Actical Accelerometer cut-points for quantifying levels of exertion: comparing normal and overweight adults. Int. J. Exerc. Sci. 179.
Goldstein, S., Askanazi, J., Weissman, C., Thomashow, B., Kinney, J.M., 1987. Energy expenditure in persons with multiple sclerosis: disability matters, but why? Mult. Scler. Int. 2014, e167.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
Gillen, W., Schmid, E., Mädel, S., Okely, T., 2015. Evaluation of actical equations and thresholds to predict physical activity in adults. J. Sports Sci. 498.
neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. J. Appl. Physiol. 107, 1300–1307.

Suzuki, M., Mitoma, H., Yoneyama, M., 2017. Quantitative analysis of motor status in Parkinson’s disease using wearable devices: from methodological considerations to problems in clinical applications. Parkinsons Dis. 2017, 6139716.

Troiano, R.P., Berrigan, D., Dodd, K.W., Masse, L.C., Tilert, T., McDowell, M., 2008. Physical activity in the United States measured by accelerometer. Med. Sci. Sports Exerc. 40, 181–188.

Trost, S.G., 2007. State of the art reviews: measurement of physical activity in children and adolescents. Am. J. Lifestyle Med. https://doi.org/10.1177/1559827607301686.

Trost, S.G., Fraga-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.

Valenti, G., Camps, S.G., Verhoef, S.P., Bonomi, A.G., Westerterp, K.R., 2014. Validating measures of free-living physical activity in overweight and obese subjects using an accelerometer. Int. J. Obes. (Lond.) 38, 1011–1014.

Warlop, T., Detrembleur, C., Stoquart, G., Lejeune, T., Jeanjean, A., 2018. Gait complexity and regularity are differently modulated by treadmill walking in Parkinson’s disease and healthy population. Front. Physiol. 9.

Weikert, M., Dlugonski, D., Sub, Y., Fernhall, B., Motl, R.W., 2011. The impact of gait disability on the calibration of accelerometer output in adults with multiple sclerosis. Int. J. MS Care 13, 170–176.

Weikert, M., Sub, Y., Lane, A., Sandroff, B., Dlugonski, D., Fernhall, B., Motl, R.W., 2012. Accelerometry is associated with walking mobility, not physical activity, in persons with multiple sclerosis. Med. Eng. Phys. 34, 590–597.

Welk, G.J., 2005. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med. Sci. Sports Exerc. 37, S501–S511.

Wens, I., Daligas, U., Vanderbeke, F., Krekels, M., Grevenveld, L., Eijnde, B.O., 2014. Multiple sclerosis affects skeletal muscle characteristics. PLoS One 9, e108158.