Psychometrics of Wearable Devices Measuring Physical Activity in Ambulant Children With Gait Abnormalities
A Systematic Review and Meta-analysis

Author(s)
van Moorsel, Huib; Engels, Barbara; Buczny, Jacek; Gorter, Jan Willem; Arbour-Nicitopoulos, Kelly; Takken, Tim; Engelbert, Raoul H.H.; Bloemen, Manon A.T.

Document Version
Final published version

Creating Tomorrow

Review Article

Psychometrics of Wearable Devices Measuring Physical Activity in Ambulant Children With Gait Abnormalities: A Systematic Review and Meta-analysis

Check for updates

Huib van Moorsel, MSc [a],*, Barbara Engels, MSc [b,c],*,
Jacek Buczny, PhD [d], Jan Willem Gorter, MD [e,f],
Kelly Arbour-Nicitopoulos, PhD [g], Tim Takken, PhD [h],
Raoul H.H. Engelbert, PhD [i,j],**,
Manon A.T. Bloemen, PhD [b],**

[a] *Institute for Human Movement Studies, University of Applied Sciences, Utrecht, The Netherlands*
[b] *Research Centre Healthy and Sustainable Living, Research group Lifestyle and Health, Utrecht University of Applied Sciences, Utrecht, The Netherlands*
[c] *University Medical Center Utrecht Brain Center and Center of Excellence for Rehabilitation Medicine, Utrecht University, Utrecht, The Netherlands*
[d] *Faculty of Behavioral and Movement Sciences, Department of Experimental and Applied Psychology, Amsterdam University Medical Center, Amsterdam, The Netherlands*
[e] *Department of Rehabilitation, Physical Therapy Science and Sports, University Medical Center Utrecht Brain Center, Utrecht, The Netherlands*
[f] *CanChild, Department of Pediatrics, McMaster University, Hamilton, Ontario, Canada*
[g] *Faculty of Kinesiology and Physical Education, University of Toronto, Toronto, Ontario, Canada*
[h] *Child Development & Exercise Center, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands*
[i] *Centre of Expertise Urban Vitality, Faculty of Health, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands*
[j] *Department of Rehabilitation Medicine, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, The Netherlands*

KEYWORDS
Accelerometry;
Gait abnormality;
Meta-analysis;
Physical activity;
Psychometric
properties;
Rehabilitation;
Systematic review

Abstract   *Objective:* To evaluate psychometrics of wearable devices measuring physical activity (PA) in ambulant children with gait abnormalities due to neuromuscular conditions.
Data Sources: We searched PubMed, Embase, PsycINFO, CINAHL, and SPORTDiscus in March 2023.
Study Selection: We included studies if (1) participants were ambulatory children (2-19y) with gait abnormalities, (2) reliability and validity were analyzed, and (3) peer-reviewed studies in the English language and full-text were available. We excluded studies of children with primarily visual conditions, behavioral diagnoses, or primarily cognitive disability. We performed independent screening and inclusion, data extraction, assessment of the data, and grading of results with 2 researchers.
Data Extraction: Our report follows Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. We assessed methodological quality with Consensus-based Standards for the selection of health measurement instruments. We extracted data on reported reliability, measurement error, and validity. We performed meta-analyses for reliability and validity coefficient values.
Data Synthesis: Of 6911 studies, we included 26 with 1064 participants for meta-analysis. Results showed that wearables measuring PA in children with abnormal gait have high to very high reliability (intraclass correlation coefficient [ICC]$_{+, \text{ test-retest reliability}}$=0.81; 95% confidence interval [CI], 0.74-0.89; I^2=88.57%; ICC$_{+, \text{ interdevice reliability}}$=0.99; 95% CI, 0.98-0.99; I^2=71.01%) and moderate to high validity in a standardized setting ($r_{+, \text{ construct validity}}$=0.63; 95% CI, 0.36-0.89; I^2=99.97%; $r_{+, \text{ criterion validity}}$=0.68; 95% CI, 0.57-0.79; I^2=98.70%; $r_{+, \text{ criterion validity cutoff point based}}$=0.69; 95% CI, 0.58-0.80; I^2=87.02%). The methodological quality of all studies included in the meta-analysis was moderate.
Conclusions: There was high to very high reliability and moderate to high validity for wearables measuring PA in children with abnormal gait, primarily due to neurological conditions. Clinicians should be aware that several moderating factors can influence an assessment.
© 2024 The Authors. Published by Elsevier Inc. on behalf of American Congress of Rehabilitation Medicine. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Promoting physical activity (PA) in children to stimulate beneficial health effects, such as physical and psychological development, well-being, and socialization, is an important international target.[1-4] Promoting PA is advised for typically developing children, as well as for those with chronic diseases and disorders with consequent disability. In 2022, the World Health Organization (WHO) reported that more than 80% of children do not meet the recommended levels of PA for optimum health.[5]

There is a lack of reports on levels of PA among children of all ages living with disability.[5] Children living with disability are less likely to be physically active compared with those without a disability.[5-7] PA guidelines report that children from 5 to 17 years of age with and without disabilities should be active for an average of at least 60 minutes per day in moderate-to-vigorous activity (MVPA), as well as engaged in activities designed to strengthen muscle and bone at least 3 days a week and should limit sedentary time.[5,8] The WHO defines the age range of 2-19 years old as encompassing childhood and adolescence.[9] Accordingly, this study will refer to this population as children.

Health care professionals can use different instruments to assess and inform children and their caretakers about reaching recommended levels of PA. Clinicians can use these instruments and data to propose individual monitoring of PA to enhance or maintain a physically active lifestyle. The availability of instruments with demonstrated psychometric properties for children with and without disabilities is essential. Self-report instruments, such as questionnaires and activity diaries, are cost-effective and feasible for users but subject to social desirability and recall bias and tend to have low reliability and weak to moderate validity.[5,10-12] Wearables such as accelerometers or pedometers seem more reliable and valid for measuring PA in children without disabilities.[13,14] Among children with disabilities, there are several patient groups that may show gait abnormalities due to neurological conditions, including cerebral palsy (CP), spina bifida (SB), or muscular dystrophy.[15-17] Variability in movement, asymmetrical walking pattern, and slow walking pace influence psychometric properties of wearables, resulting in overestimation or underestimation of PA outcomes.[18-21]

"Normal gait" relates to neuromotor development and movement parameters.[22] The International Classification of Diseases and Related Health Problems (version 2019) defines "gait with abnormalities" as ataxic and paralytic gait or difficulty in walking.[23] Individual studies of wearable devices measuring PA in children with gait abnormalities are inconclusive about their psychometric properties and clinical implications because of variations within assessed patient populations, placement of the devices, suitability of existing cutoff points related to PA metrics in specific populations, and different measurement protocols.[10,16,24-26] A systematic review or meta-analysis of the psychometric properties of existing studies reporting on wearable devices measuring PA in children with gait abnormalities is lacking.

Therefore, we aim to critically appraise, compare, summarize, and generalize psychometric properties of studies reporting on wearables assessing PA in ambulant children (2-19y of age) with gait abnormalities associated with neuromuscular conditions. Based on these findings, clinicians can be informed about wearable devices measuring PA using existing wearables as guidance for tailored treatment in clinical practice for children with disabilities and consequent gait abnormalities.

Methods

We registered our review protocol at the International Prospective Register for Systematic Reviews, registration number CRD42022313297. We conducted this review according to the quality assessment of patient-reported outcome measurements, the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN),[27,28] and reported following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses.[29] All appendices and supplemental appendices are available on the Open Science Framework (OSF, https://osf.io/kgse9/).

We performed study screening, data extraction data assessment and grading of the results, with 2 independent researchers and, if indicated, with a third researcher for consensus. We performed the meta-analysis with 2 independent researchers using *R* version 4.3.2 (Peter Dalgaard) (2024).[30]

Search strategy

We systematically conducted a literature search in March 2023 (supplemental appendix S1, available online only at http://www.archives-pmr.org/, OSF) in the following electronic databases: PubMed (MEDLINE), Embase (Elsevier), PsycINFO, CINAHL, and SPORTDiscus (EBSCOhost). We divided key search words into 3 blocks: (1) words related to the population (children with gait abnormalities), (2) measuring methods (wearable devices to measure PA), and (3) measurement properties (reliability, measurement error, validity). We used previously published search filters[31-33] and we consulted a panel of clinicians to check for completeness. Also, we hand-searched the reference lists of the included studies for the inclusion of relevant additional studies. We conducted the search without any restrictions regarding the publication date.

Eligibility criteria

Inclusion criteria were: (1) ambulatory children between 2 and 19 years of age with gait abnormalities; (2) examination of measurement properties: reliability, measurement error, and validity of instruments for wearable devices measuring PA as defined by Mokkink et al[28]; and (3) peer-reviewed studies published in English and with full-text available. We included studies with mixed populations (eg, children with normal and abnormal gait) if we were able to extract data on children with abnormal gait patterns. We excluded studies of children with primarily visual conditions, behavioral diagnoses, or primarily cognitive disability.

Screening process and study selection

We imported all identified studies into Covidence systematic review software[a] and removed duplicates. We screened all titles/abstracts and full texts for eligibility using the systematic review software Rayyan.[b]

Data extraction

We used a data extraction form based on COSMIN.[27,28] The extracted information included details regarding the PA setting (whether laboratory, free-living), frequency, intensity, and duration (PA dimension), as well as the classification of physical behavior (sedentary, low PA, MVPA). Moreover, we documented the types of physical behavior, including individual categories (such as lying, sitting, standing, walking, running), and combined categories observed during activities such as free play or structured protocols. We collected information on the name of the author(s), year of publication, study characteristics, population, type of measurement instrument, comparator instrument, measurement protocol and properties, details of units of measure, statistical and psychometrical information, and results.

Evaluation of the measurement properties

Methodological quality
We assessed the methodological quality of included studies with the COSMIN Risk of Bias tool for outcome measurement instruments, following the standards for reliability, measurement error, and validity studies.[34,35] We scored each item following the COSMIN with a 4-point scale: very good (V), adequate (A), doubtful (D), and inadequate (I). To determine the overall rating of the methodological quality of each single study, we used the worst-score-counts principle.[36]

Evaluation of results and grading of the quality of evidence
We categorized the results of each measurement property: reliability (test-retest reliability, inter-device reliability), measurement error, and validity (construct, criterion, criterion cutoff point based). As our analytical approach is of a quantitative nature, we only rated the pooled results of the meta-analysis.[34,37]

Meta-analysis

We employed a multilevel random-effects meta-analysis using *metafor*, version 4.4-0 (Wolfgang Viechtbauer).[38,39] This analysis considered sampling variance (level 1), within-study variance (level 2), and between-study variance (level 3), and reported concordance between standard reference and wearables measuring PA with pooled effect sizes.[40,41] For reliability studies, we considered the intraclass correlation coefficient (ICC) as the effect size with generalizability coefficients and Pearson *r* coefficient as equivalents.[42] For validity studies, Pearson *r* served as the effect size, with the concordance correlation coefficient, ICC and Spearman ρ considered equivalents.[43] The materials of the meta-analysis (appendices 1 and 2) and calculation of effect sizes

(supplemental appendix S2, available online only at http://www.archives-pmr.org/) are available at OSF.

We assessed heterogeneity levels for each level to explain the within and between-study variance with R packages clubSandwich,[44] dmetar,[45] meta,[40] orchaRd,[46] and, for cross-validating, we used robumeta.[47] We evaluated the level of heterogeneity with Q, τ^2, and I^2 statistics. I^2 indicates how much variance of true effects is reflected in the observed variance of the effect sizes,[48] for simplification we assumed that a value of I^2 greater than 75% suggests high heterogeneity in the effect sizes.[49] The scripts of the meta-analyses are available (supplemental appendices S3 and S4, available online only at http://www.archives-pmr.org/, OSF).

Moderator coding and analytical strategy

Through the multilevel meta-analysis, we could show relationships between moderators and effect sizes. Moderator analysis was used to investigate variation in the evidence and heterogeneity. It could further lead to the understanding of findings and help produce practical insights that guide further decision-making.[48] We performed moderator analyses to assess both reliability and validity. Table 1 presents the moderator coding for each psychometric property. Specific moderators were coded for test-retest reliability, inter-device reliability, construct validity, criterion validity, and cutoff point criterion validity, as per our data extraction (supplemental appendices S5 and S6, available online only at http://www.archives-pmr.org/, OSF).

Sensitivity analysis and publication bias were evaluated using a funnel plot, a p-curve analysis,[50] and the R package PublicationBias.[51] Furthermore, we conducted moderator analysis by examining the relationship between year of publication and effect sizes.[52] By means of R package metaplus,[41] we ran analysis to identify possible outliers. Lastly, as the number of effect sizes for all the meta-analyses could be considered low, a post hoc power analysis was conducted using POMADE.[53]

For interpretation of pooled correlation coefficients with 95% confidence interval (CI), we interpreted 0-0.3 as negligible correlation; 0.3-0.5 as low correlation; 0.5-0.7 as moderate correlation; 0.7-0.9 as high correlation; 0.9-1.0 as very high correlation.[54] We chose different interpretations than COSMIN based "sufficient," "indeterminate," and "insufficient" because we wanted to be able to differentiate, if necessary, borderline values with more detail, as we included negative as well as positive results in our meta-analysis.

Grading the pooled evidence

Finally, we evaluated the quality of evidence for the pooled results of all studies using the modified Grading of Recommendations Assessment, Development, and Evaluation, as proposed in the COSMIN guideline and rated the quality of evidence as "high," "moderate," "low," or "very low."[28]

Results

Description of studies

We identified 8777 studies resulting in 6911 unique studies after duplicate removal, of which we included 47 for full text screening. Through reference search we included 5 additional ones.[24,55-58] Eventually, for data extraction we included 30 studies with 1145 children with gait abnormalities between the lowest mean age of 2.3 years[56] and the highest mean age of 18.7 years.[59] Figure 1 presents a flow diagram of the full selection process.

Eight studies focused on test-retest reliability,[59-66] and 3 examined inter-device reliability.[67-69] Two examined test-retest reliability as well as inter-device reliability.[20,57] Seven studies reported measurement error,[57,59,62-66] whereas 8 assessed reliability and validity.[20,57,59,60,66-69]

For reliability, a total of 670 children were included (with the lowest mean age of 6.3y[60] and the highest mean age of 18.7y[59]). Eleven studies described wearables measuring PA in children with CP,[20,57,59,60,63-69] whereas the others investigated children with amyoplasia, distal arthrogryposis,[62] or acquired brain injury (ABI).[61] Six studies were conducted in a laboratory setting,[20,60,66-69] 5 in free-living settings,[57,59,62-64] and 2 in both settings.[61,65] One study assessed test-retest reliability with 1 device on the thigh and inter-device reliability with 2 devices in different locations.[57] Twice, the device was placed on the ankle,[62,64] and once on the thigh.[59] In 4 studies, the device was placed on the trunk,[20,61,65,67] or multiple devices were placed in different locations.[60,63,66,69]

Measurement error was assessed in 440 children with the lowest mean age of 8.3 years[64] and the highest mean age of 18.7 years.[59] A detailed description of reliability and measurement error studies is presented in table 2.

Validity was assessed in 25 studies.[16,18,20,21,24,26,55-60,66-78] Three evaluated construct validity,[21,66,74] 13 evaluated criterion validity,[16,20,57-60,67-69,72,75,76,78] 4 focused on criterion validity based on cutoff points,[24,56,70,71] and 4 described criterion validity and criterion validity based on cutoff points.[18,26,73,77] One study described criterion validity and construct validity.[55]

Validity was assessed in 661 children with the lowest mean age of 2.3 years[56] and the highest mean age of 18.7 years.[59] Eighteen studies described wearables measuring PA in children with CP,[18,20,24,55-57,59,60,67-75,78] and 2 in children with SB.[16,76] Three studies included children with ABI,[77] juvenile arthritis and inherited muscle disease,[26] or congenital muscular dystrophy, respectively.[21] In 24 out of 30 studies, a laboratory setting was used.[16,18,20,21,24,26,55,56,58-60,66-78] In 10 studies the wearable device was placed on the trunk,[20,21,24,26,56,70,71,73,75,77] in 7 on the thigh,[16,57-59,67,72,78] and in 1 on the arm.[68] Seven studies placed multiple devices in different locations.[18,55,60,66,69,74,76] Tables 3A and B summarize a detailed description of the studies' validity.

Quality of studies

Reliability and measurement error

We rated 2 test-retest reliability studies[64,65] as adequate, 6 as doubtful,[59-63,66] and 1 was inadequate.[57] For inter-device reliability studies, 1 achieved adequate methodological quality,[67] whereas 3 were rated as doubtful,[20,68,69] and 1 was rated as inadequate (table 2).[57] Insufficient description of population characteristics, missing data, or test-retest

Table 1 Moderator coding for reliability and validity.

Moderators	Aggregation	Reliability		Validity		
		Test-Retest	Interdevice	Construct Validity	Criterion Validity	Criterion Validity Cutoff Point Based
PA setting	Laboratory	•	•			
	Free-living	•	•			
PA dimension	Frequency	•	•		•	•
	Intensity	•	•		•	•
	Duration	•	•		•	•
Physical behavior type	Single category	•		•	•	
	Combined categories	•		•	•	
Physical behavior type specified	Walking	•	•		•	
	Lying and sitting	•	•		•	
	Standing	•			•	
	Running		•		•	
Physical behavior class	SED					•
	LPA					•
	MVPA					•
%GMFCS level I		•	•	•	•	•
Placement	Leg	•	•	•	•	
	Trunk	•	•		•	
	Arm	•	•	•	•	
	Multi	•			•	
Age	>6 y					•
	<6 y					•
	>13 y	•	•		•	•
	<13 y	•	•		•	•
Interval test-retest	>2 wk	•				
	<2 wk	•				
Placement body side	Same		•			
	Opposite		•			
Cutoff points	Population specific					•
	General					•

NOTE. A dot (•) corresponds to the examined moderators; GMFCS was used in studies with CP and ABI; when GMFCS levels were not specified, the average percentage GMFCS level across all studies was used.
Abbreviations: LPA, low physical activity; SED, sedentary.

duration intervals longer than 2 weeks influenced methodological quality.

Two measurement error studies demonstrated adequate methodological quality,[64,65] whereas 4 were rated doubtful,[59,62,63,69] and 1 showed inadequate methodological quality (table 2).[57]

Validity

We rated 5 criterion validity studies as very good,[20,67,69,75,76] 6 as doubtful,[55,57,59,60,68,78] and 3 as inadequate.[16,58,72] For criterion validity based on cutoff points, we rated 5 studies as very good,[18,24,26,70,77] and 3 as doubtful.[71,73,79] All studies examining construct validity achieved doubtful methodological quality.[21,55,66,74] Not reporting the length of the epoch, missing a description of the population, or inadequate statistical calculations resulted in lower levels of methodological quality (tables 4A and B).

Meta-analysis

From 30 studies, we excluded 4 that did not present absolute values, therefore we were not able to convert results into ICC or Pearson r.[16,58,66,75] We conducted separate meta-analyses for 26 studies, deriving from test-retest reliability, inter-device reliability, criterion validity, criterion validity of cutoff point based methods and construct validity.[18,20,21,24,26,55-57,59-65,67-74,76-78] Two studies presented partial data: one described newly developed cutoff points but no comparator cutoff points,[26] whereas another described newly developed cutoff points but no absolute values regarding the validity of those.[73] We used reported partial data for analyses.

We excluded all the studies analyzing measurement error because of the impossibility to convert to a standard effect size and because of heterogeneity in outcome measures evaluating measurement error, for example, levels of agreement,[57,59] coefficient of variation,[62] standard error of the mean with or without the smallest detectable change,[63,65] and within-subject SD.[66]

For meta-analysis, we included a total of 1064 children with the lowest mean age of 2.3 years[56] and the highest mean age of 18.7 years.[59] In the studies, children were diagnosed with CP (n=22)[18,20,24,55-57,59,60,63-65,67-75,78]; ABI (n=2)[61,77]; amyoplasia and distal arthrogryposis (n=1)[62];
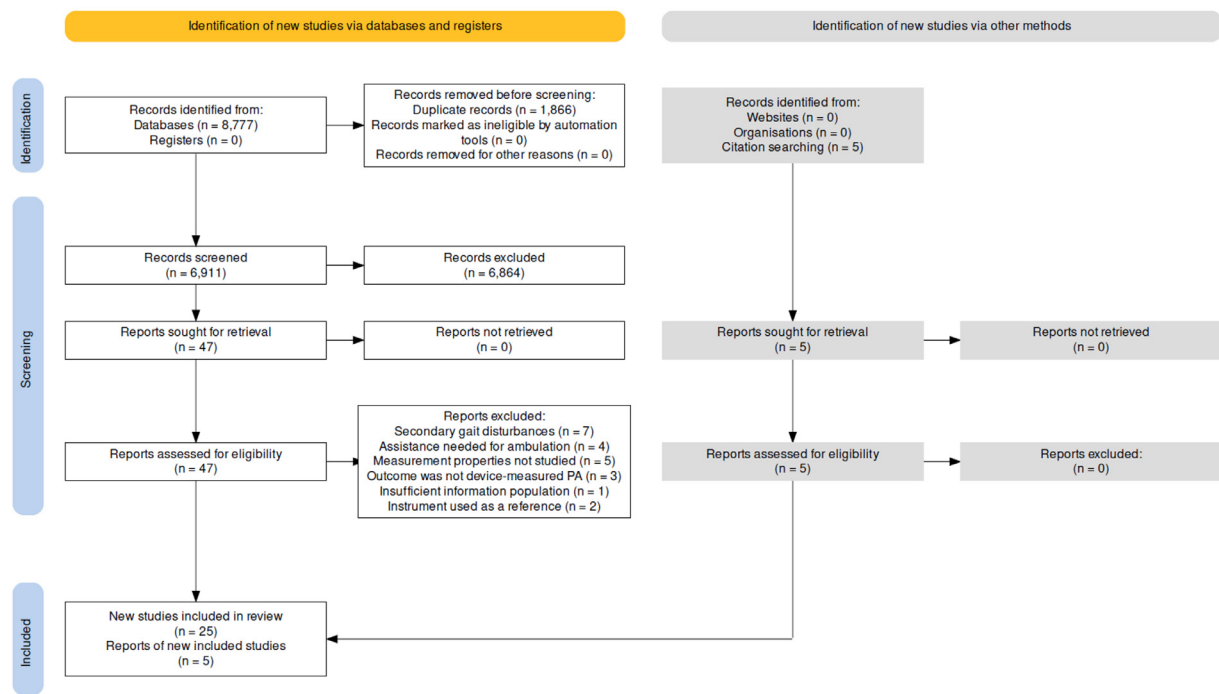
Fig 1 Preferred Reporting Items for Systematic Reviews and Meta-analyses flow diagram of study inclusion. For the quantitative analysis, we included 26 studies.

juvenile arthritis and inherited muscle disease (n=1)[26]; SB (n=1)[76]; and congenital muscular atrophy (n=1).[21]

Meta-analysis: Test-retest reliability
The overall effect size of test-retest reliability was high with $ICC_+=0.81$ (95% CI, 0.74-0.89; I^2=88.57%; see appendix 1; fig 2).

Meta-analysis: Interdevice reliability
The overall effect size of inter-device reliability in wearables measuring PA was very high with an $ICC_+=0.99$ (95% CI, 0.98-1.00; I^2=71.01%; see appendix 1; fig 3).

Meta-analysis: Construct validity
The overall effect size of construct validity to measure PA was moderate with $r_+=0.63$ (95% CI, 0.36-0.89; I^2=99.97%; see appendix 1; fig 4).

Meta-analysis: Criterion validity
The overall effect size of criterion validity was high with $r_+=0.70$ (95% CI, 0.59-0.82; I^2=98.70%; see appendix 1; fig 5).

Meta-analysis: Criterion validity of cutoff point based methods
The overall effect size of criterion validity of cutoff point based methods was moderate with $r_+=0.69$ (95% CI, 0.58-0.80; I^2=87.02%; see appendix 1; fig 6).

Meta-analysis: Moderator analyses
For test-retest reliability, PA setting was a statistically significant moderator with a very high correlation for laboratory settings with $ICC_+=0.91$ (95% CI, 0.81-1.00) and a high correlation for free-living settings with $ICC_+=0.77$ (95% CI, 0.69-0.84). Frequency (steps, PA counts) showed a very high correlation with $ICC_+=1$ (95% CI, 0.77-1.00), and intensity (energy expenditure rate) and duration (seconds, minutes of PA) showed a high correlation with $ICC_+=0.88$ (95% CI, 0.78-0.97) and $ICC_+=0.71$ (95% CI, 0.60-0.81), respectively. Specified physical behavior type showed a high correlation for walking with $ICC_+=0.87$ (95% CI, 0.76-0.98), lying and sitting with $ICC_+=0.73$ (95% CI, 0.56-0.89), and standing with $ICC_+=0.70$ (95% CI, 0.53-0.87). Percentage Gross Motor Function Classification System (GMFCS) level, placement of the device, difference in age, physical behavior type, and duration interval between test and retest were statistically nonsignificant moderators (see appendix 2).

For inter-device reliability, PA dimensions showed a very high correlation for intensity (energy expenditure rate; $ICC_+=1$; 95% CI, 1.00-1.00), frequency (steps, PA counts; $ICC_+=0.98$; 95% CI, 0.98-0.99), and duration (seconds, minutes of PA; $ICC_+=0.99$; 95% CI, 0.98-0.99). Regarding age, a very high correlation was found for children younger than 13 years ($ICC_+=0.99$; 95% CI, 0.98-0.99) and children older than 13 years ($ICC_+=1$; 95% CI, 0.99-1.00). PA setting, specified and unspecified physical behavior type, placement of the device and placement on the same or opposite body side were nonsignificant moderators (see appendix 2).

We did not find any statistically significant moderators (percentage GMFCS level I, physical behavior type, placement of the device) related to construct validity (see appendix 2).

For criterion validity, moderator analysis showed that only the specified physical behavior type was a significant moderator with a high correlation for walking and running with $r_+=0.73$ (95% CI, 0.56-0.90) and $r_+=0.76$ (95% CI, 0.56-0.95), respectively. Lying and sitting and standing showed a moderate correlation with $r_+=0.46$ (95% CI, 0.23-0.70) and $r_+=0.64$ (95% CI, 0.43-0.86). PA dimensions, percentage

Table 2 Results reliability and measurement error of device-measured physical activity.

Study Author(s) YoP	Devices		Population			Device-Measured PA	Reliability	Measurement Error	Quality	
	Name Device Epoch Length Placement Body	PB Domain Activity Duration	Condition GMFCS Level (n)	Sample Age Gender	Group Specs				Risk of Bias	GRADE Reliability
Test-retest reliability										⊕⊕⊕
Aviram et al[60] 2011	IDEEA ND Chest, thigh (2), foot (2)	Laboratory 2 × 4 min	Cerebral palsy I=4 II=5 III=3	n=12 Age=6.3±1.3 y 6 girls		EE rate (Kcal/min) slow walking	r=0.998, P≤.001		D	
Bania[59] 2014	activPAL ND Thigh	Free living 7 × 24 h	Cerebral palsy II=15 III=9	n=21 Age=18.7±2.9 y 13 girls	n=21 Meas>2 d n=17 Meas>2 d	Duration (h/d) standing Duration (h/d) lying+sitting Steps (ND) walking Duration (h/d) standing Duration (h/d) lying+sitting Steps (ND) walking	ICC (2,1)=0.60 ICC (2,1)=0.66 ICC (2,1)=0.87 ICC (2,1)=0.56 ICC (2,1)=0.63 ICC (2,1)=0.89	MD=−0.31 h/d, 95% LoA (−2.6 to 1.98) MD=0.37 h/d, 95% LoA (−2.30 to 3.04) MD=−411 steps/d, 95% LoA (−3125 to 2303) MD=−0.40 h/d, 95% LoA (−2.9 to 2.1) MD=0.48 h/d, 95% LoA (−2.4 to 3.3) MD=−516 steps/d, 95% LoA (−3191 to 2159)	D	
Baque et al[61] 2016	ActiGraph GT3X+ 5 s Waist (LNI)	Laboratory 4 × 5 min	ABI I=17 II=15	n=32 Age=12.1±2.3 y ND		AC−VM (n/min) slow walking AC−VM (n/min) moderate walking AC−VM (n/min) fast walking AC−VM (n/min) stepping	ICC (2,1)=0.90; 95% CI, 0.79-0.95 ICC (2,1)=0.83; 95% CI, 0.66-0.92 ICC (2,1)=0.91; 95% CI, 0.82-0.96 ICC (2,1)=0.89; 95% CI, 0.77-0.95		D	
Baque et al[61] 2016	ActiGraph GT3X+ 60 s Waist (LNI)	Free living 4 × >7 h (2 wk+2 we)	Acquired brain injury I=26 II=25	n=51 Age=12.1±2.4 y ND	Total (n=51) Meas 2-3-4 d Children (<13 y; n=33) Meas 2-3-4 d Adoles (≥13 y; n=18) Meas 2-3-4 d	Duration (h) unstructured activities (MVI)* Duration (h) unstructured activities (MVI)* Duration (h) unstructured activities (MVI)*	ICC=0.72; 95% CI, 0.52-0.84 ICC=0.78; 95% CI, 0.66-0.87 ICC=0.79; 95% CI, 0.67-0.87 ICC=0.57; 95% CI, 0.13-0.79 ICC=0.73; 95% CI, 0.52-0.86 ICC=0.70; 95% CI, 0.48-0.84 ICC=0.81; 95% CI, 0.49-0.93 ICC=0.76; 95% CI, 0.48-0.90 ICC=0.79; 95% CI, 0.56-0.91		D	
Braun et al[62] 2016	StepWatch ND Ankle (l)	Free living 7 d × walking h	Amyoplasia, distal arthrogryposis ND	14 Age=10.9±3.8 y 4 girls	n=13	Steps (n) unstructured activities	G coeff=0.67; 95% CI, 0.56-0.78	CV=0.36 steps	D	
Gerber et al[63] 2021	Physilog4 ND Trunk, thigh (2), shank (2)	Free living 2 d × 8 h (walking h)	Cerebral palsy I=6 II=3 III=4	n=15 Age=13.7±3.4 y 9 girls	n=13 wk-wk d n=12 wk-we d	Duration (% total duration) lying +sitting Duration (% total duration) standing Duration (% total duration) walking Duration (% total duration) lying +sitting Duration (% total duration) standing Duration (% total duration) walking	ICC (3,1)=0.47; 95% CI, 0.00-0.80 ICC (3,1)=0.18; 95% CI, 0.00-0.64 ICC (3,1)=0.76; 95% CI, 0.39-0.92 ICC (3,1)=0.30; 95% CI, 0.00-0.73 ICC (3,1)=0.31; 95% CI, 0.00-0.72 ICC (3,1)=0.00; 95% CI, 0.00-0.42	SEM/SDC=7.88/21.8% total duration SEM/SDC=8.22/21.8% total duration SEM/SDC=1.89/5.24% total duration SEM=10.73% total duration SEM=8.06% total duration SEM=4.12% total duration	D	
Gerber et al[63] 2021			Cerebral palsy I=3 II=2 III=5	n=10 Age=13.1±3.7 y 6 girls	Same wk d, 2-4 wk apart	Duration (% total duration) lying +sitting Duration (% total duration) standing Duration (% total duration) walking	ICC (3,1)=0.90; 95% CI, 0.64-0.97 ICC (3,1)=0.90; 95% CI, 0.65-0.97 ICC (3,1)=0.91; 95% CI, 0.67-0.98	SEM/SDC=4.77/13.22% total duration SEM/SDC=3.64/10.10% total duration SEM/SDC=1.39/3.86% total duration	D	
Ishikawa et al[64] 2013	StepWatch ND Ankle (l)	Free living 7 d × walking h	Cerebral palsy I=75 II=78 II=48	n=209 Age=8.3±3.3 y 91 girls	Total (n=201) 2-5 y (n=56) 6-14 y (n=145)	Steps (n) unstructured activities	G coeff=0.77; 95% CI, 0.74-0.81 G coeff=0.77; 95% CI, 0.71-0.83 G coeff=0.78; 95% CI, 0.74-0.82	CV=28% (GMFCS I) CV=40% (GMFCS II) CV=70% (GMFCS III) CV=38% (GMFCS I) CV=40% (GMFCS II) CV=63% (GMFCS III)	A	

Table 2 (*Continued*)

Study Author(s) YoP	Name Device Epoch Length Placement Body	PB Domain Activity Duration	Condition GMFCS Level (n)	Sample Age Gender	Group Specs	Device-Measured PA	Reliability	Measurement Error	Risk of Bias	GRADE Reliability
Mackey et al[66] 2009	IDEEA ND Chest, thigh (2), foot (2),	Laboratory 5 × 30 s	Cerebral palsy I+II=16 III=9	n=25 Age 14.1; range=8-12 17 girls		Duration (min) lying Duration (min) standing Duration (min) sitting Duration (min) walking (overground and stairs)		sw=15% sensitivity; 2% specificity sw=21% sensitivity; 6% specificity sw=21% sensitivity; 13% specificity sw=6% sensitivity; 2% specificity	D	
Mitchell et al[65] 2014	ActiGraph GT3X+ 5 s Waist (LNI)	Laboratory 4 × 5 min	Cerebral palsy I=16 II=14	n=30 Age=11.9±2.6 y 14 girls		AC-VA (n/min) slow walking AC-VA (n/min) moderate walking (MW) AC-VA (n/min) fast walking (FW) AC-VA (n/min) stepping (STEP) AC-VA (n/min) MW+FW+STEP (MVi)	ICC=0.80; 95% CI, 0.76-0.83 ICC=0.80; 95% CI, 0.77-0.83 ICC=0.70; 95% CI, 0.66-0.74 ICC=0.66; 95% CI, 0.60-0.72 ICC=0.80; 95% CI, 0.77-0.83	SEM/SDC=333.6/926 counts/min SEM/SDC=438/1214 counts/min SEM/SDC=702/1945 counts/min SEM/SDC=771.6/2139 counts/min SEM/SDC=510/1412 counts/min	A	
Mitchell et al[65] 2014	ActiGraph GT3X+ 5 s Waist (LNI)	Free living 4 d × >8 h	Cerebral palsy I=44 II=58	n=102 Age=11.3±2.3 y 50 girls	Total (n=81) Meas 2-3-4 d Children (<13 y; n=58) Meas 2-3-4 d Adoles (≥13 y; n=23) Meas 2-3-4 d	Duration (h) unstructured activities (MVI)[i] Duration (h) unstructured activities (MVI)[i] Duration (h) unstructured activities (MVI)[i]	ICC=0.63; 95% CI, 0.44-0.76 ICC=0.57; 95% CI, 0.28-0.74 ICC=0.72; 95% CI, 0.48-0.87 ICC=0.69; 95% CI, 0.56-0.79 ICC=0.63; 95% CI, 0.43-0.77 ICC=0.74; 95% CI, 0.44-0.88 ICC=0.73; 95% CI, 0.61-0.82 ICC=0.73; 95% CI, 0.58-0.83 ICC=0.78; 95% CI, 0.58-0.88		D	
O'Neil et al[20] 2014	ActiGraph GT3X 1 s Waist (l+r)	Laboratory 3 × 6 min	Cerebral palsy I=4 II=1 III=3	n=8 Age=11.9±3.2 y 2 girls	Meas left BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS Meas right BS	AC (n) slow walking AC (n) moderate walking AC (n) fast walking Steps (n) slow walking Steps (n) moderate walking Steps (n) fast walking	ICC=0.75; 95% CI, 0.02-0.95 ICC=0.86; 95% CI, 0.48-0.97 ICC=0.88; 95% CI, 0.52-0.98 ICC=0.95; 95% CI, 0.59-0.99 ICC=0.89; 95% CI, 0.47-0.98 ICC=0.95; 95% CI, 0.81-0.99 ICC=0.88; 95% CI, 0.49-0.98 ICC=0.96; 95% CI, 0.82-0.99 ICC=0.88; 95% CI, 0.47-0.98 ICC=0.98; 95% CI, 0.92-0.99 ICC=0.88; 95% CI, 0.48-0.98 ICC=0.97; 95% CI, 0.88-0.99		D	
Pirpiris and Graham[57] 2004	PAL 1 5 s Thigh (l)	Free living 6 × 24 h	Cerebral palsy ND	n=300 Age=11±3.9 y 151 girls	1 × 1 wk-wk d; (n=50) 1 × 1 wk-we d; (n=50) 2 × 2 wk-wk d; (n=50) 2 × 2 wk-we d; (n=50)	Duration (s) upright during ADL	ICC=0.83; 95% CI, 0.76-0.89 ICC=0.80; 95% CI, 0.74-0.87 ICC=0.87; 95% CI, 0.80-0.96 ICC=0.83; 95% CI, 0.78-0.93	MD=−133.13 s, LoA (-973.1 to 969.0) MD=70.6013 s, LoA (-682.5 to 668.5) MD=65.2313 s, LoA (-778.3 to 753.0)	I	

Inter-device reliability   ⊕⊕⊕

Study Author(s) YoP	Name Device Epoch Length Placement Body	PB Domain Activity Duration	Condition GMFCS Level (n)	Sample Age Gender	Group Specs	Device-Measured PA	Reliability	Measurement Error	Risk of Bias	GRADE Reliability
Koehler et al[68] 2015	SenseWear ND Upper portion of the arm	Laboratory 5 × 5 min	Cerebral palsy II=10	n=10 13.4±1.6 y 3 girls	n=10 n=9 n=10 n=6	EE sitting EE slow walking EE moderate walking EE fast walking		MD=0.0 Kcal/min, SD=0.2 MD=−0.1 Kcal/min, SD=0.6 MD=−0.2 Kcal/min, SD=0.5 MD=0.0 Kcal/min, SD=1.1		
Maher et al[67] 2013	NL−1000 ND Waist (l+r)	Laboratory 3 × 6 min	Cerebral palsy I=8 II=9	n=17 Age=12.3±3.2 y 9 girls	Dominant BS Nondominant BS	AC (n) walking AC (n) running AC (n) walking AC (n) running	ICC=0.88; 95% CI, 0.71-0.96 ICC=0.98; 95% CI, 0.94-0.99 ICC=0.94; 95% CI, 0.84-0.98 ICC=0.99; 95% CI, 0.96-1.00		A	
O'Neil et al[20] 2014	ActiGraph GT3X 1 s Waist (l+r)	Laboratory 3 × 6 min	Cerebral palsy I=4 II=1 III=3	n=8 Age=11.9±3.2 y 2 girls	Left vs right	AC (n) slow walking AC (n) moderate walking AC (n) fast walking Steps (n) slow walking Steps (n) moderate walking Steps (n) fast walking	ICC=0.98; 95% CI, 0.92-0.99 ICC=0.99; 95% CI, 0.98-1.00 ICC=0.98; 95% CI, 0.87-0.99 ICC=0.99; 95% CI, 0.98-1.00 ICC=0.96; 95% CI, 0.81-0.99 ICC=0.99; 95% CI, 0.99-1.00		D	

(*continued*)

Table 2 (*Continued*)

Study Author(s) YoP	Devices		Condition GMFCS Level (n)	Population		Device-Measured PA	Reliability	Measurement Error	Quality	
	Name Device Epoch Length Placement Body	PB Domain Activity Duration		Sample Age Gender	Group Specs				Risk of Bias	GRADE Reliability
O'Neil et al[69] 2016	StepWatch 3 s Ankle (l+r) SenseWear 60 s Upper portion of the arm (l+r) ActiGraph GT3X 1 s Waist (l+r)	Laboratory 6 × 5 min+3 × 6 min	Cerebral palsy I=28 II=16 III=13	n=57 Age=12.5±3.3 y 29 girls	StepWatch SenseWear ActiGraph GT3X	Steps (n/min) during activity protocol AC_VA (n/min) structured activities (II) AC_VM (n/min) structured activities (II)	ICC=0.977; 95% CI, 0.969-0.982 ICC=0.940; 95% CI, 0.929-0.950 ICC=0.986; 95% CI, 0.983-0.989 ICC=0.985; 95% CI, 0.982-0.987 ICC=0.981; 95% CI, 0.978-0.984		D	
Pirpiris and Graham[57] 2004	PAL 1 5 s Trunk, thigh (l+r)	Free living 6 × 24 h	Cerebral palsy GMFCS ND	n=300 Age=11±3.9 y 151 girls	Same limb (n=20) Opposite limb, n=20 Limb-trunk (n=20)	Duration (s) upright unstructured Activities	ICC=0.99; 95% CI, 0.89-0.99 ICC=0.99; 95% CI, 0.88-0.99 ICC=0.52; 95% CI, 0.45-0.67	MD=1 s, LoA (-120 to 131) MD=2 s, LoA (-118 to 125) MD=−27196 steps/d, LoA (-32,500 to -2189)	I	

Abbreviations: "?", indeterminate; "-", insufficient; "+", sufficient; "⊕⊕⊕", moderate; A, adequate; AC, activity counts; ADL, activities of daily living; adoles, adolescent; BS, body side; CV, coefficient of variation; coeff, coefficient; D, doubtful; EE, energy expenditure; G, generalizability; GRADE, Grading of Recommendations Assessment, Development and Evaluation; I, inadequate; IDEEA, Intelligent Device for Energy Expenditure and Activity; II, increased intensity; l, left; LNI, least neurologic impaired; LoA, levels of agreement; MD, mean difference; ME, measurement error; meas, measured; min, minute(s); MVI, moderate-to-vigorous intensity; n, number; ND, not described; PAL 1, Positional Activity Logger version 1; PB, physical behavior; R, reliability; r, Pearson correlation coefficient; r, right; SB, Spearman-Brown; SDC, smallest detectable change; stat, statistics; sw, within-subject SD; VA, vertical axis; VM, vector magnitude; we, weekend; YoP, year of publication;
* Baque (2015)[87] cutoff points were used.
† Evenson (2008)[88] cutoff points were used.

Table 3A Results validity of device-measured physical activity.

Study	Device			Population	
Author(s) YoP	Name Device Epoch Length Placement Body	Comparison Measure Specs Outcome	PB Domain Activity Duration	Condition GMFCS level (n)	Sample Age Gender
Criterion validity					
Aviram et al[60] 2011	IDEEA ND Chest, thigh (2), foot (2)	PIC Cosmed K4b2	Laboratory 1 × ±7 min+3 × 4 min	Cerebral palsy I=8 II=6 III=7	n=21 6.4±1.9 y ND
Bania[59] 2014	activPAL ND Thigh	Video DO ND	Laboratory 2 × 3 min+6 min	Cerebral palsy II=5 III=5	n=10 18.6±2.7 y 4 girls
Baque et al[77] 2017	ActiGraph GTX3+ 15 s Waist (LNI)	PIC Cortex Metamax	Laboratory 1 × 5 min+3 × 6 min+1 × 3 min	Acquired brain injury I=16 II=11	n=27 13.6±2.4 y 12 girls
Capio et al[75] 2010	MTI (ActiGraph) 15 s Waist (r)	SOFIT ND	Laboratory 6 × 2 min (structured) 10 min (free play)	Cerebral palsy I=14 II=9 III=8	n=31 9.7±2.5 y 17 girls
de Groot et al[76] 2013	Actical 60 s Waist (l) Actiheart 15 s Trunk (front)	PIC Cortex Metamax	Laboratory 6 min	Spina bifida Hoffer normal=7 Hoffer community=19	n=39 10.6±2.8 y 11 girls
Koehler et al[68] 2015	SenseWear ND Upper portion of the arm (l+r)	PIC ZAN 600	Laboratory 5 × 5 min	Cerebral palsy II=10	n=10 13.4±1.6 y 3 girls
Kuo et al[55] 2009	AMP 331 ND Ankle (r) DynaPort Minimod ND Trunk (back)	Video DO ND	Laboratory ND	Cerebral palsy I=5 II=12 III=3	n=20 10.5±3.0 y 7 girls
Lankhorst et al[16] 2019	Activ8 5 s Thigh (DBS)	Video DO ND	Free living 45 min (1:30 min)	Spina bifida Hoffer community=2 Hoffer household=2	n=10 12.9±2.1 y 1 girl
Maher et al[67] 2013	NL-1000 pedometer ND Thigh (nDBS, DBS)	Video DO FlipVideo UltraHD Camcorder	Laboratory 2 × 3 min	Cerebral palsy I=8 II=9	n=17 12.3±3.2 y 9 girls
McAloon et al[72] 2014	activPAL ND Thigh (NI)	Video DO ND	Laboratory ND	Cerebral palsy I=4 II=5 III=1	n=10 4-18 y ND
O'Donoghue and Kennedy[78] 2014	activPAL ND Thigh (NI, LNI)	Video DO 2D Sony mini digital Video camera	Laboratory ND	Cerebral palsy I=17	n=17 9.4±3.9 y 8 girls
O'Neil et al[20] 2014	ActiGraph GT3X 1 s Waist (r)	PIC Cosmed K4b2	Laboratory 3 × 6 min	Cerebral palsy I=4 II=1 III=3	n=8 11.9±3.2 y 2 girls
O'Neil et al[69] 2016	ActiGraph GT3X 1 s Waist (l+r) StepWatch 3 s Ankle (l+r) SenseWear 60 s Upper portion of the arm (l+r)	PIC Cosmed K4b2	Laboratory 2-2.5 h	Cerebral palsy I=28 II=16 III=13	n=57 12.5±3.3 y 29 girls
Pirpiris and Graham[57] 2004	PAL 1 1 s Thigh (l)	Video DO ND	Free living 1 h	Cerebral palsy	n=50 11.0±3.2 y 24 girls
Pirpiris and Graham[57] 2004	PAL 1 1 s Thigh (l)	PIC Cosmed K4b2	ND ND	Cerebral palsy	n=35 ND ND
Stephens et al[26] 2016	Actical 15 s Waist (r) ActiGraph 15 s Waist (r)	PIC Cosmed K4b2/ Cortex Metamax	Laboratory 7 × 6 min	Juvenile arthritis	n=31 12.7±2.6 y 23 girls
Stephens et al[26] 2016	Actical 15 s Waist (r) ActiGraph 15 s Waist (r)	PIC Cosmed K4b2/ Cortex Metamax	Laboratory 7 × 6 min	Inherited muscle disease	n=30 12.0±3.4 y 8 girls

Table 3A (*Continued*)

Study	Device	Comparison Measure	PB Domain	Population	
Author(s) YoP	Name Device Epoch Length Placement Body	Specs Outcome	Activity Duration	Condition GMFCS level (n)	Sample Age Gender
Tang et al[58] 2013	activPAL ND Thigh (nDBS)	Video DO ND	Laboratory 71±49 min	Cerebral palsy I=9 II=4 III=2	n=15 10.9±4.3 y 4 girls
Xing et al[18] 2021	ActiGraph GT3X 15 s Waist (r) activPAL3C 15 s Thigh (front)	PIC Sensor Medics	Laboratory 7 × 5 min	Cerebral palsy I=3 II=6 III=1	n=10 6 girls
Criterion validity of cutoff point based methods					
Baque et al[77] 2017	ActiGraph GTX3+ 15 s Waist (LNI)	DO ND	Laboratory 1 × 5 min+3 × 6 min+ 1 × 3 min		
Clanchy et al[70] 2011	ActiGraph 7164 1 s Waist (LNI or DBS)	PIC Cosmed K4b2 3 min	Laboratory 10 min+2 × 6 min+ 3 min	Cerebral palsy I=11 II=15 III=13	n=29 12.5±2.0 y 13 girls
Keawutan et al[71] 2016	Actigraph GT3X(+) 5 s Trunk (back)	Video DO ND	Laboratory 17.4 min	Cerebral palsy I=17 II=13 ND III=10	n=40 4.7 y
Keawutan et al[71] 2016	Actigraph GT3X(+) 5 s Trunk (back)	Laboratory	Cerebral palsy 19.7 min II=7	n=21 I=9 ND III=5	4.6 y
Oftedal et al[56] 2014	ActiGraph GT1M 5 s Trunk (back) ActiGraph GT3X and GT3X+ 5 s Trunk (back)	Video DO ND	Laboratory 20 à 30 min	Cerebral palsy I-III 26 girls	n=39 2.3±0.5 y
Oftedal et al[56] 2014	ActiGraph GT1M 5 s Trunk (back) ActiGraph GT3X and GT3X+ 5 s Trunk (back)	Laboratory	20 à 30 min	n=23 15 girls	2.5±0.6 y
Ryan et al[73] 2014	RT3 1 Hz ND	PIC Oxycon	Laboratory 6 min	Cerebral palsy I=10 II=4	n=18 11.4±3.2 y 8 girls
Stephens et al[26] 2016	Actical 15 s Hip (mid-line, r) ActiGraph 15 s Waist (r)	PIC Cosmed K4b2/ Cortex Metamax	Laboratory 7 × 6 min	Juvenile arthritis	n=31 12.7±2.6 y 23 girls
Stephens et al[26] 2016	Actical 15 s Hip (mid-line, r) ActiGraph 15 s Waist (r)	PIC Cosmed K4b2/ Cortex Metamax	Laboratory 7 × 6 min	Inherited muscle disease	n=30 12.0±3.4 y 8 girls
Trost et al[24] 2016	ActiGraph GT3X 1 s Waist (r)	PIC Cosmed K4b2	Laboratory 4 × 5 min+3 × 6 min	Cerebral palsy I=27 II=12	n=51 age=12.5 24 girls
Xing et al[18] 2021	ActiGraph GT3X 15 s Waist (r) activPAL3C 15 s Thigh	PIC Sensor Medics	Laboratory 7 × 5 min	Cerebral palsy I=3 II=6 III=1	n=10 6 girls
Convergent validity					
Kuo et al[55] 2009	AMP 331 ND Ankle (r) DynaPort Minimod ND Trunk (back)	Measuring wheel ND	Laboratory ND III=3	Cerebral palsy I=5 II=12	n=20 Age=10.5±3.0 y 7 girls
Lawal et al[21] 2020	ActiGraph GT3X 60 s Waist	Observation Tally counter	Laboratory 6 min	Congenital muscular Dystrophy ND	n=9 Age=ND 6 girls

Table 3A (*Continued*)

Study	Device			Population	
Author(s) YoP	Name Device Epoch Length Placement Body	Comparison Measure Specs Outcome	PB Domain Activity Duration	Condition GMFCS level (n)	Sample Age Gender
Mackey et al[66] 2009	IDEEA ND Chest, thigh (2), foot (2)	Time recording	Laboratory 5 × 30 s	Cerebral palsy I or II=16 III=9	n=25 Age=14.1 y 17 girls
Sala et al[74] 2019	Fitbit One ND Waist Fitbit Flex ND Wrist (DBS)	Observation Tally counter	Laboratory 3 × ±3 min	Cerebral palsy I=22 II=5 III=11	n=39 Age=9.6 y Range 4-15 y 16 girls

Abbreviations: DBS: dominant bodyside; DO, direct observation; IDEEA: Intelligent Device for Energy Expenditure and Activity; l: left; LNI: least neurological impaired; n: number; ND: not described; nDBS: nondominant bodyside; NI: neurological impaired; PAL 1: Positional Activity Logger version 1; PB: physical behavior; PIC: portable indirect calorimetry; r: right; SOFIT: System for Observing Fitness Instruction Time; YoP: year of publication.

Table 3B   Results validity of device-measured physical activity.

Study Author(s) YoP	Population Group Specs	Cutoff points/pred.eq.	Device-Measured PA	Validity		Quality	
						Risk of Bias	GRADE
Criterion validity							⊕⊕⊕
Aviram et al[60] 2011	n=13		EE rate (Kcal/min) slow walking EE rate (Kcal/min) moderate walking EE rate (Kcal/min) stepping total EE (Kcal) structured activities (II)	r=0.7 r=0.88 r=0.75 r=0.72		D	
Bania[59] 2014			Duration sitting Duration standing Steps walking	MD=0.5 min, 95% LoA (−0.6 to 1.5) MD=−0.06 min, 95% LoA (−0.4 to 0.3) MD=−13.8 steps, 95% LoA (−36.9 to 9.4)		D	
Baque et al[77] 2017			AC-VM (n) structured activities (II)	r=0.89		V	
Capio et al[75] 2010			AC (n) structured activities (II) AC (n) unstructured activities (RI)	R^2=0.56 R^2=0.45		V	
de Groot et al[76] 2013	Actical Actiheart	Corder (AC) Corder (AC) Corder (HRAR) Corder (AC+HRAR) Corder (AC+predHRAS) Corder (AC+trueHRAS) Takken (predHRAS) Takken (trueHRAS)	EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking	ICC=0.6; 95% CI, 0.28-0.80 ICC=0.49 (0.12-0.74) ICC=0.74 (0.49-0.88) ICC=0.8; 95% CI, 0.59-0.91 ICC=0.12 (−0.28 to 0.48) ICC=0.71 (0.45-0.86) ICC=0.018; 95% CI, −0.37 to 0.40 ICC=0.73 (0.48-0.87)		V	
	Actical Actiheart	Current study (AC) Current study (HRAR)	EE (J/kg/min) walking EE (J/kg/min) walking	R^2=0.68 R^2=0.72			
Koehler et al[68] 2015	n=10 (NHS) n=9 n=10 n=6 n=2 n=10 (HS) n=9 n=10 n=6 n=2		EE sitting (rest) EE walking (0.85 m/s) EE walking (1.35 m/s) EE walking (1.85 m/s) EE (Kcal/min) running (2.35 m/s) EE sitting (rest) EE walking (0.85 m/s) EE walking (1.35 m/s) EE walking (1.85 m/s) EE running (2.35 m/s)	MD=0.3 Kcal/min, 95% LoA (−0.7 to 1.3) MD=−0.6 Kcal/min, 95% LoA (−3.2 to 2.0) MD=−0.3 Kcal/min, 95% LoA (−3.7 to 3.1) MD=0.8 Kcal/min, 95% LoA (−4.1 to 5.7) NA MD=0.4 Kcal/min, 95% LoA (−0.7 to 1.5) MD=−0.6 Kcal/min, 95% LoA (−2.7 to 1.4) MD=−0.1 Kcal/min, 95% LoA (−2.9 to 2.6) MD=0.7 Kcal/min, 95% LoA (−5.9 to 7.2) MD=−0.6 Kcal/min, 95% LoA (−9.4 to 8.3)		D	
Kuo et al[55] 2009	AMP Minimod AMP Minimod AMP Minimod AMP Minimod		Steps continuous walking Steps continuous walking Steps intermittent walking Steps intermittent walking Steps downstair climbing Steps downstair climbing Steps upstair climbing Steps upstair climbing	MD=−3.5 steps, 95% LoA (−16.9 to 10.0) MD=−11.2 steps, 95% LoA (−40.0 to 17.7) MD=−0.4 step, 95% LoA (−4.1 to 3.3) MD=−38.7 steps, 95% LoA (−87.8 to 104) MD=−10.1 steps, SD=9.5 MD=−1 step, SD=1 MD=−8.1 steps, SD=9.2 MD=−1 step, SD=1.7		D	
Lankhorst et al[16] 2019	n=8 (slow pace); n=7 (fast pace) n=8 n=8		Duration basic sitting Duration basic standing Duration basic walking Duration basic bicycling Duration basic running Duration complex sitting Duration complex standing Duration complex walking Duration complex bicycling Duration complex running	MD=83 s MD=−70 s MD=−385 s MD=357 s MD=15 s MD=−106 s MD=−296 s MD=−372 s MD=713 s MD=61 s	Error=6.9% Error=−11.6% Error=−21% Error=39.7% Error=3.4% Error=−8.1% Error=−84.6% Error=−12.4% NA Error=265.2%	I	

Table 3B (*Continued*)

Study Author(s) YoP	Population Group Specs	Cutoff points/pred.eq.	Device-Measured PA	Validity	Quality Risk of Bias	GRADE
Maher et al[67] 2013	Dominant body side		Steps (n) walking	ICC=0.94; 95% CI, 0.83-0.98	V	
			Steps (n) running	ICC=0.94; 95% CI, 0.84-0.98		
	Non-dominant body side		Steps (n) walking	ICC=0.78; 95% CI, 0.49-0.92		
			Steps (n) running	ICC=0.95; 95% CI, 0.87-0.98		
McAloon et al[72] 2014			Duration sitting	MD=−6.8 s, 95% LoA (18.5 to −32.1)	I	
			Duration standing	MD=5.9 s, 95% LoA (19.1 to −7.3)		
			Duration walking	MD=−2.2 s, 95% LoA (7.8 to −12.3)		
			Steps steps	MD=−3.2 steps, 95% LoA (4.5 to −10.9)		
			Duration (s) structured activities (RI)	Agreement=86.5%		
O'Donoghue and Kennedy[78] 2014	NI		Duration (s) sitting	ICC (3.1)=0.49	D	
			Duration (s) standing	ICC (3.1)=0.59		
			Duration (s) walking	ICC (3.1)=0.99		
			Steps (n) walking	ICC (3.1)=0.96		
	LNI		Duration (s) sitting	ICC (3.1)=0.95		
			Duration (s) standing	ICC (3.1)=0.98		
			Duration (s) walking	ICC (3.1)=0.94		
			Steps (n) walking	ICC (3.1)=0.95		
O'Neil et al[20] 2014			AC (n) walking (slow, brisk, fast)	rs=0.67	V	
			Steps (n) walking (slow, brisk, fast)	rs=0.29		
O'Neil et al[69] 2016	ActiGraph		Steps (n) structured activities	rs=0.82	V	
			AC-VA (n) structured activities	rs=0.835		
			AC-VM (n) structured activities	rs=0.84		
	StepWatch		Steps (n) structured activities	rs=0.78		
	SenseWear		Steps (n) structured activities	rs=0.74		
Pirpiris and Graham[57] 2004			Duration upright (s) unstructured activities	MD=5 s, 95% LoA (−37 to 47)	D	
Pirpiris and Graham[57] 2004			Duration upright (s)	rs=0.61		
Stephens et al[26] 2016	Actical	Puyau	EE (METs) structured activities (II)	ICC (2,1)=0.41; 95% CI, 0.30-0.52	V	
	ActiGraph	Corder	EE (METs)	ICC (2,1)=0.28; 95% CI, 0.14-0.40		
		Freedson	EE (METs)	ICC (2,1)=0.35; 95% CI, 0.22-0.47		
		Puyau	EE (METs)	ICC (2,1)=0.22; 95% CI, 0.09-0.35		
	Actical	Current study	EE (METs)	ICC (2,1)=0.75; 95% CI, 0.74-0.76		
		Current study (HR)	EE (METs)	ICC (2,1)=0.85; 95% CI, 0.84-0.86		
	ActiGraph	Current study	EE (METs)	ICC (2,1)=0.7; 95% CI, 0.69-0.71		
		Current study (HR)	EE (METs)	ICC (2,1)=0.84; 95% CI, 0.83-0.85		
Stephens et al[26] 2016	Actical	Puyau	EE (METs) structured activities (II)	ICC (2,1)=0.38; 95% CI, 0.25-0.50		
	ActiGraph	Corder	EE (METs)	ICC (2,1)=0.31; 95% CI, 0.17-0.45		
		Freedson	EE (METs)	ICC (2,1)=0.36; 95% CI, 0.21-0.49		
		Puyau	EE (METs)	ICC (2,1)=0.25; 95% CI, 0.10-0.39		
	Actical	Current study	EE (METs)	ICC (2,1)=0.74; 95% CI, 0.73-0.75		
		Current study (HR)	EE (METs)	ICC (2,1)=0.78; 95% CI, 0.77-0.79		
	ActiGraph	Current study	EE (METs)	ICC (2,1)=0.71; 95% CI, 0.70-0.72		
		Current study	EE (METs)	ICC (2,1)=0.68; 95% CI, 0.66-0.70		
Tang et al[58] 2013			Duration (min) sitting/lying	Error=−2.6%	I	
			Duration (min) upright	Error=1.1%		
			Duration (min) standing	Error=−0.5%		
			Duration (min) stepping	Error=5.6%		
			Steps (n) walking	Error=3.8%		
Xing et al[18] 2021	ActiGraph	Freedman (VA)	EE sitting	MD=−0.58 METs; 95% CI, −0.64 to −0.44	V	
		Trost (VA)	EE sitting	MD=−0.24 Kcal/min; 95% CI, −0.46 to −0.02		
		Treuth (VA)	EE sitting	MD=−1.13 METs; 95% CI, −1.18 to −1.07		
	activPAL		EE sitting	MD=−0.37 METs; 95% CI, −0.42 to −0.31		
	ActiGraph	Freedman	EE standing	MD=−0.13 METs; 95% CI, −0.27 to 0.02		
		Trost	EE standing	MD=0.13 Kcal/min; 95% CI, −0.09 to 0.33		
		Treuth	EE standing	MD=−0.71 METs; 95% CI, −0.81 to −0.62		
	activPAL		EE standing	MD=−0.16 METs; 95% CI, −0.28 to −0.05		
	ActiGraph	Freedman	EE slow walking	MD=0.46 METs; 95% CI, 0.21-0.70		
		Trost	EE slow walking	MD=0.75 Kcal/min; 95% CI, 0.59-0.91		
		Treuth	EE slow walking	MD=0.02 METs; 95% CI −0.14 to 0.19		
	activPAL		EE slow walking	MD=−0.44 METs; 95% CI −0.60 to −0.27]		
	ActiGraph	Freedman	EE moderate walking	MD=−0.17 METs; 95% CI, −0.50 to 0.17		
		Trost	EE moderate walking	MD=0.51 Kcal/min, 95% CI, 0.28-0.75		
		Treuth	EE moderate walking	MD=−0.29 METs, 95% CI, −0.47 to −0.10		
	activPAL		EE moderate walking	MD=0.39 METs, 95% CI, 0.09, 0.68		
	ActiGraph	Freedman	EE fast walking	MD=1.84 METs; 95% CI, 1.26, 2.43		
		Trost	EE fast walking	MD=2.17 Kcal/min, 95% CI, 1.33, 3.01		
		Treuth	EE fast walking	MD=1.48 METs, 95% CI, 0.94, 2.02		
	activPAL		EE fast walking	MD=3.54 METs, 95% CI, 2.60, 4.48		

Table 3B (*Continued*)

Study Author(s) YoP	Population Group Specs	Cutoff points/pred.eq.	Device-Measured PA	Validity			Quality Risk of Bias	GRADE
Criterion validity for cutoff point based methods								⊕⊕⊕
Baque et al[77] 2017		Baque	AC-VM (n/15 s) structured activities (SED)	AUC=1	Se=100%	Sp=100%	V	
			AC-VM (n/15 s) structured activities (MI)	AUC=0.98	Se=90.7%	Sp=92.6%		
			AC-VM (n/15 s) structured activities (VI)	AUC=0.99	Se=93.8%	Sp=96.3%		
		Clanchy	PA intensity levels (SED, LPA, MPA, VPA) (METs) structured activities (II)	κ=0.73 (SE=−0.08)				
		Evenson	PA intensity levels (SED, LPA, MPA, VPA) (METs) structured activities (II)	κ=0.77 (SE=−0.07)				
		Baque	PA intensity levels (SED, LPA, MPA, VPA) (METs) structured activities (II)	κ=0.92 (SE=−0.07)				
Clanchy et al[70] 2011		Freedson	SED≤1.5 (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%	V	
		Puyau	SED (METs) structured activities (II)	AUC=0.90, 95% CI, 0.84-0.96	Se=90.3%	Sp=89.4%		
		Treuth	SED (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%		
		Evenson	SED (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%		
		Clanchy	SED (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%		
		Freedson	LPA=[1.5, 4] (METs) structured activities (II)	AUC=0.66, 95% CI, 0.58-0.74	Se=45.3%	Sp=85.7%		
		Puyau	LPA (METs) structured activities (II)	AUC=0.68, 95% CI, 0.59-0.77	Se=63.2%	Sp=73.1%		
		Treuth	LPA (METs) structured activities (II)	AUC=0.68, 95% CI, 0.60-0.75	Se=60.4%	Sp=76.2%		
		Evenson	LPA (METs) structured activities (II)	AUC=0.67, 95% CI, 0.58-0.75	Se=49.1%	Sp=84.1%		
		Clanchy	LPA (METs) structured activities (II)	AUC=0.63, 95% CI, 0.56-0.71	Se=35.9%	Sp=90.5%		
		Freedson	MVPA=≥4 (METs) structured activities (II)	AUC=0.75, 95% CI, 0.66-0.83	Se=81.8%	Sp=67.5%		
		Puyau	MVPA (METs) structured activities (II)	AUC=0.75, 95% CI, 0.68-0.83	Se=59.6%	Sp=91.3%		
		Treuth	MVPA (METs) structured activities (II)	AUC=0.73, 95% CI, 0.63-0.82	Se=66.7%	Sp=78.3%		
		Evenson	MVPA (METs) structured activities (II)	AUC=0.91, 95% CI, 0.84-0.97	Se=81.8%	Sp=100%		
		Clanchy	MVPA (METs) structured activities (II)	AUC=0.94, 95% CI, 0.88-0.98	Se=91.4%	Sp=86.2%		
Keawutan et al[71] 2016	GMFCS I	Keawutan	AC-VM (n) structured activities (SED)	AUC=0.79, 95% CI, 0.77-081	Se=74%	Sp=73%	D	
	GMFCS II		AC-VM (n) structured activities (SED)	AUC=0.78, 95% CI, 0.76-0.80	Se=72%	Sp=73%		
	GMFCS III		AC-VM (n) structured activities (SED)	AUC=0.81, 95% CI, 0.79-0.82	Se=74%	Sp=74%		
Keawutan et al[71] 2016	GMFCS I	Keawutan	Duration SED (min) structured activities (II)	MD=−13.3%, 95% LoA [−34.0 to 7.4]	Se=78.6%	Sp=84%		
		Butte	Duration SED (min) structured activities (II)	MD=−6.2%, 95% LoA [−22.8 to 10.4]	Se=84.1%	Sp=79.6%		
	GMFCS II	Keawutan	Duration SED (min) structured activities (II)	MD=−15.6%, 95% LoA [−33.5 to 2.3]	Se=78.2%	Sp=86.7%		
		Butte	Duration SED (min) structured activities (II)	MD=−10.4%, 95% LoA [−27.2 to 6.4]	Se=81.7%	Sp=83.2%		
	GMFCS III	Keawutan	Duration SED (min) structured activities (II)	MD=−1%, 95% LoA [−25.6 to 23.7]	Se=72.5%	Sp=76.2%		
		Butte	Duration SED (min) structured activities (II)	MD=−2.4%, 95% LoA [−26.8 to 21.9]	Se=71.8%	Sp=77%		
Oftedal et al[56] 2014	n=39	Oftedal (VM)	AC-VA (n) structured activities (SED)	AUC=0.77, 95% CI, 0.76-0.78	Se=71%	Sp=77%	D	
	n=18		AC-VM (n) structured activities (SED)	AUC=0.81, 95% CI, 0.80-0.82	Se=75%	Sp=76%		
Oftedal et al[56] 2014	n=23	Oftedal (VA)	Duration SED (min) structured activities (II)	MD=−10.5%, 95% LoA [−30.2 to 9.1]	Se=74%	Sp=80%	D	
	n=18		Duration SED (min) structured activities (II)	MD=−1.5%, 95% LoA [−20.0 to 16.8]	Se=79%	Sp=72%		
Ryan et al[73] 2014		Ryan	AC (n) structured activities (LI)	AUC=0.965, 95% CI, 84.6-99.8			D	
			AC (n) structured activities (MVI)	AUC=0.896, 95% CI, 77.4-96.6				
		VanHelst	SED (<2) (METs) structured activities (II)	κ=0.92, 95% CI, 0.82-1.00	Se=89.5%	Sp=100%		
		Ryan	SED (<2) (METs) structured activities (II)	κ=0.96, 95% CI, 0.89-1.00	Se=94.7%	Sp=100%		
		VanHelst	LPA ([2, 3]) (METs) structured activities (II)	κ=0.57, 95% CI, 0.38-0.77	Se=88.9%	Sp=79.6%		
		Ryan	LPA (METs) structured activities (II)	κ=0.71, 95% CI, 0.52-0.90	Se=83.3%	Sp=89.8%		
		Rowlands	MVPA (≥3) (METs) structured activities (II)	κ=0.69, 95% CI, 0.52-0.86	Se=70%	Sp=97.3%		
		VanHelst	MVPA (METs) structured activities (II)	κ=0.66, 95% CI, 0.48-0.84	Se=70%	Sp=94.6%		
		Ryan	MVPA (METs) structured activities (II)	κ=0.79, 95% CI, 0.64-0.94	Se=86.7%	Sp=91.9%		
Stephens et al[26] 2016	Actical	Stephens	AC-VM (n) structured activities (SED)	AUC=0.84, 95% CI, 0.78-0.90	Se=78%	Sp=88%	V	
			AC-VM (n) structured activities (MI)	AUC=0.82, 95% CI, 0.77-0.88	Se=72%	Sp=75%		
			AC-VM (n) structured activities (VI)	AUC=0.98, 95% CI, 0.96-1.0	Se=100%	Sp=94%		
	ActiGraph	Stephens	AC-VA (n) structured activities (SED)	AUC=0.82, 95% CI, 0.74-0.90	Se=75%	Sp=91%		
			AC-VA (n) structured activities (MI)	AUC=0.78, 95% CI, 0.71-0.86	Se=86%	Sp=63%		
			AC-VA (n) structured activities (VI)	AUC=0.78, 95% CI, 0.52-1.0	Se=83%	Sp=79%		
	Actical	Evenson	SED (n) structured activities (RI)		Se=72%	Sp=92%		
			MPA (n) structured activities (RI)		Se=49%	Sp=94%		
			VPA (n) structured activities (RI)		Se=100%	Sp=92%		
	ActiGraph	Evenson	SED (n) structured activities (RI)		Se=75%	Sp=90%		
			MPA (n) structured activities (RI)		Se=41%	Sp=90%		
			VPA (n) structured activities (RI)		Se=50%	Sp=95%		
Stephens et al[26] 2016	Actical	Stephens	AC-VM (n) structured activities (SED)	AUC=0.96, 95% CI, 0.93-0.98	Se=82%	Sp=97%	V	
			AC-VM (n) structured activities (MI)	AUC=0.89, 95% CI, 0.83-0.94	Se=82%	Sp=81%		
			AC-VM (n) structured activities (VI)	AUC=0.91, 95% CI, 0.87-0.95	Se=100%	Sp=90%		
	ActiGraph	Stephens	AC-VA (n) structured activities (SED)	AUC=0.9, 95% CI, 0.86-0.95	Se=78%	Sp=91%		
			AC-VA (n) structured activities (MI)	AUC=0.91, 95% CI, 0.85-0.97	Se=81%	Sp=94%		
			AC-VA (n) structured activities (VI)	AUC=0.92, 95% CI, 0.88-0.96	Se=100%	Sp=92%		
	Actical	Evenson	SED (n) structured activities (RI)		Se=80%	Sp=97%		
			MPA (n) structured activities (RI)		Se=47%	Sp=96%		
			VPA (n) structured activities (RI)		Se=0%	Sp=96%		

(*continued*)

Table 3B (*Continued*)

Study Author(s) YoP	Population Group Specs	Cutoff points/pred.eq.	Device-Measured PA	Validity			Quality Risk of Bias	GRADE
	ActiGraph	Evenson	SED (n) structured activities (RI)		Se=75%	Sp=91%		
			MPA (n) structured activities (RI)		Se=81%	Sp=90%		
			VPA (n) structured activities (RI)		Se=0%	Sp=92%		
Trost et al[24] 2016		Evenson	SED (<1.5) (METs) structured activities (II)	AUC=0.93, 95% CI, 0.91-0.96	Se=98.9%	Sp=87.6%	V	
		Clanchy	SED (METs) structured activities (II)	AUC=0.93, 95% CI, 0.91-0.96	Se=98.9%	Sp=87.6%		
		Trost (VA)	SED (METs) structured activities (II)	AUC=0.97, 95% CI, 0.95-0.99	Se=97.9%	Sp=96.1%		
		Trost (VM)	SED (METs) structured activities (II)	AUC=0.96, 95% CI, 0.94-0.99	Se=96.9%	Sp=96.1%		
		Evenson	LPA ([1.5, 3]) (METs) structured activities (II)	AUC=0.68, 95% CI, 0.63-0.73	Se=61.5%	Sp=74.6%		
		Clanchy	LPA (METs) structured activities (II)	AUC=0.68, 95% CI, 0.62-0.73	Se=58.1%	Sp=77.1%		
		Trost (VA)	LPA (METs) structured activities (II)	AUC=0.82, 95% CI, 0.77-0.86	Se=77.8%	Sp=86.4%		
		Trost (VM)	LPA (METs) structured activities (II)	AUC=0.8, 95% CI, 0.76-0.85	Se=72.7%	Sp=87.8%		
		Evenson	MVPA (≥3) (METs) structured activities (II)	AUC=0.75, 95% CI, 0.71-0.80	Se=57.1%	Sp=93.4%		
		Clanchy	MVPA (METs) structured activities (II)	AUC=0.76, 95% CI, 0.72-0.81	Se=61.4%	Sp=91.6%		
		Trost (VA)	MVPA (METs) structured activities (II)	AUC=0.86, 95% CI, 0.82-0.89	Se=78.6%	Sp=92.5%		
		Trost (VM)	MVPA (METs) structured activities (II)	AUC=0.86, 95% CI, 0.82-0.89	Se=81.4%	Sp=89.7%		
Xing et al[18] 2021		Puyau (VA)	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.84	κ=0.458		V	
		Evenson	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.888	κ=0.585			
		Romanzini (VA)	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.886	κ=0.56			
		Romanzini (VM)	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.886	κ=0.675			
		Clanchy	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.935	κ=0.721			
		Baque	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.896	κ=0.773			

Convergent validity								⊕⊕⊕
Kuo et al[55] 2009	AMP		Distance continuous walking	MD=−4.8 m, 95% LoA (-20.1 to 10.5)			D	
	Minimod		Distance continuous walking	MD=−0.4 m, 95% LoA (-4.7 to 4.0)				
	AMP		Distance intermittent walking	MD=−3.6 m, 95% LoA (-19.2 to 12.0)				
	Minimod		Distance intermittent walking	MD=−2.3 m, 95% LoA (-27.9 to 23.3)				
	AMP		Distance downstair climbing	MD=−1.3 m, SD -2.5				
	Minimod		Distance downstair climbing	MD=8.9 m, SD -2.5				
	AMP		Distance upstair climbing	MD=−2 m, SD -2.5				
	Minimod		Distance upstair climbing	MD=3.3 m, SD -2.2				
Lawal et al[21] 2020	ActiGraph GT3X		Steps (n) walking	ICC=0.29, 95% CI, −0.42 to 0.78			D	
	LFE-ActiGraph GT3X		Steps (n) walking	ICC=0.52, 95% CI, −0.16 to 0.87				
Mackey et al[66] 2009			Duration (min) lying	Se=100%	Sp=100%		D	
			Duration (min) sitting	Se=100%	Sp=100%			
			Duration (min) standing	Se=100%	Sp=97%			
	n=12 for stair climbing		Duration (min) walking (overground and stairs)	Se=78.5%	Sp=100%			
Sala et al[74] 2019	Hip (n=38)		Steps walking	MAE=7 steps, range=−52 to 6	r=0.991		D	
			Distance walking	MAE=0.07 miles, range=0.01 to 0.16				
	GMFCS I+II (n=27)		Steps walking	MAE=6 steps, range=−20 to 6	r=0.998			
			Distance walking	MAE=0.07 miles, range=0.01 to 0.16				
	GMFCS III (n=11)		Steps walking	MAE=12 steps, range=−52 to 1	r=0.981			
			Distance walking	MAE=0.07 miles, range=0.02 to 0.14				
	Wrist (n=38)		Steps walking	MAE=88 steps, range=−484 to 35	r=−0.033			
			Distance walking	MAE=0.06 miles, range=−0.13 to 0.16				
	GMFCS I+II (n=27)		Steps walking	MAE=27 steps, range=−177 to 23	r=0.837			
			Distance walking	MAE=0.04 miles, range=−0.02 to 0.16				
	GMFCS III (n=11)		Steps walking	MAE=238 steps, range=−484 to 35	r=−0.242			
			Distance walking	MAE=0.08 miles, range=−0.13 to 0.11				

Abbreviations: "?", indeterminate; "-", insufficient; "+", sufficient; "⊕⊕⊕", moderate; AC, activity counts; AUC, area under the curve; D, doubtful; EE, energy expenditure; GRADE, Grading of Recommendations Assessment, Development and Evaluation; HR, heart rate; HRAR, heart rate above rest; HRAS, heart rate above sleep; HS, hemiparetic side; I, inadequate; II, increased intensity; LNI, least neurological impaired; LoA, levels of agreement; LPA, low physical activity; MAE, mean absolute error; MD, mean difference; MET, metabolic equivalent of task; MI, moderate intensity; MPA, moderate physical activity; MVI, moderate-to-vigorous; n, number; NA, not applicable; NHS, nonhemiparetic side; NI, neurological impaired; pred, prediction; r, Pearson correlation coefficient; R^2, coefficient of determination; RI, random intensity; rs, Spearman rank correlation coefficient; Se, sensitivity; SED, sedentary; Sp, specificity; V, very good; VA, vertical axis; VI, vigorous intensity; VM, vector magnitude; VPA, vigorous physical activity; YoP, year of publication; κ, Cohen's kappa.

Table 4A   Validity of device-measured physical activity.

Study	Instrument			Population	
Author(s)	Name	Comparison Measure	PB Domain	Condition	Sample
YoP	Epoch	Specs	Activity Duration	GMFCS level (n)	Age
	Placement	Outcome			Gender
Criterion validity					
Aviram et al[60]	IDEEA	PIC	Laboratory	Cerebral palsy	n=21
2011	ND	Cosmed K4b2	1 × ±7 min+3 × 4 min	I=8	6.4±1.9 y
	Chest, thigh (2),		II=6	ND	
	Foot (2)			III=7	
Bania[59]	activPAL	Video DO	Laboratory	Cerebral palsy	n=10
2014	ND	ND	2 × 3 min+6 min	II=5	18.6±2.7 y
	Thigh			III=5	4 girls
Baque et al[77]	ActiGraph GTX3+	PIC	Laboratory	Acquired brain injury	n=27
2017	15 s	Cortex Metamax	1 × 5 min+3 ×	I=16	13.6±2.4 y
	Waist (LNI)		6 min+1 × 3 min	II=11	12 girls
Capio et al[75]	MTI (ActiGraph)	SOFIT	Laboratory	Cerebral palsy	n=31
2010	15 s	ND	6 × 2 min (structured)	I=14	9.7±2.5 y
	Waist (r)		10 min (free play)	II=9	17 girls
				III=8	
de Groot et al[76]	Actical	PIC	Laboratory	Spina bifida	n=39
2013	60 s	Cortex Metamax	6 min	Hoffer normal=7	10.6±2.8 y
	Waist (l)			Hoffer community=19	11 girls
	Actiheart				
	15 s				
	Trunk (front)				
Koehler et al[68]	SenseWear	PIC	Laboratory	Cerebral palsy	n=10
2015	ND	ZAN 600	5 × 5 min	II=10	13.4±1.6 y
	Upper portion of the arm (l+r)				3 girls
Kuo et al[55]	AMP 331	Video DO	Laboratory	Cerebral palsy	n=20
2009	ND	ND	ND	I=5	10.5±3.0 y
	Ankle (r)			II=12	7 girls
	DynaPort Minimod		III=3		
	ND				
	Trunk (back)				
Lankhorst et al[16]	Activ8	Video DO	Free living	Spina bifida	n=10
2019	5 s	ND	45 min (1:30 min)	Hoffer community=2	12.9±2.1 y
	Thigh (DBS)			Hoffer household=2	1 girl
Maher et al[67]	NL-1000 pedometer	Video DO	Laboratory	Cerebral palsy	n=17
2013	ND	FlipVideo UltraHD	2 × 3 min	I=8	12.3±3.2 y
	Thigh (nDBS, DBS)	Camcorder		II=9	9 girls
McAloon et al[72]	activPAL	Video DO	Laboratory	Cerebral palsy	n=10
2014	ND	ND	ND	I=4	4-18 y
	Thigh (NI)			II=5	ND
				III=1	
O'Donoghue and Kennedy[78]	activPAL	Video DO	Laboratory	Cerebral palsy	n=17
2014	ND	2D Sony mini digital	ND	I=17	9.4±3.9 y
	Thigh (NI, LNI)	Video camera			8 girls
O'Neil et al[20]	ActiGraph GT3X	PIC	Laboratory	Cerebral palsy	n=8
2014	1 s	Cosmed K4b2	3 × 6 min	I=4	11.9±3.2 y
	Waist (r)			II=1	2 girls
				III=3	
O'Neil et al[69]	ActiGraph GT3X	PIC	Laboratory	Cerebral palsy	n=57
2016	1 s	Cosmed K4b2	2-2.5 h	I=28	12.5±3.3 y
	Waist (l+r)			II=16	29 girls
	StepWatch			III=13	
	3 s				
	Ankle (l+r)				
	SenseWear				
	60 s				
	Upper portion of the arm (l+r)				
Pirpiris and Graham[57]	PAL 1	Video DO	Free living	Cerebral palsy	n=50
2004	1 s	ND	1 h		11.0±3.2 y
	Thigh (l)				24 girls
Pirpiris and Graham[57]	PAL 1	PIC	ND	Cerebral palsy	n=35
2004	1 s	Cosmed K4b2	ND		ND
	Thigh (l)				ND
Stephens et al[26]	Actical	PIC	Laboratory	Juvenile arthritis	n=31
2016	15 s	Cosmed K4b2/	7 × 6 min		12.7±2.6 y
	Waist (r)	Cortex Metamax			23 girls
	ActiGraph				
	15 s				
	Waist (r)				
Stephens et al[26]	Actical	PIC	Laboratory	Inherited muscle disease	n=30
2016	15 s	Cosmed K4b2/	7 × 6 min		12.0±3.4 y
	Waist (r)	Cortex Metamax			8 girls
	ActiGraph				
	15 s				
	Waist (r)				

Table 4A (*Continued*)

Study	Instrument			Population	
Author(s) YoP	Name Epoch Placement	Comparison Measure Specs Outcome	PB Domain Activity Duration	Condition GMFCS level (n)	Sample Age Gender
Tang et al[58] 2013	activPAL ND Thigh (nDBS)	Video DO ND	Laboratory 71±49 min	Cerebral palsy I=9 II=4 III=2	n=15 10.9±4.3 y 4 girls
Xing et al[18] 2021	ActiGraph GT3X 15 s Waist (r) activPAL3C 15 s Thigh (front)	PIC Sensor Medics	Laboratory 7 × 5 min	Cerebral palsy I=3 II=6 III=1	n=10 6 girls
Criterion validity of cutoff point based methods					
Baque et al[77] 2017	ActiGraph GTX3+ 15 s Waist (LNI)	DO ND	Laboratory 1 × 5 min+3 × 6 min+ 1 × 3 min		
Clanchy et al[70] 2011	ActiGraph 7164 1 s Waist (LNI or DBS)	PIC Cosmed K4b2 3 min	Laboratory 10 min+2 × 6 min+ II=15	Cerebral palsy I=11 III=13	n=29 12.5±2.0 y 13 girls
Keawutan et al[71] 2016	Actigraph GT3X(+) 5 s Trunk (back)	Video DO ND	Laboratory 17.4 min II=13	Cerebral palsy I=17 ND III=10	n=40 4.7 y
Keawutan et al[71] 2016	Actigraph GT3X(+) 5 s Trunk (back)	Laboratory	Cerebral palsy 19.7 min II=7	n=21 I=9 ND III=5	4.6 y
Oftedal et al[56] 2014	ActiGraph GT1M 5 s Trunk (back) ActiGraph GT3X and GT3X+ 5 s Trunk (back)	Video DO ND	Laboratory 20 à 30 min	Cerebral palsy I-III 26 girls	n=39 2.3±0.5 y
Oftedal et al[56] 2014	ActiGraph GT1M 5 s Trunk (back) ActiGraph GT3X and GT3X+ 5 s Trunk (back)	Laboratory	20 à 30 min	n=23 15 girls	2.5±0.6 y
Ryan et al[73] 2014	RT3 1 Hz ND	PIC Oxycon	Laboratory 6 min	Cerebral palsy I=10 II=4	n=18 11.4±3.2 y 8 girls
Stephens et al[26] 2016	Actical 15 s Hip (mid-line, r) ActiGraph 15 s Waist (r)	PIC Cosmed K4b2/ Cortex Metamax	Laboratory 7 × 6 min	Juvenile arthritis	n=31 12.7±2.6 y 23 girls
Stephens et al[26] 2016	Actical 15 s Hip (mid-line, r) ActiGraph 15 s Waist (r)	PIC Cosmed K4b2/ Cortex Metamax	Laboratory 7 × 6 min	Inherited muscle disease	n=30 12.0±3.4 y 8 girls
Trost et al[24] 2016	ActiGraph GT3X 1 s Waist (r)	PIC Cosmed K4b2	Laboratory 4 × 5 min+3 × 6 min	Cerebral palsy I=27 II=12	n=51 Age=12.5 24 girls
Xing et al[18] 2021	ActiGraph GT3X 15 s Waist (r) activPAL3C 15 s Thigh	PIC Sensor Medics	Laboratory 7 × 5 min	Cerebral palsy I=3 II=6 III=1	n=10 6 girls
Convergent validity					
Kuo et al[55] 2009	AMP 331 ND Ankle (r) DynaPort Minimod ND Trunk (back)	Measuring wheel ND	Laboratory ND III=3	Cerebral palsy I=5 II=12	n=20 Age=10.5±3.0 y 7 girls
Lawal et al[21] 2020	ActiGraph GT3X 60 s Waist	Observation Tally counter	Laboratory 6 min	Congenital muscular Dystrophy ND	n=9 Age=ND 6 girls

Table 4A (*Continued*)

Study	Instrument			Population	
Author(s) YoP	Name Epoch Placement	Comparison Measure Specs Outcome	PB Domain Activity Duration	Condition GMFCS level (n)	Sample Age Gender
Mackey et al[66] 2009	IDEEA ND Chest, thigh (2), foot (2)	Time recording	Laboratory 5 × 30 s	Cerebral palsy I or II=16 III=9	n=25 Age=14.1 y 17 girls
Sala et al[74] 2019	Fitbit One ND Waist Fitbit Flex ND Wrist (DBS)	Observation Tally counter	Laboratory 3 × ±3 min	Cerebral palsy I=22 II=5 III=11	n=39 Age=9.6 y Range 4-15 y 16 girls

Abbreviations: DBS, dominant bodyside; DO, direct observation; IDEEA, Intelligent Device for Energy Expenditure and Activity; l, left; LNI, least neurological impaired; n, number; ND, not described; nDBS, nondominant bodyside; NI, neurological impaired; PAL 1, Positional Activity Logger version 1; PB, physical behavior; PIC, portable indirect calorimetry; r, right; SOFIT, System for Observing Fitness Instruction Time; YoP, year of publication.

Table 4B   Validity of device-measured physical activity.

Study	Population		Device-Measured PA	Validity		Quality		
Author(s) YoP	Group specs	Cutoff points/ pred.eq.				Overall Rating	Risk of Bias	GRADE
Criterion validity								⊕⊕⊕
Aviram et al[60] 2011	n=13		EE rate (Kcal/min) slow walking EE rate (Kcal/min) moderate walking EE rate (Kcal/min) stepping Total EE (Kcal) structured activities (II)	r=0.7 r=0.88 r=0.75 r=0.72		+	D	
Bania[59] 2014			Duration sitting Duration standing Steps walking	MD=0.5 min, 95% LoA (-0.6 to 1.5) MD=−0.06 min, 95% LoA (-0.4 to 0.3) MD=−13.8 steps, 95% LoA (-36.9 to 9.4)		?	D	
Baque et al[77] 2017			AC-VM (n) structured activities (II)	r=0.89		+	V	
Capio et al[75] 2010			AC (n) structured activities (II) AC (n) unstructured activities (RI)	R^2=0.56 R^2=0.45		?	V	
de Groot et al[76] 2013	Actical Actiheart Actical Actiheart	Corder (AC) Corder (AC) Corder (HRAR) Corder (AC+HRAR) Corder (AC+predHRAS) Corder (AC+trueHRAS) Takken (predHRAS) Takken (trueHRAS) Current study (AC) Current study (HRAR)	EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking EE (J/kg/min) walking	ICC=0.6, 95% CI, 0.28-0.80 ICC=0.49 (0.12-0.74) ICC=0.74 (0.49-0.88) ICC=0.8, 95% CI, 0.59-0.91 ICC=0.12 (-0.28 to 0.48) ICC=0.71 (0.45-0.86) ICC=0.018, 95% CI, −0.37 to 0.40 ICC=0.73 (0.48-0.87) R^2=0.68 R^2=0.72		-/?/+	V	
Koehler et al[68] 2015	n=10 (NHS) n=9 n=10 n=6 n=2 n=10 (HS) n=9 n=10 n=6 n=2		EE sitting (rest) EE walking (0.85 m/s) EE walking (1.35 m/s) EE walking (1.85 m/s) EE (Kcal/min) running (2.35 m/s) EE sitting (rest) EE walking (0.85 m/s) EE walking (1.35 m/s) EE walking (1.85 m/s) EE running (2.35 m/s)	MD=0.3 Kcal/min, 95% LoA (-0.7 to 1.3) MD=−0.6 Kcal/min, 95% LoA (-3.2 to 2.0) MD=−0.3 Kcal/min, 95% LoA (-3.7 to 3.1) MD=0.8 Kcal/min, 95% LoA (-4.1 to 5.7) NA MD=0.4 Kcal/min, 95% LoA (-0.7 to 1.5) MD=−0.6 Kcal/min, 95% LoA (-2.7 to 1.4) MD=−0.1 Kcal/min, 95% LoA (-2.9 to 2.6) MD=0.7 Kcal/min, 95% LoA (-5.9 to 7.2) MD=−0.6 Kcal/min, 95% LoA (-9.4 to 8.3)		?	D	
Kuo et al[55] 2009	AMP Minimod AMP Minimod AMP Minimod AMP Minimod		Steps continuous walking Steps continuous walking Steps intermittent walking Steps intermittent walking Steps downstair climbing Steps downstair climbing Steps upstair climbing Steps upstair climbing	MD=−3.5 steps, 95% LoA (-16.9 to 10.0) MD=−11.2 steps, 95% LoA (-40.0 to 17.7) MD=−0.4 step, 95% LoA (-4.1 to 3.3) MD=−38.7 steps, 95% LoA (-87.8 to 104) MD=−10.1 steps, SD=9.5 MD=−1 step, SD=1 MD=−8.1 steps, SD=9.2 MD=−1 step, SD=1.7		?	D	
Lankhorst et al[16] 2019	 n=8 (slow pace); n=7 (fast pace) n=8 n=8		Duration basic sitting Duration basic standing Duration basic walking Duration basic bicycling Duration basic running Duration complex sitting Duration complex standing Duration complex walking Duration complex bicycling Duration complex running	MD=83 s MD=−70 s MD=−385 s MD=357 s MD=15 s MD=−106 s MD=−296 s MD=−372 s MD=713 s MD=61 s	Error=6.9% Error=−11.6% Error=−21% Error=39.7% Error=3.4% Error=−8.1% Error=−84.6% Error=−12.4% NA Error=265.2%	?	I	
Maher et al[67] 2013	Dominant body side Nondominant body side		Steps (n) walking Steps (n) running Steps (n) walking Steps (n) running	ICC=0.94, 95% CI, 0.83-0.98 ICC=0.94, 95% CI, 0.84-0.98 ICC=0.78, 95% CI, 0.49-0.92 ICC=0.95, 95% CI, 0.87-0.98		+	V	

Table 4B (*Continued*)

Study	Population		Device-Measured PA	Validity			Quality		
Author(s) YoP	Group specs	Cutoff points/ pred.eq.					Overall Rating	Risk of Bias	GRADE
McAloon et al[72] 2014			Duration sitting	MD=−6.8 s, 95% LoA (18.5 to -32.1)			?	I	
			Duration standing	MD=5.9 s, 95% LoA (19.1 to -7.3)					
			Duration walking	MD=−2.2 s, 95% LoA (7.8 to -12.3)					
			Steps steps	MD=−3.2 steps, 95% LoA (4.5 to -10.9)					
			Duration (s) structured activities (RI)		Agreement=86.5%				
O'Donoghue and Kennedy[78] 2014	NI		Duration (s) sitting	ICC (3.1)=0.49			-/+	D	
			Duration (s) standing	ICC (3.1)=0.59					
			Duration (s) walking	ICC (3.1)=0.99					
			Steps (n) walking	ICC (3.1)=0.96					
	LNI		Duration (s) sitting	ICC (3.1)=0.95					
			Duration (s) standing	ICC (3.1)=0.98					
			Duration (s) walking	ICC (3.1)=0.94					
			Steps (n) walking	ICC (3.1)=0.95					
O'Neil et al[20] 2014			AC (n) walking (slow, brisk, fast)	rs=0.67			-	V	
			Steps (n) walking (slow, brisk, fast)	rs=0.29					
O'Neil et al[69] 2016	ActiGraph		Steps (n) structured activities	rs=0.82			+	V	
			AC-VA (n) structured activities	rs=0.835					
			AC-VM (n) structured activities	rs=0.84					
	StepWatch		Steps (n) structured activities	rs=0.78					
	SenseWear		Steps (n) structured activities	rs=0.74					
Pirpiris and Graham[57] 2004			Duration upright (s) unstructured activities	MD=5 s, 95% LoA (-37 to 47)			?	D	
Pirpiris and Graham[57] 2004			Duration upright (s)	rs=0.61			-		
Stephens et al[26] 2016	Actical	Puyau	EE (METs) structured activities (II)	ICC (2,1)=0.41, 95% CI, 0.30-0.52			-/+	V	
	ActiGraph	Corder	EE (METs)	ICC (2,1)=0.28, 95% CI, 0.14-0.40					
		Freedson	EE (METs)	ICC (2,1)=0.35, 95% CI, 0.22-0.47					
		Puyau	EE (METs)	ICC (2,1)=0.22, 95% CI, 0.09-0.35					
	Actical	Current study	EE (METs)	ICC (2,1)=0.75, 95% CI, 0.74-0.76					
		Current study (HR)	EE (METs)	ICC (2,1)=0.85, 95% CI, 0.84-0.86					
	ActiGraph	Current study	EE (METs)	ICC (2,1)=0.7, 95% CI, 0.69-0.71					
		Current study (HR)	EE (METs)	ICC (2,1)=0.84, 95% CI, 0.83-0.85					
Stephens et al[26] 2016	Actical	Puyau	EE (METs) structured activities (II)	ICC (2,1)=0.38, 95% CI, 0.25-0.50			-/+		
	ActiGraph	Corder	EE (METs)	ICC (2,1)=0.31, 95% CI, 0.17-0.45					
		Freedson	EE (METs)	ICC (2,1)=0.36, 95% CI, 0.21-0.49					
		Puyau	EE (METs)	ICC (2,1)=0.25, 95% CI, 0.10-0.39					
	Actical	Current study	EE (METs)	ICC (2,1)=0.74, 95% CI, 0.73-0.75					
		Current study (HR)	EE (METs)	ICC (2,1)=0.78, 95% CI, 0.77-0.79					
	ActiGraph	Current study	EE (METs)	ICC (2,1)=0.71, 95% CI, 0.70-0.72					
		Current study	EE (METs)	ICC (2,1)=0.68, 95% CI, 0.66-0.70					
Tang et al[58] 2013			Duration (min) sitting/lying		Error=−2.6%		?	I	
			Duration (min) upright		Error=1.1%				
			Duration (min) standing		Error=−0.5%				
			Duration (min) stepping		Error=5.6%				
			Steps (n) walking		Error=3.8%				
Xing et al[18] 2021	ActiGraph	Freedman (VA)	EE sitting	MD=−0.58 METs, 95% CI, −0.64 to −0.44			?	V	
		Trost (VA)	EE sitting	MD=−0.24 Kcal/min, 95% CI, −0.46 to −0.02					
		Treuth (VA)	EE sitting	MD=−1.13 METs, 95% CI, −1.18 to −1.07]					
	activPAL		EE sitting	MD=−0.37 METs, 95% CI, −0.42 to −0.31]					
	ActiGraph	Freedman	EE standing	MD=−0.13 METs, 95% CI, −0.27 to 0.02					
		Trost	EE standing	MD=0.13 Kcal/min, 95% CI, −0.09 to 0.33					
		Treuth	EE standing	MD=−0.71 METs, 95% CI, −0.81 to −0.62					
	activPAL		EE standing	MD=−0.16 METs, 95% CI, −0.28 to −0.05					
	ActiGraph	Freedman	EE slow walking	MD=0.46 METs, 95% CI, 0.21 to 0.70					
		Trost	EE slow walking	MD=0.75 Kcal/min, 95% CI, 0.59 to 0.91					
		Treuth	EE slow walking	MD=0.02 METs, 95% CI, −0.14 to 0.19					
	activPAL		EE slow walking	MD=−0.44 METs, 95% CI, −0.60 to −0.27					
	ActiGraph	Freedman	EE moderate walking	MD=−0.17 METs, 95% CI, −0.50 to 0.17					
		Trost	EE moderate walking	MD=0.51 Kcal/min, 95% CI, 0.28 to 0.75					
		Treuth	EE moderate walking	MD=−0.29 METs, 95% CI, −0.47 to −0.10					
	activPAL		EE moderate walking	MD=0.39 METs, 95% CI, 0.09 to 0.68					
	ActiGraph	Freedman	EE fast walking	MD=1.84 METs, 95% CI, 1.26 to 2.43					
		Trost	EE fast walking	MD=2.17 Kcal/min, 95% CI, 1.33 to 3.01					
		Treuth	EE fast walking	MD=1.48 METs, 95% CI, 0.94 to 2.02					
	activPAL		EE fast walking	MD=3.54 METs, 95% CI, 2.60 to 4.48					

Criterion validity for cutoff point based methods									⊕⊕⊕
Baque et al[77] 2017		Baque	AC-VM (n/15 s) structured activities (SED)	AUC=1	Se=100%	Sp=100%	+	V	
			AC-VM (n/15 s) structured activities (MI)	AUC=0.98	Se=90.7%	Sp=92.6%			
			AC-VM (n/15 s) structured activities (VI)	AUC=0.99	Se=93.8%	Sp=96.3%			
		Clanchy	PA intensity levels (SED, LPA, MPA, VPA) (METs) structured activities (II)	κ=0.73 (SE=−0.08)					
		Evenson	PA intensity levels (SED, LPA, MPA, VPA) (METs) structured activities (II)	κ=0.77 (SE=−0.07)					
		Baque	PA intensity levels (SED, LPA, MPA, VPA) (METs) structured activities (II)	κ=0.92 (SE=−0.07)					

(*continued*)

Table 4B (*Continued*)

Study	Population		Device-Measured PA	Validity			Quality		
Author(s) YoP	Group specs	Cutoff points/ pred.eq.					Overall Rating	Risk of Bias	GRADE
Clanchy et al[70] 2011		Freedson	SED≤1.5 (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%	-/+	V	
		Puyau	SED (METs) structured activities (II)	AUC=0.90, 95% CI, 0.84-0.96	Se=90.3%	Sp=89.4%			
		Treuth	SED (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%			
		Evenson	SED (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%			
		Clanchy	SED (METs) structured activities (II)	AUC=0.92, 95% CI, 0.85-0.98	Se=86.7%	Sp=96.5%			
		Freedson	LPA=[1.5, 4] (METs) structured activities (II)	AUC=0.66, 95% CI, 0.58-0.74	Se=45.3%	Sp=85.7%			
		Puyau	LPA (METs) structured activities (II)	AUC=0.68, 95% CI, 0.59-0.77	Se=63.2%	Sp=73.1%			
		Treuth	LPA (METs) structured activities (II)	AUC=0.68, 95% CI, 0.60-0.75	Se=60.4%	Sp=76.2%			
		Evenson	LPA (METs) structured activities (II)	AUC=0.67, 95% CI, 0.58-0.75	Se=49.1%	Sp=84.1%			
		Clanchy	LPA (METs) structured activities (II)	AUC=0.63, 95% CI, 0.56-0.71	Se=35.9%	Sp=90.5%			
		Freedson	MVPA=≥4 (METs) structured activities (II)	AUC=0.75, 95% CI, 0.66-0.83	Se=81.8%	Sp=67.5%			
		Puyau	MVPA (METs) structured activities (II)	AUC=0.75, 95% CI, 0.68-0.83	Se=59.6%	Sp=91.3%			
		Treuth	MVPA (METs) structured activities (II)	AUC=0.73, 95% CI, 0.63-0.82	Se=66.7%	Sp=78.3%			
		Evenson	MVPA (METs) structured activities (II)	AUC=0.91, 95% CI, 0.84-0.97	Se=81.8%	Sp=100%			
		Clanchy	MVPA (METs) structured activities (II)	AUC=0.94, 95% CI, 0.88-0.98	Se=91.4%	Sp=86.2%			
Keawutan et al[71] 2016	GMFCS I	Keawutan	AC-VM (n) structured activities (SED)	AUC=0.79, 95% CI, 0.77-081	Se=74%	Sp=73%	+	D	
	GMFCS II		AC-VM (n) structured activities (SED)	AUC=0.78, 95% CI, 0.76-0.80	Se=72%	Sp=73%			
	GMFCS III		AC-VM (n) structured activities (SED)	AUC=0.81, 95% CI, 0.79-0.82	Se=74%	Sp=74%			
Keawutan et al[71] 2016	GMFCS I	Keawutan	Duration SED (min) structured activities (II)	MD=−13.3%, 95% LoA [−34.0 to 7.4]	Se=78.6%	Sp=84%	?		
		Butte	Duration SED (min) structured activities (II)	MD=−6.2%, 95% LoA [−22.8 to 10.4]	Se=84.1%	Sp=79.6%			
	GMFCS II	Keawutan	Duration SED (min) structured activities (II)	MD=−15.6%, 95% LoA [−33.5 to 2.3]	Se=78.2%	Sp=86.7%			
		Butte	Duration SED (min) structured activities (II)	MD=−10.4%, 95% LoA [−27.2 to 6.4]	Se=81.7%	Sp=83.2%			
	GMFCS III	Keawutan	Duration SED (min) structured activities (II)	MD=−1%, 95% LoA [−25.6 to 23.7]	Se=72.5%	Sp=76.2%			
		Butte	Duration SED (min) structured activities (II)	MD=−2.4%, 95% LoA [−26.8 to 21.9]	Se=71.8%	Sp=77%			
Oftedal et al[56] 2014	n=39	Oftedal (VM)	AC-VA (n) structured activities (SED)	AUC=0.77, 95% CI, 0.76-0.78	Se=71%	Sp=77%	+	D	
	n=18		AC-VM (n) structured activities (SED)	AUC=0.81, 95% CI, 0.80-0.82	Se=75%	Sp=76%			
Oftedal et al[56] 2014	n=23	Oftedal (VA)	Duration SED (min) structured activities (II)	MD=−10.5%, 95% LoA [−30.2 to 9.1]	Se=74%	Sp=80%	?	D	
	n=18		Duration SED (min) structured activities (II)	MD=−1.5%, 95% LoA [−20.0 to 16.8]	Se=79%	Sp=72%			
Ryan et al[73] 2014		Ryan	AC (n) structured activities (LI)	AUC=0.965, 95% CI, 84.6-99.8			?/+	D	
			AC (n) structured activities (MVI)	AUC=0.896, 95% CI, 77.4-96.6					
		VanHelst	SED (<2) (METs) structured activities (II)	κ=0.92, 95% CI, 0.82-1.00	Se=89.5%	Sp=100%			
		Ryan	SED (<2) (METs) structured activities (II)	κ=0.96, 95% CI, 0.89-1.00	Se=94.7%	Sp=100%			
		VanHelst	LPA ([2, 3]) (METs) structured activities (II)	κ=0.57, 95% CI, 0.38-0.77	Se=88.9%	Sp=79.6%			
		Ryan	LPA (METs) structured activities (II)	κ=0.71, 95% CI, 0.52-0.90	Se=83.3%	Sp=89.8%			
		Rowlands	MVPA (≥3) (METs) structured activities (II)	κ=0.69, 95% CI, 0.52-0.86	Se=70%	Sp=97.3%			
		VanHelst	MVPA (METs) structured activities (II)	κ=0.66, 95% CI, 0.48-0.84	Se=70%	Sp=94.6%			
		Ryan	MVPA (METs) structured activities (II)	κ=0.79, 95% CI, 0.64-0.94	Se=86.7%	Sp=91.9%			
Stephens et al[26] 2016	Actical	Stephens	AC-VM (n) structured activities (SED)	AUC=0.84, 95% CI, 0.78-0.90	Se=78%	Sp=88%	-/?/+	V	
			AC-VM (n) structured activities (MI)	AUC=0.82, 95% CI, 0.77-0.88	Se=72%	Sp=75%			
			AC-VM (n) structured activities (VI)	AUC=0.98, 95% CI, 0.96-1.0	Se=100%	Sp=94%			
	ActiGraph	Stephens	AC-VA (n) structured activities (SED)	AUC=0.82, 95% CI, 0.74-0.90	Se=75%	Sp=91%			
			AC-VA (n) structured activities (MI)	AUC=0.78, 95% CI, 0.71-0.86	Se=86%	Sp=63%			
			AC-VA (n) structured activities (VI)	AUC=0.78, 95% CI, 0.52-1.0	Se=83%	Sp=79%			
	Actical	Evenson	SED (n) structured activities (RI)		Se=72%	Sp=92%			
			MPA (n) structured activities (RI)		Se=49%	Sp=94%			
			VPA (n) structured activities (RI)		Se=100%	Sp=92%			
	ActiGraph	Evenson	SED (n) structured activities (RI)		Se=75%	Sp=90%			
			MPA (n) structured activities (RI)		Se=41%	Sp=90%			
			VPA (n) structured activities (RI)		Se=50%	Sp=95%			
Stephens et al[26] 2016	Actical	Stephens	AC-VM (n) structured activities (SED)	AUC=0.96, 95% CI, 0.93-0.98	Se=82%	Sp=97%	-/?/+	V	
			AC-VM (n) structured activities (MI)	AUC=0.89, 95% CI, 0.83-0.94	Se=82%	Sp=81%			
			AC-VM (n) structured activities (VI)	AUC=0.91, 95% CI, 0.87-0.95	Se=100%	Sp=90%			
	ActiGraph	Stephens	AC-VA (n) structured activities (SED)	AUC=0.9, 95% CI, 0.86-0.95	Se=78%	Sp=91%			
			AC-VA (n) structured activities (MI)	AUC=0.91, 95% CI, 0.85-0.97	Se=81%	Sp=94%			
			AC-VA (n) structured activities (VI)	AUC=0.92, 95% CI, 0.88-0.96	Se=100%	Sp=92%			
	Actical	Evenson	SED (n) structured activities (RI)		Se=80%	Sp=97%			
			MPA (n) structured activities (RI)		Se=47%	Sp=96%			
			VPA (n) structured activities (RI)		Se=0%	Sp=96%			
	ActiGraph	Evenson	SED (n) structured activities (RI)		Se=75%	Sp=91%			
			MPA (n) structured activities (RI)		Se=81%	Sp=90%			
			VPA (n) structured activities (RI)		Se=0%	Sp=92%			
Trost et al[24] 2016		Evenson	SED (<1.5) (METs) structured activities (II)	AUC=0.93, 95% CI, 0.91-0.96	Se=98.9%	Sp=87.6%	-/+	V	
		Clanchy	SED (METs) structured activities (II)	AUC=0.93, 95% CI, 0.91-0.96	Se=98.9%	Sp=87.6%			
		Trost (VA)	SED (METs) structured activities (II)	AUC=0.97, 95% CI, 0.95-0.99	Se=97.9%	Sp=96.1%			
		Trost (VM)	SED (METs) structured activities (II)	AUC=0.96, 95% CI, 0.94-0.99	Se=96.9%	Sp=96.1%			
		Evenson	LPA ([1.5, 3]) (METs) structured activities (II)	AUC=0.68, 95% CI, 0.63-0.73	Se=61.5%	Sp=74.6%			
		Clanchy	LPA (METs) structured activities (II)	AUC=0.68, 95% CI, 0.62-0.73	Se=58.1%	Sp=77.1%			
		Trost (VA)	LPA (METs) structured activities (II)	AUC=0.82, 95% CI, 0.77-0.86	Se=77.8%	Sp=86.4%			
		Trost (VM)	LPA (METs) structured activities (II)	AUC=0.8, 95% CI, 0.76-0.85	Se=72.7%	Sp=87.8%			
		Evenson	MVPA (≥3) (METs) structured activities (II)	AUC=0.75, 95% CI, 0.71-0.80	Se=57.1%	Sp=93.4%			
		Clanchy	MVPA (METs) structured activities (II)	AUC=0.76, 95% CI, 0.72-0.81	Se=61.4%	Sp=91.6%			
		Trost (VA)	MVPA (METs) structured activities (II)	AUC=0.86, 95% CI, 0.82-0.89	Se=78.6%	Sp=92.5%			
		Trost (VM)	MVPA (METs) structured activities (II)	AUC=0.86, 95% CI, 0.82-0.89	Se=81.4%	Sp=89.7%			

(*continued*)

Table 4B (*Continued*)

Study	Population		Device-Measured PA	Validity			Quality		
Author(s) YoP	Group specs	Cutoff points/ pred.eq.					Overall Rating	Risk of Bias	GRADE
Xing et al[18]		Puyau (VA)	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.84	κ=0.458		+	V	
2021		Evenson	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.888	κ=0.585				
		Romanzini (VA)	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.886	κ=0.56				
		Romanzini (VM)	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.886	κ=0.675				
		Clanchy	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.935	κ=0.721				
		Baque	PA intensity levels (SED, LPA, MVPA) (METs) structured activities (II)	rs=0.896	κ=0.773				
Convergent validity									⊕⊕⊕
Kuo et al[55] 2009	AMP		Distance continuous walking	MD=−4.8 m, 95% LoA (-20.1 to 10.5)			?	D	
	Minimod		Distance continuous walking	MD=−0.4 m, 95% LoA (-4.7 to 4.0)					
	AMP		Distance intermittent walking	MD=−3.6 m, 95% LoA (-19.2 to 12.0)					
	Minimod		Distance intermittent walking	MD=−2.3 m, 95% LoA (-27.9 to 23.3)					
	AMP		Distance downstair climbing	MD=−1.3 m, SD -2.5					
	Minimod		Distance downstair climbing	MD=8.9 m, SD -2.5					
	AMP		Distance upstair climbing	MD=−2 m, SD -2.5					
	Minimod		Distance upstair climbing	MD=3.3 m, SD −2.2					
Lawal et al[21] 2020	ActiGraph GT3X		Steps (n) walking	ICC=0.29, 95% CI, −0.42 to 0.78			-	D	
	LFE-ActiGraph GT3X		Steps (n) walking	ICC=0.52, 95% CI, −0.16 to 0.87					
Mackey et al[66] 2009			Duration (min) lying	Se=100%	Sp=100%		?	D	
			Duration (min) sitting	Se=100%	Sp=100%				
			Duration (min) standing	Se=100%	Sp=97%				
	n=12 for stair climbing		Duration (min) walking (overground and stairs)	Se=78.5%	Sp=100%				
Sala et al[74] 2019	Hip (n=38)		Steps walking	MAE=7 steps, range=−52 to 6	r=0.991		-/+	D	
			Distance walking	MAE=0.07 miles, range=0.01 to 0.16					
	GMFCS I+II (n=27)		Steps walking	MAE=6 steps, range=−20 to 6	r=0.998				
			Distance walking	MAE=0.07 miles, range=0.01 to 0.16					
	GMFCS III (n=11)		Steps walking	MAE=12 steps, range=−52 to 1	r=0.981				
			Distance walking	MAE=0.07 miles, range=0.02 to 0.14					
	Wrist (n=38)		Steps walking	MAE=88 steps, range=−484 to 35	r=−0.033				
			Distance walking	MAE=0.06 miles, range=−0.13 to 0.16					
	GMFCS I+II (n=27)		Steps walking	MAE=27 steps, range=−177 to 23	r=0.837				
			Distance walking	MAE=0.04 miles, range=−0.02 to 0.16					
	GMFCS III (n=11)		Steps walking	MAE=238 steps, range=−484 to 35	r=−0.242				
			Distance walking	MAE=0.08 miles, range=−0.13 to 0.11					

Abbreviations: "?", indeterminate; "-", insufficient; "+", sufficient; "⊕⊕⊕", moderate; AC, activity counts; AUC, area under the curve; D, doubtful; EE, energy expenditure; GRADE, Grading of Recommendations Assessment, Development and Evaluation; HR, heart rate; HRAR, heart rate above rest; HRAS, heart rate above sleep; HS, hemiparetic side; I, inadequate; II, increased intensity; LNI, least neurological impaired; LoA, levels of agreement; LPA, low physical activity; MAE, mean absolute error; MD, mean difference; MET, metabolic equivalent of task; MI, moderate intensity; MPA, moderate physical activity; MVI, moderate-to-vigorous; n, number; NA, not applicable; NHS, nonhemiparetic side; NI, neurological impaired; pred, prediction; r, Pearson correlation coefficient; R^2, coefficient of determination; RI, random intensity; rs, Spearman rank correlation coefficient; Se, sensitivity; SED, sedentary; Sp, specificity; V, very good; VA, vertical axis; VI, vigorous intensity; VM, vector magnitude; VPA, vigorous physical activity; YoP, year of publication; κ, Cohen's kappa.

GMFCS level, age difference, physical behavior type, and placement were non-significant moderators (see appendix 2).

For criterion validity based on cutoff points, moderator analysis showed that physical behavior class and age were statistically significant in cutoff point based methods with a high correlation with r_+=0.76 (95% CI, 0.62-0.89) for sedentary behavior, r_+=0.47 (95% CI, 0.29-0.65) for light PA, and r_+=0.61 (95% CI, 0.47-0.76) for MVPA. Children ≥13 years old showed a very high correlation with r_+=0.91 (95% CI, 0.79-1.00), compared with children <13 years of age with a moderate correlation of r_+=0.61 (95% CI, 0.54-0.68). PA dimensions, percentage GMFCS level, age <6 years or >6 years, and cutoff points based on a specific population or general were nonsignificant moderators (see appendix 2).

Grading the pooled evidence

The quality of evidence of all psychometric properties was moderate.

Publication bias

Funnel plots, *p*-curve analysis, year of publication as analysis as a moderator, and outlier identification reveal no substantial evidence of publication bias. Notably, recent studies on cutoff point based criterion validity exhibit a larger effect size (slope, 0.03; 95% CI, 0.01-0.07; $F[1,22]$=4.56; $P<.04$). Simulation detected outliers for interdevice reliability (observed simulation outliers, 0.50). However, there was no significant difference between the observed and simulated data (P=.14), indicating no anticipated impact on the overall effect size. Additionally, corrected effect sizes were computed ($ICC_{test\text{-}retest\ reliability}$=0.76; 95% CI, 0.61-0.91; $ICC_{interdevice\ reliability}$=analysis not possible; $r_{construct\ validity}$=0.34; 95% CI, −0.09 to 0.78; $r_{criterion\ validity}$=0.49; 95% CI, 0.27-0.72; $r_{criterion\ validity\ cutoff\ point\ based\ methods}$=0.59; 95% CI, 0.18-0.99), considering the assumption that confirmatory findings are 5 times more likely to be published than nonconfirmatory findings.
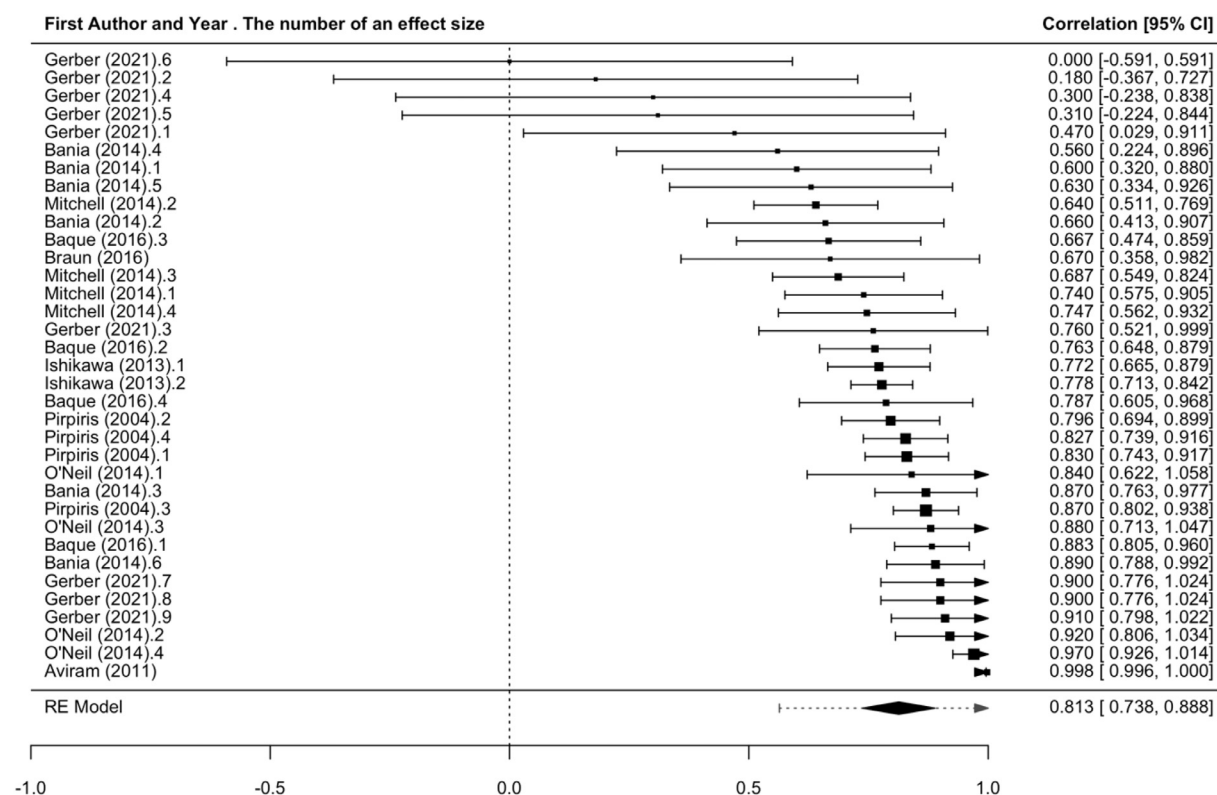
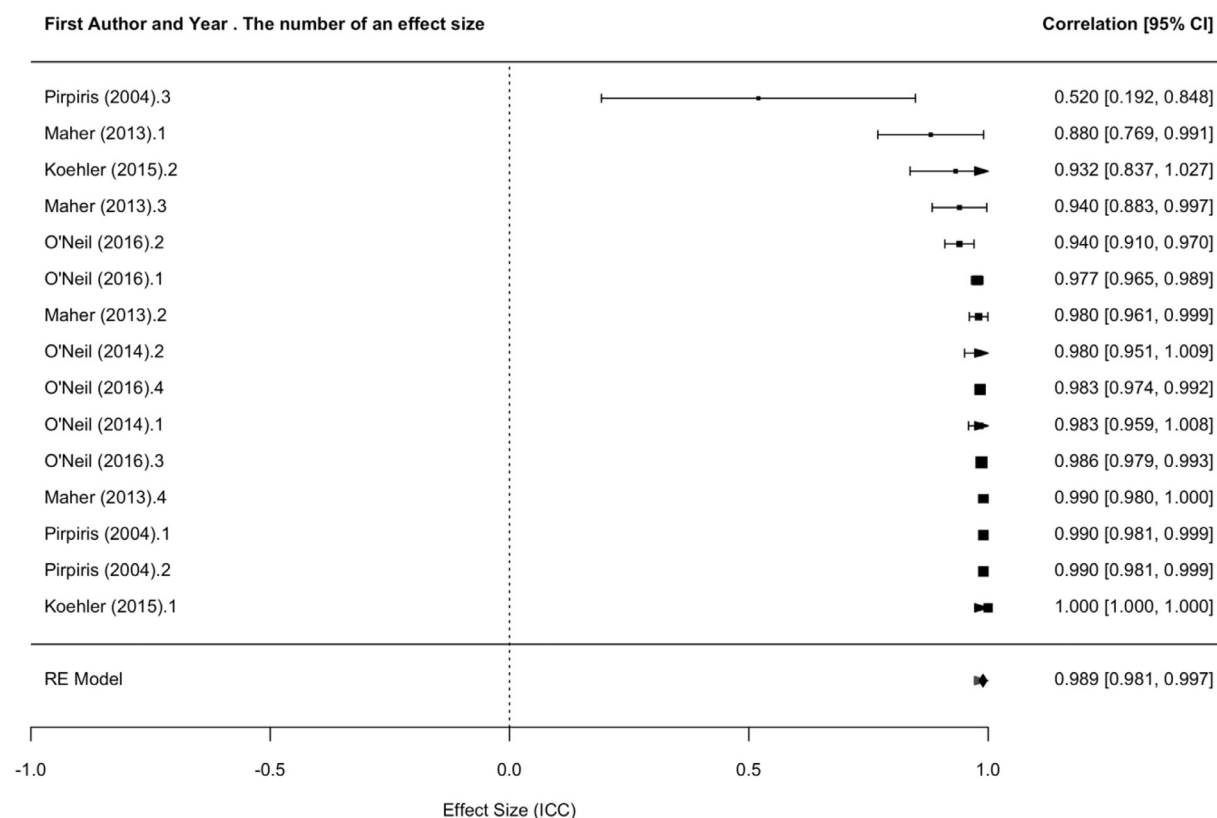**Fig 2** Forest plot for test-retest reliability. RE, Random Effects Model.

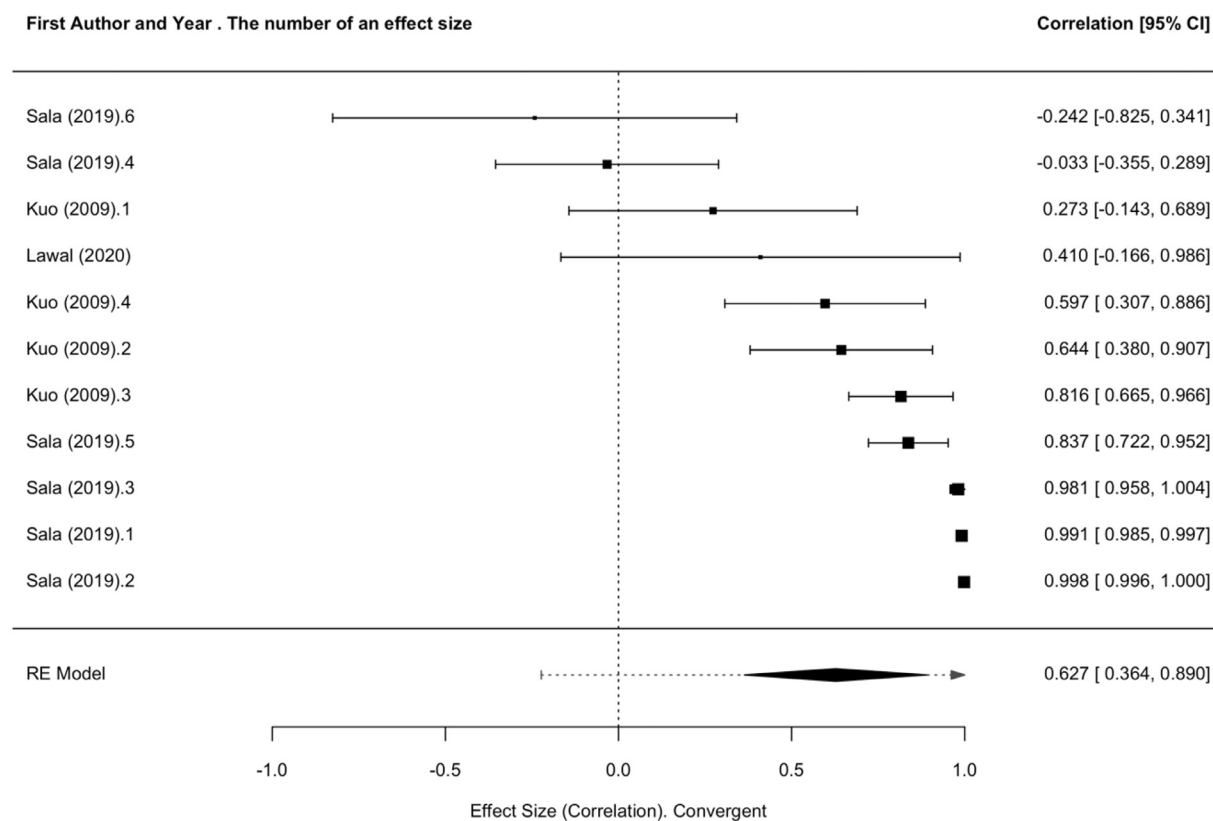**Fig 3** Forest plot for inter-device reliability. RE, Random Effects Model.

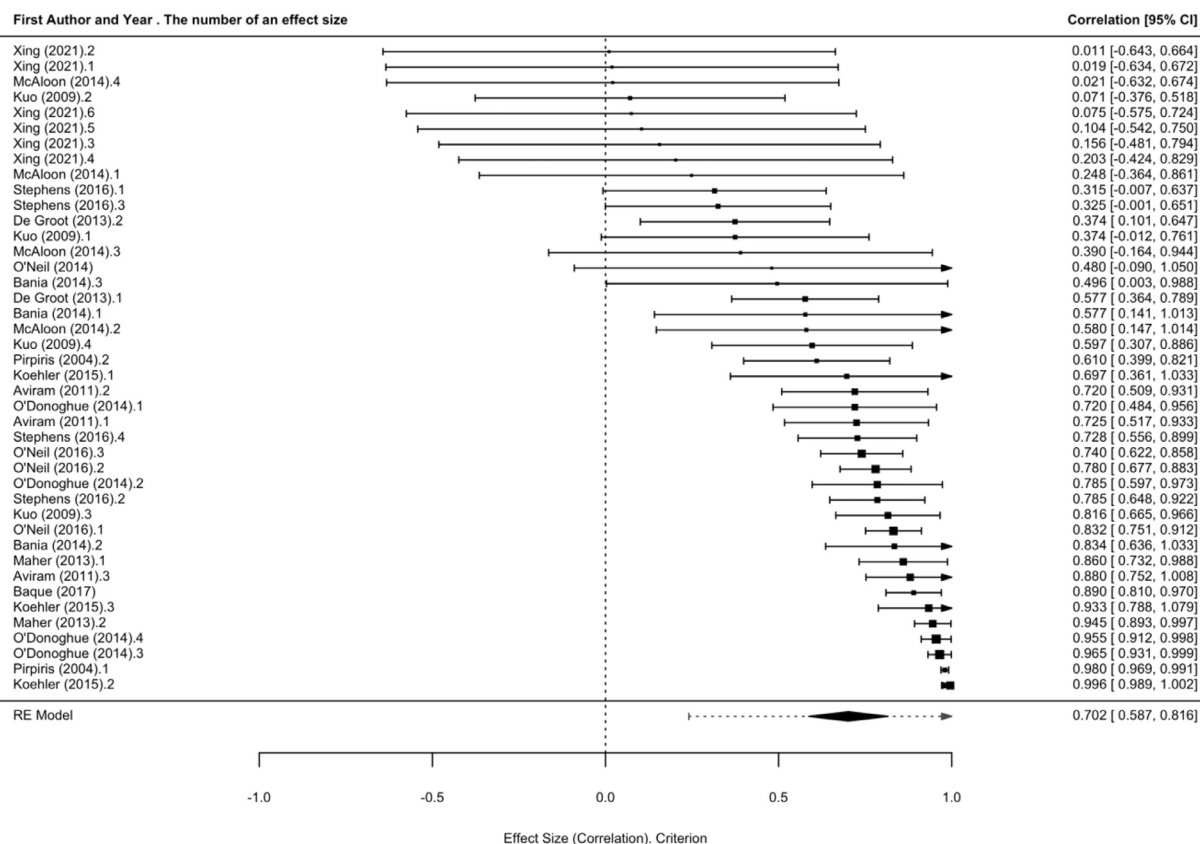**Fig 4**   Forest plot for construct validity. RE, Random Effects Model.

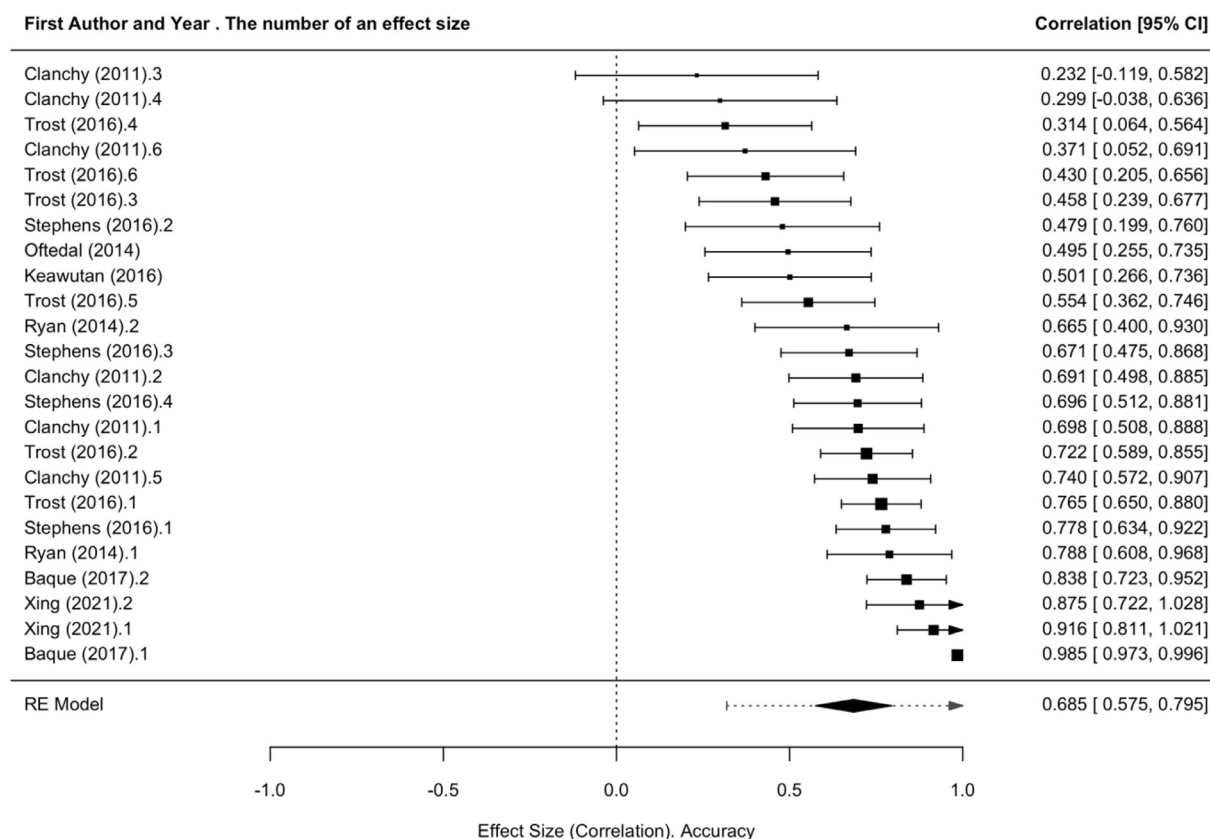**Fig 5**   Forest plot for criterion validity. RE, Random Effects Model.

First Author and Year . The number of an effect size		Correlation [95% CI]
Clanchy (2011).3		0.232 [-0.119, 0.582]
Clanchy (2011).4		0.299 [-0.038, 0.636]
Trost (2016).4		0.314 [0.064, 0.564]
Clanchy (2011).6		0.371 [0.052, 0.691]
Trost (2016).6		0.430 [0.205, 0.656]
Trost (2016).3		0.458 [0.239, 0.677]
Stephens (2016).2		0.479 [0.199, 0.760]
Oftedal (2014)		0.495 [0.255, 0.735]
Keawutan (2016)		0.501 [0.266, 0.736]
Trost (2016).5		0.554 [0.362, 0.746]
Ryan (2014).2		0.665 [0.400, 0.930]
Stephens (2016).3		0.671 [0.475, 0.868]
Clanchy (2011).2		0.691 [0.498, 0.885]
Stephens (2016).4		0.696 [0.512, 0.881]
Clanchy (2011).1		0.698 [0.508, 0.888]
Trost (2016).2		0.722 [0.589, 0.855]
Clanchy (2011).5		0.740 [0.572, 0.907]
Trost (2016).1		0.765 [0.650, 0.880]
Stephens (2016).1		0.778 [0.634, 0.922]
Ryan (2014).1		0.788 [0.608, 0.968]
Baque (2017).2		0.838 [0.723, 0.952]
Xing (2021).2		0.875 [0.722, 1.028]
Xing (2021).1		0.916 [0.811, 1.021]
Baque (2017).1		0.985 [0.973, 0.996]
RE Model		0.685 [0.575, 0.795]

Effect Size (Correlation). Accuracy

Fig 6   Forest plot for criterion validity based on cutoff points. RE, Random Effects Model.

Discussion

Our systematic review represents a multilevel meta-analysis appraising, comparing, summarizing, and generalizing levels of reliability and validity of instruments for wearables measuring PA in ambulatory children with gait abnormalities due to neuromuscular conditions. The meta-analysis of 26 studies showed high to very high reliability and moderate to high validity of wearables measuring PA. Despite the heterogeneous quality of the studies (risk of bias), the overall quality of evidence (Grading of Recommendations Assessment, Development and Evaluation) of reliability studies, and the quality of validity studies, was moderate. Significant moderators for test-retest reliability were PA setting (laboratory vs free-living setting), PA dimensions (frequency, intensity, duration), and specified physical behavior type (walking, lying and sitting, standing). For inter-device reliability, PA dimensions (frequency, intensity, duration), and age (<13y, >13y) were statistically significant moderators. Specified physical behavior type (running, lying and sitting, standing, walking) was a significant moderator for criterion validity, whereas criterion validity based on cutoff points was affected by physical behavior class (sedentary, low PA, MVPA), as well as age (<13y, >13y). Levels of measurement error could not be compared or summarized because of heterogeneity in outcome measures.

Practical implications

Most studies were conducted in laboratory settings. For test-retest reliability, we differentiated between laboratory and free-living settings. Studies conducted in laboratories showed higher test-retest reliability than those conducted in free-living settings. Similar findings from other reviews on wearables in children without disabilities highlight a lack of free-living validations with child-specific movements in their protocols.[13,80] Future research should prioritize validating wearables in free-living settings, incorporating child-specific movements to improve clinical relevance. This will enable clinicians to analyze better and interpret results in daily practice.

Interestingly, correlation for free-living setting for reliability is still high, indicating that wearables measuring PA in children with gait abnormalities shows consistent scores over time. We understand that the measurement error in free-living settings is higher than in laboratory settings because children are more prone to factors of physical environment, social environment and personal factors, influencing their level of PA.[81] To estimate the true change of a patient when assessed in repeated measurements, a clinician also needs to know information about the measurement error. Unfortunately, we were not able to comprehensively summarize the measurement error reported in the studies found, as there was a broad variability in calculating and reporting measurement error.[28] Consequently, the interpretation of the individual study findings regarding measurement error must be situated within the context of each individual study and cannot be generalized for children with abnormal walking patterns.

In our study, we did not find a moderating effect of the percentage of GMFCS level I on the psychometric properties,

and therefore, we cannot draw conclusions about the accuracy of wearable devices measuring PA related to increasing gait abnormalities, based on the pooled results of our study. We defined the moderator as percentage of GMFCS level I, as many studies did not report the contrast between levels I, II, and III. In individual studies, there is evidence that SDs in step counts are higher in children with GMFCS III compared with GMFCS I because of higher energy cost and variability in increasing abnormal gait.[17,64] A reason for higher misclassification in children with a GMFCS level III may be that they use functional walking aids such as crutches or walkers, which can influence the accuracy of a device.[74]

The moderator of age (>13y of age or younger) was only significant in cutoff point based validity: older children reached higher correlations than younger children. This finding has consequences for clinical assessment and treatment as accuracy drops to a moderate level in children younger than 13 years, which induces more uncertainty about the actual level of PA. Given that children younger than 13 years tend to be more physically active, both age and higher activity levels serve as moderators that adversely affect cutoff point based criterion validity. More research in younger children is needed for a better understanding of their levels of PA.[5] This is especially true since the age-related decrease in PA and increased sedentary time seems to become apparent at 6-7 years of age.[4] We are aware that walking may be delayed in children with disabilities, and therefore, the age of 2 is the first age of inclusion with possible PA measurements of gait pattern.

For test-retest reliability, steps and PA counts (frequency) showed very high correlations and energy expenditure rate (intensity) and duration were correlated highly. Among physical behavior types, duration has a negative impact on test-retest reliability compared with frequency and intensity. Duration differs statistically from frequency and intensity, implying that clinicians might have the most optimal assessment when measuring frequency (PA counts) or intensity (energy expenditure rate). The high to very high values of test-retest reliability imply that the clinician should use devices to measure PA in children with abnormal gait as robust tools when repeating an assessment.

When comparing the same devices worn at the same time on different locations, for example, right and left hip, correlations were very high for intensity (energy expenditure rate). This trend was consistent across measures of intensity (energy expenditure rate), frequency (steps, PA counts), as well as duration (seconds, minutes of PA). Those findings show that the devices are suitable for repeated measures and robust in construct.

The specified physical behavior type (running, lying and sitting, standing, walking) was a significant moderator. When interested in step counts and supporting children to increase their daily walking time, clinicians can use devices that are validated well for step counts. If clinicians are more interested in intensity levels of PA, they have to get familiar with devices that can distinguish between physical behavior classes as sedentary, low PA, or MVPA. Cutoff point based methods try to shed light into the best algorithms to provide valid assessments. Physical behavior class was a statistically significant moderator for cutoff point based methods with a low correlation for sedentary activities. Sedentary behavior, as measured by lying or sitting, shows statistical differences from walking and running. This indicates that cutoff point based criterion validity is higher during activities as walking or running, indicating that cutoff points may positively impact the validity for sedentary behavior. Our finding is contrary to Lynch et al,[82] who assessed the accuracy of accelerometers for measuring PA and levels of sedentary behavior in children in 11 studies: accuracy appeared to be highest when detecting sedentary activities and least with low PA. Based on our findings we believe that cutoff point based values can be used best to distinguish between sedentary and active behavior in children with abnormal gait.

To the best of our knowledge, this systematic review represents the first multilevel meta-analysis appraising, comparing, summarizing, and generalizing the levels of reliability and validity of instruments for wearable devices measuring PA in children with gait abnormalities due to neuromuscular conditions. A notable strength of our research is the predetermined and Prospective Register for Systematic Reviews—registered methodology, which aligns with COSMIN guidelines. Through this meta-analysis, we overcame the influence of small samples (n<100) in determining the evidence of effect size, a common limitation in systematic literature reviews on the measurement properties of wearable PA instruments. In addition, our meta-analysis enabled us to analyze the effects of publication bias, for which we found no substantial evidence. We suppose that the analysis techniques used will be future solutions for pooling and analyzing data. When grading the level of evidence with COSMIN guidelines, publication bias is not taken into account because it is difficult to assess studies on measurement properties because of a lack of registries for these types of studies.[28]

As there is a literature gap about wearables measuring PA and surveys of parents in children under 5 years of age,[5] this review makes an important contribution given its inclusion of studies of children with a mean age of 2.3 years and could not find a moderating effect of children younger than 6 years on wearable devices measuring PA.

Study limitations

A limitation is the generalizability of the results to all children with abnormal gait. Our research question was narrow, targeting children with gait abnormality due to neuromuscular conditions and therefore excluding children who use a wheelchair and children with conditions other than neuromuscular etiology. Most studies (71%) focused on children with CP, similar to a previous review about activity instruments in children with physical disabilities, which also found 64% of studies focusing on CP.[10] This aligns with the higher global prevalence of CP[83] compared with, for example, muscular dystrophies[84] or neural tube defects,[85] and therefore reflects a realistic proportion in pediatric rehabilitation.

In our study, we analyzed reliability and validity and did not assess responsiveness ("longitudinal validity"), which is important as it refers to the ability of a measurement instrument to detect change over time in the construct to be measured.[28,86] Responsiveness was included in the search strategy and screening, but we did not find studies assessing responsiveness in children with abnormal gait. We assume that research is still concentrating on reliability and validity

studies in this group of children before focusing on responsiveness.

For construct validity, results should be interpreted cautiously because of the low number of studies (n=3) included. Measuring PA in children with gait abnormalities with other than criterion standard reference methods is prone to errors. Therefore, better instruments should be developed to catch the construct they are supposed to assess. Nonetheless, Sala et al[74] provide the clinician with important knowledge: they assessed children with CP with a wrist-based and hip-based activity monitor (AM) to investigate the number of steps and distance walked. The number of steps detected by a wearable device showed a very low correlation with steps registered with a tally counter ($r=-0.03$). This correlation is based on the placement of the device (wrist), while 11 of 13 children walked with crutches and posterior walkers, which affected the natural sway of the arm.[74] Although wrist-worn wearables may be better accepted among children, they seem to be unable to detect steps when using walking aids. With this important finding, we urge clinicians to consider the intended outcome and target population of interest when choosing suitable wearables. Unfortunately, based on this systematic review, the best placement of a device remains unclear, as moderator analyses were statistically nonsignificant.

The WHO recently stated that there is insufficient device-measured data assessing PA in children with a disability,[5] which may result in insufficient action from policymakers because there is no factual problem.[3] Wearables used by clinicians to measure PA are essential in diagnostics and individually tailored treatment of functional (dis)ability in children with neuromuscular conditions with consequent abnormal gait. With more evidence that children with a developmental disability show lower levels of PA than typically developing peers, we can elucidate the problem and apply for structural support from policymakers.

Conclusions

There is high to very high reliability for wearable devices measuring instruments in children with abnormal gait, and moderate to high validity in children with primarily neurological conditions. The use of wearables in clinical practice can support a clinician's clinical reasoning process and help assess tailored PA interventions. Clinicians should be aware that several moderators, for example, setting of measurement (laboratory, free-living), dimensions (frequency, intensity, duration), specified physical behavior type (walking, lying and sitting, running, standing), physical behavior class (sedentary, low PA, MVPA), or age can influence an assessment. As we found barely any studies that were conducted in free-living settings, we encourage researchers to conduct psychometric studies of devices measuring PA in more functional, free-living environments and include child-specific variability of movements. We advise researchers to report their methodology clearly and provide access to raw data, which can be used for pooling.

Suppliers

a. Covidence systematic review software, Veritas Health Innovation.
b. Rayyan software for systematic reviews, Rayyan.

Corresponding author

Manon A.T. Bloemen, PhD, HU University of Applied Sciences Utrecht, Heidelberglaan 7, 3584 CJ Utrecht, The Netherlands. *E-mail address:* manon.bloemen@hu.nl.

Disclosure

Acknowledgments

Appendix 1 Meta-analysis Results

Tables 1−5

Table 1 Meta-analysis results for test-retest reliability.

Model (k)	ICC (SE)	95% CI LB	95% CI UB	t (df)	P	Q	P	I²
Random-effects (35)*	0.813 (0.037)	0.738	0.888	22.010 (34)	<.001	320.262	<.001	88.574
Random-effects (35)†	0.807 (0.021)	0.765	0.849	38.902 (34)	<.001			0.001
Random-effects (35)‡	0.813 (0.037)	0.738	0.888	22.010 (34)	<.001			88.573

Abbreviations: LB, lower bound; UB, upper bound.
* The multilevel model that accounted for the within and between-study variance.
† A model that accounted for the within-study variance only (level 2 only).
‡ A model that accounted for the between-study variance (level 3 only).

Table 2 Meta-analysis results for interdevice reliability.

Model (k)	ICC (SE)	95% CI LB	95% CI UB	t (df)	P	i	P	I^2
Random-effects (15)*	0.989 (0.004)	0.981	0.997	261.809 (14)	<.001	96.876	<.001	71.014
Random-effects (13)†	0.985 (0.003)	0.978	0.992	307.124 (14)	<.001			0.001
Random-effects (13)‡	0.989 (0.004)	0.981	0.997	261.809 (14)	<.001			71.013

Abbreviations: LB, lower bound; UB, upper bound.
* The multilevel model that accounted for the within and between-study variance.
† A model that accounted for the within-study variance only (level 2 only).
‡ A model that accounted for the between-study variance (level 3 only).

Table 3 Meta-analysis results for construct validity.

Model (k)	r (SE)	95% CI LB	95% CI UB	t (df)	P	Q	P	I^2
Random-effects (11)*	0.627 (0.118)	0.364	0.890	5.311 (10)	<.001	106.882	<.001	99.967
Random-effects (11)†	0.627 (0.118)	0.364	0.890	5.331 (10)	<.001			99.966
Random-effects (11)‡	0.788 (0.144)	0.467	1.110	5.463 (10)	<.001			0.001

Abbreviations: LB, lower bound; UB, upper bound.
* The multilevel model that accounted for the within and between-study variance.
† A model that accounted for the within-study variance only (level 2 only).
‡ A model that accounted for the between-study variance (level 3 only)

Table 4 Meta-analysis results for criterion validity.

Model (k)	r (SE)	95% CI LB	95% CI UB	t (df)	P	Q	P	I^2
Random-effects (42)*	0.702 (0.057)	0.587	0.816	12.373 (41)	<.001	308.638	<.001	98.703
Random-effects (42)†	0.689 (0.038)	0.612	0.766	18.084 (41)	<.001			19.443
Random-effects (42)‡	0.732 (0.057)	0.617	0.848	12.805 (41)	<.001			79.261

Abbreviations: LB, lower bound; UB, upper bound.
* The multilevel model that accounted for the within and between-study variance.
† A model that accounted for the within-study variance only (level 2 only).
‡ A model that accounted for the between-study variance (level 3 only).

Table 5 Meta-analysis results for criterion validity of cutoff point based methods.

Model (k)	r (SE)	95% CI LB	95% CI UB	t (df)	P	Q	P	I^2
Random-effects (24)*	0.685 (0.053)	0.575	0.795	12.905 (23)	<.001	266.948	<.001	87.021
Random-effects (24)†	0.657 (0.041)	0.573	0.741	16.215 (23)	<.001			44.629
Random-effects (24)‡	0.717 (0.055)	0.603	0.831	12.973 (23)	<.001			42.392

Abbreviations: LB, lower bound; UB, upper bound.
* The multilevel model that accounted for the within and between-study variance.
† A model that accounted for the within-study variance only (level 2 only).
‡ A model that accounted for the between-study variance (level 3 only).

Appendix 2: Moderator Analysis

Tables 1—5

Table 1 Moderator analysis test-retest reliability.

Test-Retest Reliability

Moderator (k)	Summary Effect and 95% CI					Test of Moderation			
	ICC (SE)/ Slope (SE)	95% CI LB	95% CI UB	t	P	F (df1, df2)	P	R^2	Total I^2/Level 2 I^2/Level 3 I^2
PA domain (33)				19.739	<.001	5.630 (1, 33)	.024	0.302	82.39%/0.01%/82.38%
Laboratory	0.915 (0.050)	0.814	1.016	18.359	<.001				
Community	0.767 (0.037)	0.692	0.843	20.690	<.001				
PA dimension (35)				5.922 (2, 32)	.007	0.388	86.72%/0.01%/86.71%
Frequency*	0.998 (0.108)	0.779	1.217	9.284	<.001				
Intensity[†]	0.876 (0.045)	0.783	0.968	19.268	<.001				
Duration*,[†]	0.713 (0.046)	0.620	0.806	15.632	<.001				
PB type specified (22)						3.934 (2, 19)	.004	0.270	89.44%/6.83%/82.61%
Lying and sitting[‡]	0.727 (0.079)	0.562	0.893	9.217	<.001				
Standing[§]	0.704 (0.081)	0.533	0.874	8.645	<.001				
Walking[‡,§]	0.872 (0.053)	0.760	0.984	16.318	<.001				
PB type (41)						0.492 (1, 27)	.489	0.053	89.87%/0.01%/89.86%
Single	0.852 (0.049)	0.751	0.952	17.412	<.001				
Multi	0.786 (0.080)	0.622	0.950	9.858	<.001				
% GMFCS level I (35)	−0.001 (0.003)	−0.007	0.004	−0.526	.602	0.277 (1, 33)	.602	0.039	89.57%/0.01%/89.56%
Placement (33)						0.072 (2, 32)	.930	0.016	90.61%/0.01%/90.60%
Leg	0.796 (0.072)	0.649	0.943	11.016	<.001				
Trunk	0.807 (0.062)	0.680	0.934	12.951	<.001				
Multi	0.836 (0.081)	0.672	1.000	10.382	<.001				
Age (35)						0.599 (1, 31)	.445	0.059	88.83%/0.01%/88.82%
<13 y	0.811 (0.038)	0.734	0.888	21.410	<.001				
>13 y	0.880 (0.081)	0.714	1.047	10.810	<.001				
Time interval test-retest (33)						0.410 (1, 33)	.527	0.050	89.55%/0.01%/89.54%
<2 wk	0.801 (0.042)	0.714	0.887	18.742	<.001				
>2 wk	0.864 (0.090)	0.681	1.048	9.564	<.001				

Abbreviations: LB, lower bound; PB, physical behavior; R^2, coefficient of determination; UB, upper bound.
[*,†,‡,§] Differences between effect sizes are statistically significant.

Table 2 Moderator analyses and publication bias for inter reliability.

Inter-Device Reliability

Moderator (k)	Summary Effect and 95% CI					Test of Moderation			
	ICC (SE)/slope (SE)	95% CI LB	95% CI UB	t	P	F (df1, df2)	P	R^2	Total I^2/Level 2 I^2/Level 3 I^2
PA domain (15)						0.010 (1, 13)	.921	0.003	76.97%/0.01%/76.96%
Laboratory	0.989 (0.005)	0.978	0.999	203.151	<.001				
Community	0.990 (0.009)	0.970	1.010	203.151	<.001				
PA dimensions (15)						35.070 (2, 12)	.001	—	—/—/—
Frequency*	0.983 (0.002)	0.979	0.988	456.832	<.001				
Intensity*,[†]	1.000 (0.000)	1.000	1.000	14998.97	<.001				
Duration[†]	0.990 (0.003)	0.983	0.997	306.675	<.001				
PB type specified (8)						3.003 (2, 5)	.139	0.435	93.51%/0.01%/93.50%
Lying and sitting	1.000 (0.025)	0.942	1.071	40.256	<.001				
Walking	0.957 (0.022)	0.901	1.013	44.132	<.001				
Running	1.000 (0.023)	0.945	1.063	43.948	<.001				

Table 2 (*Continued*)

	Inter-Device Reliability								
	Summary Effect and 95% CI					Test of Moderation			
Moderator (k)	ICC (SE)/slope (SE)	95% CI LB	95% CI UB	t	P	F (df1, df2)	P	R^2	Total I^2/Level 2 I^2/Level 3 I^2
Placement (14)						2.207 (2, 11)	.156	0.197	93.73%/0.01%/93.72%
Leg	0.989 (0.009)	0.969	1.010	108.275	<.001				
Trunk	0.997 (0.009)	0.977	1.017	108.032	<.001				
Arm	0.973 (0.013)	0.945	1.001	76.573	<.001				
Age (15)						15.709 (1, 13)	.002	0.828	24.28%/0.01%/24.27%
<13 y	0.986 (0.002)	0.981	0.991	415.326	<.001				
≥13 y	1.000 (0.003)	0.994	1.006	372.703	<.001				
Placement body side (15)						0.018 (1, 13)	.897	0.002	73.63%/8.42%/65.21%
Same	0.988 (0.006)	0.976	1.000	173.089	<.001				
Opposite	0.989 (0.004)	0.980	0.998	228.930	<.001				

Abbreviations: LB, lower bound; PB, physical behavior; R^2, coefficient of determination; UB, upper bound.

[*],[†] Differences between effect sizes are statistically significant.

Table 3 Moderator analysis for construct validity.

	Construct Validity								
	Summary Effect and 95% CI					Test of Moderation			
Moderator (k)	r (SE)/Slope (SE)	95% CI LB	95% CI UB	t	P	F (df1, df2)	P	R^2	Total I^2/Level 2 I^2/Level 3 I^2
% GMFCS level 1 (11)	0.004 (0.004)	−0.006	0.014	0.914	.385	0.835 (1, 9)	.385	0.085	99.97%/99.96%/0.01%
PA type (11)						0.000 (1, 9)	.997	0.001	99.97%/99.96%/0.01%
Single	0.621 (0.139)	0.307	0.936	4.468	.002				
Multi	0.620 (0.293)	−0.042	1.282	2.120	.063				
Placement (11)						0.147 (2, 8)	.866	0.033	99.97%/99.96%/0.01%
Leg	0.471 (0.314)	−0.253	1.000	1.500	<.100				
Trunk	0.630 (0.180)	0.215	1.000	3.501	<.010				
Arm	0.682 (0.241)	0.125	1.000	2.826	<.050				

Abbreviations: LB, lower bound; PB, physical behavior; r, Pearson correlation coefficient; R^2, coefficient of determination; UB, upper bound.

Table 4 Moderator analysis for criterion validity.

	Criterion Validity								
	Summary Effect and 95% CI					Test of Moderation			
Moderator (k)	r (SE)/Slope (SE)	95% CI LB	95% CI UB	t	P	F (df1, df2)	P	R^2	Total I^2/Level 2 I^2/Level 3 I^2
PA dimensions (41)						1.250 (2, 38)	.298	0.137	98.67%/30.74%/67.92%
Frequency	0.784 (0.088)	0.605	0.963	8.863	<.001				
Intensity	0.602 (0.080)	0.441	0.764	7.540	<.001				
Duration	0.759 (0.090)	0.576	0.942	8.405	<.001				
% GMFCS level 1 (42)	0.002 (0.003)	−0.003	0.007	0.703	.486	0.494 (1, 40)	.486	0.037	98.77%/18.97%/79.80%
PB type specified (29)						3.582 (3, 25)	.028	0.177	96.95%/0.01%/96.94%
Lying and sitting[*],[†]	0.463 (0.113)	0.230	0.696	4.095	<.001				
Standing	0.643 (0.105)	0.427	0.859	6.123	<.001				
Walking[*]	0.733 (0.082)	0.564	0.903	8.913	<.001				
Running[†]	0.756 (0.094)	0.563	0.949	8.064	<.001				
Age (37)						0.142 (1, 35)	.709	0.012	98.92%/18.82%/80.10%
<13 yo	0.740 (0.073)	0.593	0.888	10.170	<.001				
≥13 yo	0.689 (0.116)	0.454	0.924	5.955	<.001				

Table 4 (*Continued*)

	Criterion Validity								
	Summary Effect and 95% CI					Test of Moderation			
Moderator (k)	*r* (SE)/Slope (SE)	95% CI LB	95% CI UB	*t*	*P*	F (*df*1, *df*2)	*P*	R^2	Total I^2/Level 2 I^2/Level 3 I^2
PB type (41)						0.265 (1, 39)	.610	0.008	98.90%/16.26%/82.64%
Single	0.723 (0.073)	0.577	0.870	9.977	<.001				
Multi	0.676 (0.084)	0.507	0.845	8.076	<.001				
Placement (42)						0.307 (3, 38)	.820	0.033	98.97%/17.60%/81.36%
Leg	0.650 (0.079)	0.489	0.810	8.207	<.001				
Trunk	0.727 (0.083)	0.559	0.895	8.761	<.001				
Arm	0.715 (0.130)	0.452	0.978	5.507	<.001				
Multi	0.789 (0.239)	0.306	1.000	3.305	.002				

Abbreviations: LB, lower bound; PB, physical behavior; *r*, Pearson correlation coefficient; R^2, coefficient of determination; UB, upper bound; yo, years old.
[*,†] Differences between effect sizes are statistically significant.

Table 5 Moderator analysis for criterion validity of cutoff point based methods.

	Criterion Validity of Cutoff Point Based Methods								
	Summary Effect and 95% CI					Test of Moderation			
Moderator (k)	*r* (SE)/Slope (SE)	95% CI LB	95% CI UB	*t*	*P*	F (*df*1, *df*2)	*P*	R^2	Total I^2/Level 2 I^2/Level 3 I^2
PA dimensions (24)						0.098 (1, 22)	.757	0.010	88.02%/42.82%/45.20%
Frequency	0.699 (0.073)	0.547	0.850	9.559	<.001				
Intensity	0.663 (0.086)	0.485	0.842	7.716	<.001				
Duration									
% GMFCS level 1 (24)	0.001 (0.006)	−0.012	0.014	0.188	.853	0.035 (1, 22)	.853	0.003	88.30%/40.68%/47.62%
PB class (22)						9.388 (2,19)	.002	0.340	84.26%/6.18%/78.08%
Sedentary[*,†]	0.755 (0.063)	0.623	0.888	11.907	<.001				
Light[*]	0.469 (0.087)	0.288	0.651	5.406	<.001				
Moderate-to-vigorous[†]	0.610 (0.070)	0.465	0.756	8.773	<.001				
Age (24)						2.009 (1, 22)	.170	0.126	86.00%/42.31%/43.69%
<6 yo	0.498 (0.143)	0.202	0.795	3.483	.002				
≥6 yo	0.715 (0.055)	0.601	0.830	12.952	<.001				
Age (24)						21.364 (1, 22)	.001	0.406	78.31%/78.30%/0.01%
<13 yo	0.610 (0.033)	0.542	0.678	18.629	<.001				
≥13 yo	0.911 (0.056)	0.795	1.000	16.204	<.001				
Cutoff points (24)						1.293 (1, 22)	.268	0.061	86.73%/47.18%/39.56%
Population specific	0.703 (0.054)	0.590	0.815	12.930	<.001				
General	0.611 (0.083)	0.439	0.783	7.380	<.001				

Abbreviations: LB, lower bound; PB, physical behavior; *r*, Pearson correlation coefficient; R^2, coefficient of determination; UB, upper bound; yo, years old.
[*,†] Differences between effect sizes are statistically significant.

References

1. The Lancet Child Adolescent Health. Enabling participation in physical activity. Lancet Child Adolesc Health 2022;6:71.
2. Carbone PS, Smith PJ, Lewis C, LeBlanc C. Promoting the participation of children and adolescents with disabilities in sports, recreation, and physical activity. Pediatrics 2021;148:e2021054664.
3. Olusanya BO, Davis AC, Hadders-Algra M, Wright SM. Global investments to optimise the health and wellbeing of children with disabilities: a call to action. Lancet 2023;401:175-7.
4. Steene-Johannessen J, Hansen BH, Dalene KE, et al. Variations in accelerometry measured physical activity and sedentary time across Europe - harmonized analyses of 47,497 children and adolescents. Int J Behav Nutr Phys Act 2020;17:38.
5. World Health Organization. Global status report on physical activity 2022. Available at: https://www.who.int/teams/health-promotion/physical-activity/global-status-report-on-physical-activity-2022. Accessed November 25, 2023.
6. Bloemen MAT, van den Berg-Emons RJG, Tuijt M, et al. Physical activity in wheelchair-using youth with spina bifida: an observational study. J Neuroeng Rehabil 2019;16:9.
7. Verschuren O, Peterson MD, Balemans AC, Hurvitz EA. Exercise and physical activity recommendations for people with cerebral palsy. Dev Med Child Neurol 2016;58:798-808.
8. World Health Organization. Physical activity fact sheet. Available at: https://www.who.int/publications/i/item/WHO-HEP-HPR-RUN-2021.2. Accessed October 13, 2022.
9. World Health Organization. Adolescent health. Available at: https://www.who.int/health-topics/adolescent-health/#tab=-tab_1. Accessed March 05, 2024.
10. White L, Volfson Z, Faulkner G. Arbour-Nicitopoulos K. Reliability and validity of physical activity instruments used in children

and youth with physical disabilities: a systematic review. Pediatr Exerc Sci 2016;28:240-63.

11. Trost SG, O'Neil M. Clinical use of objective measures of physical activity. Br J Sports Med 2014;48:178-81.

12. Prince SA, Cardilli L, Reed JL, et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2020;17:31.

13. Lettink A, Altenburg TM, Arts J, van Hees VT, Chinapaw MJM. Systematic review of accelerometer-based methods for 24-h physical behavior assessment in young children (0-5 years old). Int J Behav Nutr Phys Act 2022;19:116.

14. Ahmadi MN, Pfeiffer KA, Trost SG. Physical activity classification in youth using raw accelerometer data from the hip. Meas Phys Educ Exerc Sci 2020;24:129-36.

15. Downs J, Leonard H, Hill K. Initial assessment of the StepWatch Activity Monitor™ to measure walking activity in Rett syndrome. Disabil Rehabil 2012;34:1010-5.

16. Lankhorst K, van den Berg-Emons RJ, Bussmann JBJ, Horemans HLD, de Groot JF. A novel tool for quantifying and promoting physical activity in youths with typical development and youths who are ambulatory and have motor disability. Phys Ther 2019;99:354-63.

17. Paraschiv-Ionescu A, Newman CJ, Carcreff L, Gerber CN, Armand S, Aminian K. Locomotion and cadence detection using a single trunk-fixed accelerometer: validity for children with cerebral palsy in daily life-like conditions. J Neuroeng Rehabil 2019;16:24.

18. Xing R, Huang WY, Sit CH. Validity of accelerometry for predicting physical activity and sedentary time in ambulatory children and young adults with cerebral palsy. J Exerc Sci Fit 2021;19:19-24.

19. Ryan JM, Forde C, Hussey JM, Gormley J. Comparison of patterns of physical activity and sedentary behavior between children with cerebral palsy and children with typical development. Phys Ther 2015;95:1609-16.

20. O'Neil ME, Fragala-Pinkham MA, Forman JL, Trost SG. Measuring reliability and validity of the ActiGraph GT3X accelerometer for children with cerebral palsy: a feasibility study. J Pediatr Rehabil Med 2014;7:233-40.

21. Lawal TA, Todd JJ, Elliott JS, et al. Assessing motor function in congenital muscular dystrophy patients using accelerometry. J Neurosci Nurs 2020;52:172-8.

22. Bonnefoy-Mazure A, Armand S. Normal gait. In: Canavese F, Deslandes J, eds. Orthopedic management of children with cerebral palsy, Hauppauge: Nova Science Publishers, Inc; 2015. p. 211.

23. World Health Organization. ICD-10: international statistical classification of diseases and related health problems: tenth revision, 2nd ed. Available at: https://apps.who.int/iris/handle/10665/42980. Accessed May 30, 2023.

24. Trost SG, Fragala-Pinkham M, Lennon N, O'Neil ME. Decision trees for detection of activity intensity in youth with cerebral palsy. Med Sci Sports Exerc 2016;48:958-66.

25. Nooijen CF, de Groot JF, Stam HJ, van den Berg-Emons RJ, Bussmann HB. Fit for the Future Consortium. Validation of an activity monitor for children who are partly or completely wheelchair-dependent. J Neuroeng Rehabil 2015;12:11.

26. Stephens S, Takken T, Esliger DW, et al. Validation of accelerometer prediction equations in children with chronic disease. Pediatr Exerc Sci 2016;28:117-32.

27. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147-57.

28. Mokkink LB, Prinsen CAC, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs). Version 1.0. Available at: https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf. Accessed May 30, 2021.

29. Moher D, Liberati A, Tetzlaff J, Altman DG. PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.

30. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.r-project.org/ Accessed September 10, 2023.

31. Novak-Pavlic M, Abdel Malek S, Rosenbaum P, Macedo LG, Di Rezze B. A scoping review of the literature on grandparents of children with disabilities. Disabil Rehabil 2022;44:3326-48.

32. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res 2009;18:1115-23.

33. Bloemen MA, Backx FJ, Takken T, et al. Factors associated with physical activity in children and adolescents with a physical disability: a systematic review. Dev Med Child Neurol 2015;57:137-48.

34. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1171-9.

35. Mokkink LB, Boers M, van der Vleuten CPM, et al. COSMIN risk of bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. BMC Med Res Methodol 2020;20:293.

36. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651-7.

37. Mokkink LB, Elsman EBM, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. Qual Life Res 2024;33(11):2929-39.

38. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1-48.

39. Pastor DA, Lazowski RA. On the multilevel nature of meta-analysis: a tutorial, comparison of software programs, and discussion of analytic choices. Multivariate Behav Res 2018;53:74-89.

40. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 2019;22:153-60.

41. Beath KJ. Metaplus: an R package for the analysis of robust meta-analysis and meta-regression. R Journal 2016;8:5-16.

42. Fan X, Sun S. Generalizability theory as a unifying framework of measurement reliability in adolescent research. J Early Adolesc 2014;34:38-65.

43. Friedman L, Nixon MS, Komogortsev OV. Method to assess the temporal persistence of potential biometric features: application to oculomotor, gait, face and brain structure databases. PLoS One 2017;12:e0178501.

44. Pustejovsky JE. clubSandwich: cluster-robust (sandwich) variance estimators with small-sample corrections. R package version 0.5.6. 2022. Available at: https://github.com/jepusto/clubSandwich. Accessed September 10, 2023.

45. Harrer M, Cuijpers P, Furukawa T, Ebert DD. Companion R package for the guide "Doing Meta-Analysis in R." 2019. http://dmetar.protectlab.org/. Accessed July 31, 2023.

46. Nakagawa S, Lagisz M, O'Dea RE, et al. The orchard plot: cultivating a forest plot for use in ecology, evolution, and beyond. Res Synth Methods 2021;12:4-12.

47. Fisher Z, Tipton E, Hou Z. Robumeta: robust variance meta-regression. 2017. Available at: https://cran.r-project.org/. Accessed September 10, 2023.

48. Borenstein M. In a meta-analysis, the I-squared statistic does not tell us how much the effect size varies. J Clin Epidemiol 2022;152:281-4.

49. Higgins J, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions. Version 6.4. 2023. Available at: https://training.cochrane.org/handbook/current. Accessed December 11, 2023.

50. Simonsohn U, Simmons JP, Nelson LD. Better P-curves: making P-curve analysis more robust to errors, fraud, and ambitious P-hacking, a reply to Ulrich and Miller (2015). J Exp Psychol Gen 2015;144:1146-52.

51. Mathur MB, VanderWeele TJ. Sensitivity analysis for publication bias in meta-analyses. J R Stat Soc Ser C Appl Stat 2020;69:1091-119.

52. Körner R, Röseler L, Schütz A, Bushman BJ. Dominance and prestige: meta-analytic review of experimentally induced body position effects on behavioral, self-report, and physiological dependent variables. Psychol Bull 2022;148:67-85.

53. Vembye MH, Pustejovsky JE. POMADE: power for meta-analysis of dependent effects. R package version 0.1.0. 2022. Available at: https://mikkelvembye.github.io/POMADE/. Accessed September 10, 2023.

54. Cabot M, Daviet JC, Duclos N, Bernikier D, Salle JY, Compagnat M. First systematic review and meta-analysis of the validity and test-retest reliability of physical activity monitors for estimating energy expenditure during walking in individuals with stroke. Arch Phys Med Rehabil 2022;103:2245-55.

55. Kuo YL, Culhane KM, Thomason P, Tirosh O, Baker R. Measuring distance walked and step count in children with cerebral palsy: an evaluation of two portable activity monitors. Gait Posture 2009;29:304-10.

56. Oftedal S, Bell KL, Davies PS, Ware RS, Boyd RN. Validation of accelerometer cut points in toddlers with and without cerebral palsy. Med Sci Sports Exerc 2014;46:1808-15.

57. Pirpiris M, Graham HK. Uptime in children with cerebral palsy. J Pediatr Orthop 2004;24:521-8.

58. Tang KT, Richardson AM, Maxwell D, Spence WD, Stansfield BW. Evaluation of an activity monitor for the objective measurement of free-living physical activity in children with cerebral palsy. Arch Phys Med Rehabil 2013;94:2549-58.

59. Bania T. Measuring physical activity in young people with cerebral palsy: validity and reliability of the activPAL^{TM} monitor. Physiother Res Int 2014;19:186-92.

60. Aviram R, Belokopytov M, Ben-Chaim S, Rotstein A. Evaluation of energy expenditure in children with cerebral palsy using a multi-sensor accelerometer. J Sports Med Phys Fitness 2011;51:506-14.

61. Baque E, Barber L, Sakzewski L, Boyd RN. Reproducibility in measuring physical activity in children and adolescents with an acquired brain injury. Brain Inj 2016;30:1692-8.

62. Braun S, Dillon E, Sheiko M, Kang M, Bjornson K, Song K. Reliably estimating ambulatory activity in youth with arthrogryposis. Disabil Rehabil 2016;38:749-53.

63. Gerber CN, Carcreff L, Paraschiv-Ionescu A, Armand S, Newman CJ. Reliability of single-day walking performance and physical activity measures using inertial sensors in children with cerebral palsy. Ann Phys Rehabil Med 2021;64:101250.

64. Ishikawa S, Kang M, Bjornson KF, Song K. Reliably measuring ambulatory activity levels of children and adolescents with cerebral palsy. Arch Phys Med Rehabil 2013;94:132-7.

65. Mitchell LE, Ziviani J, Boyd RN. Variability in measuring physical activity in children with cerebral palsy. Med Sci Sports Exerc 2015;47:194-200.

66. Mackey AH, Hewart P, Walt SE, Stott NS. The sensitivity and specificity of an activity monitor in detecting functional activities in young people with cerebral palsy. Arch Phys Med Rehabil 2009;90:1396-401.

67. Maher C, Kenyon A, McEvoy M, Sprod J. The reliability and validity of a research-grade pedometer for children and adolescents with cerebral palsy. Dev Med Child Neurol 2013;55:827-33.

68. Koehler K, Abel T, Wallmann-Sperlich B, Dreuscher A, Anneken V. Energy expenditure in adolescents with cerebral palsy: comparison of the sensewear armband and indirect calorimetry. J Phys Act Health 2015;12:540-5.

69. O'Neil ME, Fragala-Pinkham M, Lennon N, George A, Forman J, Trost SG. Reliability and validity of objective measures of physical activity in youth with cerebral palsy who are ambulatory. Phys Ther 2016;96:37-45.

70. Clanchy KM, Tweedy SM, Boyd RN, Trost SG. Validity of accelerometry in ambulatory children and adolescents with cerebral palsy. Eur J Appl Physiol 2011;111:2951-9.

71. Keawutan P, Bell KL, Oftedal S, Davies PS, Boyd RN. Validation of accelerometer cut-points in children with cerebral palsy aged 4 to 5 years. Pediatr Phys Ther 2016;28:427-34.

72. McAloon MT, Hutchins S, Twiste M, Jones R, Forchtner S. Validation of the activPAL activity monitor in children with hemiplegic gait patterns resultant from cerebral palsy. Prosthet Orthot Int 2014;38:393-9.

73. Ryan J, Walsh M, Gormley J. Ability of RT3 accelerometer cut points to detect physical activity intensity in ambulatory children with cerebral palsy. Adapt Phys Activ Q 2014;31:310-24.

74. Sala DA, Grissom HE, Delsole EM, et al. Measuring ambulation with wrist-based and hip-based activity trackers for children with cerebral palsy. Dev Med Child Neurol 2019;61:1309-13.

75. Capio CM, Sit CH, Abernethy B. Physical activity measurement using MTI (actigraph) among children with cerebral palsy. Arch Phys Med Rehabil 2010;91:1283-90.

76. de Groot JF, de Jong AS, Visser T, Takken T. Validation of the actical and actiheart monitor in ambulatory children with spina bifida. J Pediatr Rehabil Med 2013;6:103-11.

77. Baque E, Sakzewski L, Trost SG, Boyd RN, Barber L. Validity of accelerometry to measure physical activity intensity in children with an acquired brain injury. Pediatr Phys Ther 2017;29:322-9.

78. O'Donoghue D, Kennedy N. Validity of an activity monitor in young people with cerebral palsy gross motor function classification system level I. Physiol Meas 2014;35:2307-18.

79. Oftedal S, Bell KL, Davies PS, Ware RS, Boyd RN. Sedentary and active time in toddlers with and without cerebral palsy. Med Sci Sports Exerc 2015;47:2076-83.

80. Hendry D, Rohl AL, Rasmussen CL, et al. Objective measurement of posture and movement in young children using wearable sensors and customised mathematical approaches: a systematic review. Sensors (Basel) 2023;23:9661.

81. Holsbeeke L, Ketelaar M, Schoemaker MM, Gorter JW. Capacity, capability, and performance: different constructs or three of a kind? Arch Phys Med Rehabil 2009;90:849-55.

82. Lynch BA, Kaufman TK, Rajjo TI, et al. Accuracy of accelerometers for measuring physical activity and levels of sedentary behavior in children: a systematic review. J Prim Care Community Health 2019;10:2150132719874252.

83. McIntyre S, Goldsmith S, Webb A, et al. Global prevalence of cerebral palsy: a systematic analysis. Dev Med Child Neurol 2022;64:1494-506.

84. Salari N, Fatahi B, Valipour E, et al. Global prevalence of Duchenne and Becker muscular dystrophy: a systematic review and meta-analysis. J Orthop Surg Res 2022;17:96.

85. Blencowe H, Kancherla V, Moorthie S, Darlison MW, Modell B. Estimates of global and regional prevalence of neural tube defects for 2015: a systematic analysis. Ann N Y Acad Sci 2018;1414:31-46.

86. Mokkink L, Terwee C, de Vet H. Key concepts in clinical epidemiology: responsiveness, the longitudinal aspect of validity. J Clin Epidemiol 2021;140:159-62.

87. Baque E, Barber L, Sakzewski L, Boyd R. Validity of accelerometry to measure physical activity in children and adolescents with an acquired brain injury. Dev Med Child Neurol 2015;57 (S5):77-8.

88. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci 2008;26(14):1557-65.