Calibration and Cross-Validation of the ActiGraph wGT3X+ Accelerometer for the Estimation of Physical Activity Intensity in Children with Intellectual Disabilities

Arlene M. McGarty[1]*, **Victoria Penpraze**[2], **Craig A. Melville**[1]

1 Institute of Health & Wellbeing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, Scotland, 2 School of Life Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, Scotland

* Arlene.McGarty@glasgow.ac.uk

Abstract

Background

Valid objective measurement is integral to increasing our understanding of physical activity and sedentary behaviours. However, no population-specific cut points have been calibrated for children with intellectual disabilities. Therefore, this study aimed to calibrate and cross-validate the first population-specific accelerometer intensity cut points for children with intellectual disabilities.

Methods

Fifty children with intellectual disabilities were randomly assigned to the calibration ($n = 36$; boys = 28, 9.53±1.08yrs) or cross-validation ($n = 14$; boys = 9, 9.57±1.16yrs) group. Participants completed a semi-structured school-based activity session, which included various activities ranging from sedentary to vigorous intensity. Direct observation (SOFIT tool) was used to calibrate the ActiGraph wGT3X+, which participants wore on the right hip. Receiver Operating Characteristic curve analyses determined the optimal cut points for sedentary, moderate, and vigorous intensity activity for the vertical axis and vector magnitude. Classification agreement was investigated using sensitivity, specificity, total agreement, and Cohen's kappa scores against the criterion measure of SOFIT.

Results

The optimal (AUC = .87−.94) vertical axis cut points (cpm) were ≤507 (sedentary), 1008−2300 (moderate), and ≥2301 (vigorous), which demonstrated high sensitivity (81−88%) and specificity (81−85%). The optimal (AUC = .86−.92) vector magnitude cut points (cpm) of ≤1863 (sedentary), 2610−4214 (moderate), and ≥4215 (vigorous) demonstrated comparable, albeit marginally lower, accuracy than the vertical axis cut points (sensitivity = 80−86%; specificity = 77−82%). Classification agreement ranged from moderate to almost

perfect (κ = .51−.85) with high sensitivity and specificity, and confirmed the trend that accuracy increased with intensity, and vertical axis cut points provide higher classification agreement than vector magnitude cut points.

Conclusions

This study provides the first valid methods of interpreting accelerometer output in children with intellectual disabilities. The calibrated physical activity cut points are notably higher than existing cut points, thus raising questions on the validity of previous low physical activity estimates in children with intellectual disabilities that were based on typically developing cut points.

Introduction

Physical activity is associated with many physical and mental health benefits in children, such as reduced body mass index (BMI), increased bone health, reduced risk of metabolic syndromes, and lower rates of depression [1–3]. Furthermore, childhood physical activity is a determinant of physical activity in adulthood, therefore it is important that positive behaviours are promoted in children [4, 5]. This is of primary importance in children with intellectual disabilities as this population participate in low levels of activity and have a higher prevalence of negative health outcomes in comparison with their typically developing peers [6, 7]. Therefore, there is a need to increase our understanding of physical activity in this population and develop effective interventions to increase activity levels. In accordance with best practice guidelines on the development of effective interventions, research has to be based on valid measurement of various parameters of activity, such as intensity, frequency, and duration [8, 9]. This will increase our understanding of dose-response relationships and determinants of activity, from which interventions can be developed.

Free-living physical activity can be measured using subjective methods (such as questionnaires or activity logs) or objective methods (such as accelerometers and pedometers). Subjective measures provide feasible methods for collecting data on various parameters of physical activity, but this high feasibility is at a cost to validity. Due to the cognitive demands associated with subjective measures, this lower validity is particularly apparent in children, which will be exacerbated in children with intellectual disabilities [10 11]. Of the commonly used objective measures, pedometers are affected by abnormal gait patterns and heart monitors affected by atypical heart rates, both of which are prevalent in children with intellectual disabilities [11–14]. Accelerometers therefore provide the most comprehensive and feasible method of measuring parameters of activity in children with intellectual disabilities.

Accelerometers are small, non-intrusive devices that measure acceleration of the body during movement. In general, older accelerometers measure acceleration on the vertical axis of the body, but advantaging technology now enables acceleration to be measured on up to three planes [15]. Acceleration signals are post-processed and converted into arbitrary activity "counts" for the vertical axis or combined three axes (vector magnitude), which can be calibrated to provide biologically meaningful data, such as physical activity intensity [16].

Accelerometer calibration is age- and population-specific due to the effects of maturation and between group differences, such as cardiorespiratory fitness [17, 18]. Therefore, generalising intensity cut points between populations introduces systematic measurement error and reduces validity. However, accelerometers have not been calibrated for children with

intellectual disabilities, thus raising questions on the validity of generalising typically developing cut points to children with intellectual disabilities due to movement and metabolic differences between these groups [18–20]. As a result, the calibration of accelerometer cut points is an important next step is the development of a valid knowledge-base on the physical activity behaviours of children with intellectual disabilities, from which effective interventions can be developed.

Therefore, the aims of this study are to: 1) calibrate ActiGraph wGT3X+ vertical axis and vector magnitude cut points for the estimation of sedentary, moderate, and vigorous intensity activity in children with intellectual disabilities, and; 2) cross-validate the developed cut points in a sub-sample of children with intellectual disabilities.

Methods

Ethical Consideration

This study was approved by the Medical, Veterinary, and Life Sciences College Ethics Committee, University of Glasgow. Prior to participation, written informed consent was obtained from participants and parents.

Participants

Five additional support needs primary schools in the West of Scotland, which were specifically for children with mild to moderate intellectual disabilities, were used for recruitment and data collection. To be eligible for participation, children had to be aged 8 to 11 years, have intellectual disabilities, and be independently ambulatory. No data was collected on the aetiology of disabilities of participants due to feasibility reasons, e.g. the need to conduct IQ tests. In total, 86 information packs were handed out to eligible children (60 boys, 26 girls), which resulted in 50 children taking part, representing a recruitment rate of 58.14%.

Protocol

A semi-structured physical activity session was designed specifically for this calibration study. A semi-structured protocol was used to increase and ecological validity and ensure that children participated in sufficient activity at each intensity. The session content was developed based on previous field-based calibration and validation research conducted in typically developing children and discussions with teachers to ensure the use of appropriate and familiar activities. In addition, the energy expenditure compendium for youth was used as a guide to inform the inclusion of activities that were of the required activity intensities [21]. The semi-structured session was developed with four main phases: warm-up, instruction games, obstacle games, and team games (Table 1). Although all sessions included these four phases, the semi-

Table 1. Description of activity protocol.

Phase	Description of main activities	Approximate time (min)	Intensity
Warm up	Active stretches, walking, jogging	10	Light/moderate
Instruction games	Moving around the hall following instructions from the researcher, e.g. star jump, touch floor.	10	Moderate
Obstacle games	Short courses included various aerobic activities, e.g. agility ladder, zig-zag cones, bench jumps.	10	Vigorous
Team games	Dodge ball and tag.	15	Vigorous

doi:10.1371/journal.pone.0164928.t001

structured nature of the protocol enabled the activities to be adapted to suit the participants' level of ability.

Measures

Anthropometric. Height was measured to the nearest 0.1 centimetre using a stadiometer (Seca Scales, Hamburg, Germany) and weight measured to the nearest 0.1 kilogram using digital scales (Seca Scales, Hamburg, Germany). Measurements were conducted twice to produce a mean value whilst participants were wearing light clothing and no shoes.

Accelerometry. Physical activity was measured using the ActiGraph wGT3X+ accelerometer (ActiGraph, LLC, Pensacola, FL, USA). This small, lightweight device ($46 \times 33 \times 15$ mm, 19 g) measures acceleration during movement across the vertical, horizontal, and perpendicular axes. Prior to the session, the accelerometers were initialized to record accelerations at a sampling frequency of 30Hz. Participants wore one device on their right hip at the iliac crest, attached using an elastic belt, for the duration of the activity session.

Criterion measure: Direct observation. The System for Observing Fitness Instruction Time (SOFIT) is a momentary time sampling direct observation tool [22]. The 'student activity' element of this tool categorises physical activity behaviours as: lying down, sitting, standing, walking, and very active, and has been validated for typically developing children and children with intellectual disabilities [23–25]. Activity was coded every 20 seconds using 10-second observe/record intervals, yielding 3 observations per minute, and was paced using pre-recorded audio MP4 files [26]. If the participant was transitioning from one activity to another at the end of the observe interval, the activity was recorded as the higher code; for example, if transitioning from lying down to sitting, the activity was coded as sitting [26].

To ensure the SOFIT data were reliable and valid, three raters (AMMcG, VP, CAM) were involved in data coding. Raters undertook eight hours of classroom-based training which included understanding coding procedures and definitions, video analysis practice, and validity assessment [26]. A combined accuracy of 86% was achieved with the gold standard assessment video, which exceeds the minimum recommendation of 80% [26]. Data from one rater (AMMcG) was used for data collection, with the other raters (VP & CAM) used to assess field-based reliability. Initial reliability was established using data from two randomly selected participants from session one, with inter- and intra-rater reliability scores of 79% and 89%, respectively, achieved. At the midpoint of data collection, reliability was further investigated using two randomly selected participants from session three, with inter- and intra-rater reliability of 85% and 91% achieved, respectively, thus confirming the data collected by the lead rater was reliable.

Management of data. Accelerometer data for the vertical axis and vector magnitude were downloaded using ActiLife version 6.11.5 software (ActiGraph LLC, Pensacola) in 10-second epochs. Video data was time matched to the accelerometer data to ensure the SOFIT coding started at the beginning of a 10-second accelerometer epoch. Two consecutive 10-second vertical axis and vector magnitude epochs were summed using an Excel macro to correspond with one 20-second SOFIT epoch. This produced data in the following formats: vertical axis counts (counts/20-sec), vector magnitude counts (counts/20-sec), and SOFIT classification (score/20-sec). Data were then screened for spurious scores and epochs where the participant left the gym hall were excluded from the analysis. Data in this format were used for all analyses. Prior to conducting the calibration analyses, data for 14 participants were removed to enable cross-validation analyses, with two participants randomly selected from each of the seven sessions, resulting in data from 36 participants being used for calibration.

Statistical analysis. All statistical analyses were conducted using SPSS 22 IBM statistical package (SPSS IBM, New York, NY, USA). Descriptive statistics (mean ± SD) were calculated

for all participant and session variables. Independent samples t-tests were additionally conducted to test for differences in age, height, weight, and BMI between the calibration and cross-validation groups. Receiver operating characteristic (ROC) curve analyses were conducted to determine the optimal cut points for the classification of sedentary, moderate, and vigorous intensity activity. ROC curve analysis quantifies the relationship between positive and negative scores for continuous data and allows a cut point to be identified which best discriminates between two conditions [27]. A "positive" score represents the condition of interest, whereas a "negative" score is not the condition of interest.

ROC curves were interpreted using sensitivity, specificity, and the area under the curve (AUC) of the ROC curve. Sensitivity is the accuracy of a cut point to correctly classify activity intensity (true positive) whereas specificity is the accuracy of a cut point to exclude data which is not of the specified intensity (false positive). In addition, the AUC gives a statistical representation of the accuracy of the optimal cut point. The AUC is the average true positive classification rate, independent of false positive classifications. Therefore, a cut point which perfectly classifies all scores will have an AUC of 1.0, with a cut point equivalent to chance having an AUC of .50. The AUC scores will be interpreted using the following scale: $\geq$ .90 is excellent, .80-.89 is good, .70-.79 is fair, and $<$ .70 is poor [28]. In line with previous accelerometer calibration studies, the aim of this ROC curve analysis was to identify the cut point which maximises both sensitivity and specificity. This optimal cut points was identified by manually applying Youden's index to all identified cut points: Youden's index = maximum (sensitivity + sensitivity– 1).

For the ROC curve analysis, SOFIT scores were converted into binary scores, with binary code 1 representing a positive score and binary code 0 representing a negative score. Separate analyses were conducted for the calibration of the sedentary, moderate, and vigorous cut points for the vertical axis and vector magnitude, with the vigorous cut point providing the upper boundary for the moderate intensity cut point. Table 2. summarises the binary codes used for each analysis.

For cross-validation, classification agreement between SOFIT and the calibrated cut points was investigated using sensitivity, specificity, total agreement percentages, and Cohen's kappa scores. Kappa scores (κ) provide a statistical measure of agreement, accounting for agreements which may occur by chance, and will be interpreted using the following scale: $<$ .00 is less than change agreement, .00–.20 is slight agreement, .21–.40 is fair agreement, .41–.60 is moderate agreement, .61–.80 is substantial agreement, and .81–1.00 is almost perfect agreement [29].

Results

Descriptive statistics for all participants are presented in Table 3. There were no significant differences in age (t = -.13, df = 48, p > .01, 95% CI -.74, .65) height (t = .08, df = 48, p > .01, 95% CI -.05, .06), weight (t = -.06, df = 48, p > .01, 95% CI -6.78, 6.38), or BMI (t = -.32, df = 48, p > .05, 95% CI -2.81, 2.05) between the participants in the calibration and cross-validation

Table 2. Summary of the conversion of SOFIT categories into binary codes for ROC curve analysis.

Activity intensity	SOFIT categories	
	Binary code 1 (positive)	Binary code 0 (negative)
Sedentary	1, 2	3, 4, 5
Moderate	4, 5	1, 2, 3
Vigorous	5	1, 2, 3, 4

SOFIT categories: lying down = 1; sitting = 2; lying down = 3; walking = 4; very active = 5

doi:10.1371/journal.pone.0164928.t002

Table 3. Descriptive statistics (mean ± SD) of all participants and participants in the calibration and cross-validation groups.

Characteristic	Participants		
	All participants (n = 50)	Calibration (n = 36)	Cross-validation (n = 14)
Age (yrs)	9.54 ± 1.09	9.53 ± 1.08	9.57 ± 1.16
Height (m)	1.43 ± .09	1.43 ± .09	1.43 ± .08
Weight (kg)	39.33 ± 10.28	39.28 ± 11.46	39.48 ± 6.69
BMI (kg/m^2)	19.09 ± 3.80	18.99 ± 4.15	19.37 ± 2.81
Sex	M = 37, F = 13	M = 28, F = 8	M = 9, F = 6

doi:10.1371/journal.pone.0164928.t003

groups, suggesting the cross-validation sample is representative of the calibration group. Seven activity sessions were conducted, with descriptive data for each session presented in Table 4. In total, 1057 minutes of data were used for the calibration analyses and 382 minutes used for cross-validation. This equates to 256, 257, 338, and 206 minutes for sedentary, moderate, light, and vigorous intensity activity for calibration and 92, 82, 107, and 101 minutes for each intensity, respectively, for cross-validation.

The sedentary, moderate, and vigorous count boundaries developed (cpm) were $\leq$ 507, 1008 −2300, and $\geq$ 2301 for the vertical axis and $\leq$ 1863, 2610−4214, and $\geq$ 4215 for vector magnitude, respectively. As shown in Table 5, these cut points exhibit high sensitivity (80−88%) and

Table 4. Descriptive statistics on session duration, participants, and percentage of the session spent in each SOFIT category.

Session	School	Session duration (min)	Lying down (%)	Sitting (%)	Standing (%)	Walking (%)	Very active (%)	Boys (n)	Girls (n)
1	1	25	3.96	11.08	27.70	36.15	21.12	7	1
2	2	34	6.25	26.72	17.03	29.66	20.34	5	3
3	2	38	4.50	28.23	27.93	22.02	17.32	4	5
4	3	40	4.06	12.19	26.40	24.37	32.97	7	0
5	3	29	1.64	21.51	21.67	41.87	13.30	7	0
6	4	27	3.27	22.67	17.38	29.97	26.70	3	2
7	5	16	3.47	29.86	11.11	31.94	23.61	5	1

Note: two sessions were conducted at schools 2 and 3 for feasibility reasons due to the higher number of participants recruited

doi:10.1371/journal.pone.0164928.t004

Table 5. Calibration and cross-validation statistics for the derived cut points.

	Cut point (counts/20-sec)	Cut point (cpm)	Calibration			Cross-validation			
			Sensitivity (%)	Specificity (%)	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Total agreement (%)	κ (SE)
Vertical axis									
Sedentary	$\leq$ 169	$\leq$ 507	81	81	.87 (.86−.88)	93	83	85	.66* (.02)
Moderate	336−766	1008−2300	86	83	.92 (.91−.93)	75	96	90	.74* (.02)
Vigorous	$\geq$ 767	$\geq$ 2301	88	85	.94 (.93−.95)	93	95	94	.85* (.02)
MVPA	$\geq$ 336	$\geq$ 1008				91	95	93	.85* (.02)
Vector magnitude									
Sedentary	$\leq$ 621	$\leq$ 1863	80	77	.86 (.84−.87)	83	86	85	.63* (.03)
Moderate	870−1404	2610−4214	86	82	.92 (.91−.93)	60	89	79	.51* (.03)
Vigorous	$\geq$ 1405	$\geq$ 4215	85	82	.92 (.91−.93)	89	89	90	.74* (.02)
MVPA	$\geq$ 870	$\geq$ 2610				91	84	87	.75* (.02)

* Significant at p < .001

doi:10.1371/journal.pone.0164928.t005

specificity (77–85%) scores, with the accuracy of the cut points increasing with intensity (AUC = .86–.94). For the vertical axis and vector magnitude, the sedentary cut points demonstrated good classification accuracy, with the moderate and vigorous cut points demonstrating excellent AUC scores. The high sensitivity and specificity scores illustrate that these cut points should limit misclassifications, with the vertical axis cut points providing marginally higher accuracy than that vector magnitude cut points. ROC curve graphs for each cut point are presented in S1 Fig.

When cross-validated, the cut points demonstrated fair to almost perfect classification agreement (Table 5). The moderate intensity cut point for vector magnitude was the only cut point that demonstrated a substantially lower level of accuracy than expected, based on the calibration results. Consistent with the calibration findings, the vertical axis cut points demonstrated higher classification agreement, with classification agreement also increasing with activity intensity.

Discussion

This study calibrated and cross-validate the first accelerometer cut points for the classification of sedentary, moderate, and vigorous intensity activity in children with intellectual disabilities, and empirically demonstrated the need for population-specific cut points. Due to the associated health outcomes of sedentary behaviour and MVPA, and the need for more standardised objective measurement [30, 31], the current recommendation is to use a vertical axis cut point of 100 cpm for sedentary and the Evenson et al. [32] cut point of ≥ 2296 cpm for MVPA in typically developing children [33, 34]. However, as the sedentary and MVPA cut points developed in this study are notably different to these recommended cut points, it is important that the measurement of physical activity in children with intellectual disabilities is viewed independent of typically developing research.

To further highlight the effect of generalising cut points, the lower boundary of many exiting moderate intensity cut points are higher than the upper boundary of the present moderate cut point, i.e. > 2300 cpm (Table 6). Similarly, the calibrated vigorous cut point of ≥ 2301 cpm is also lower than some existing moderate intensity cut points, i.e. the lower boundary for the moderate cut points developed by Puyau et al. [35], Treuth et al. [36], and Mattocks et al. [37], are > 2301 cpm. Therefore, if generalised to children with intellectual disabilities, potentially all moderate and vigorous intensity activity will be misclassified. Furthermore, this raises questions on the validity of existing literature which has quantified physical activity levels in children with intellectual disabilities by generalising cut points calibrated in typically developing children (e.g. Einarsson et al. [6] and Phillips & Holland [38]).

With the evolving technology of ActiGraph devices, the ability to measure three axes and calculate vector magnitude should theoretically increase the accuracy of capturing the dynamic activity behaviours of children; however, there is currently no consensus on this [30, 39, 40]. Consistent with the findings of the present study, previous empirical research has demonstrated that vector magnitude cut points do not provide a consistently higher level of validity than vertical axis cut points [41–43]. The moderate and vigorous vector magnitude cut points in the present study are similar to the existing cut points, which were all calibrated using the ActiGraph GT3X device (Table 6). Therefore, as third generation ActiGraph devices record higher count values than older devices (e.g. the GT1M), the device used for calibration could have contributed to the smaller variance identified with existing vector magnitude cut points [44, 45]. Furthermore, as most existing vertical axis cut points were calibrated using GT1M or AM7164 devices, the wGT3X+ device could have contributed to the lower vector magnitude cut points developed in the present study. With the large

Table 6. Existing ActiGraph vertical axis and vector magnitude cut points calibrated for typically developing children.

Cut points	Sedentary	Moderate	Vigorous
Vertical Axis			
Puyau (2002)	0–799	3200–8199	≥ 8200
Treuth (2004)	0–100	3000–5200	≥ 5201
Freedson (2005)	0–500*	501–4000	4001–7600
Mattocks (2007)	0–100	3581–6129	≥ 6130
Evenson (2008)	0–100	2296–4011	≥ 4012
Pulsford (2011)	0–99	2241–3840	≥ 3841
Vanhelst (2011)	0–400	1901–3918	≥ 3919
Mackintosh (2012)	0–372	2161–4806	≥ 4807
Jimmy (2013)	n/a	1596–2315	≥ 2316
Romanzini (2014)	0–184	2428–3271	≥ 3272
Current study	0–507	1008–2300	≥ 2301
Vector Magnitude			
Santos-Lozano (2013)	0–2114*	2114–6547	6548–11490
Hanggi (2013)	0–180	> 3360 (MVPA)	
Jimmy (2013)	n/a	2952–3791	≥ 3792
Romanzini (2014)	0–720	3028–4447	≥ 4448
Current study	0–1863	2610–4214	≥ 4215

* Represents sedentary and light intensity

differences identified between the cut points calibrated in the presented study, it is important to further consider why such notable differences occurred, and additional factors which could have affected calibration.

As most of the activities used in previous calibration studies were structured and constant, the accelerometer will record more consistent counts and the criterion method will more accurately measure activity intensity [32, 41, 46–48]. The free-living design of the current protocol made it more difficult to discriminate between activity intensity, with the data recorded by SOFIT affected by transitions or epochs containing more than one intensity of activity. However, this protocol design will better account for the sporadic nature of children's activity, thus increasing ecological validity.

As the criterion measure, the validity of the developed cut points is dependent on the accurate use of SOFIT. Although a criterion measure, there is some subjectivity in the classification of what is deemed "ordinary" walking, which will vary between children and may include activity which is not of a moderate intensity. This could have contributed to the lower moderate intensity boundaries, particularly in comparison with existing cut points not established using SOFIT. Furthermore, as SOFIT does not capture extraneous movements which may be detected by the accelerometer, such as foot tapping, this could at least partially account for the higher sedentary cut points calibrated [49]. There is also a lack of consensus on whether standing should be coded as sedentary activity [50].

As calibration is based on observed behaviours, these findings show that during the same biomechanical movements, children with intellectual disabilities produce a smaller acceleration, which could be a result of abnormal gait patterns [13, 14]. On the other hand, it is important to note that adults with intellectual disabilities have a significantly higher energy expenditure than adults without intellectual disabilities during the same laboratory-based

activities [51]. Therefore, there is a need to further investigate biomechanical and physiological differences between children with and without intellectual disabilities, which could empirically identify population differences that could have contributed to the variance between the present cut points and existing cut points.

Strengths and Limitations

This study was the first to calibrate and cross-validate accelerometer cut points for physical activity intensity in children with intellectual disabilities, which provides the first stages of establishing valid methods to interpret accelerometer output in this population. With vector magnitude calibration in its relative infancy, this study has placed measurement research in children with intellectual disabilities in line with the emerging research in typically developing children. The free-living design of this protocol enabled calibration and cross-validation to be conducted on representative activities, thus increasing ecological validity. Cross-validation was also in accordance with best practice recommendations, in that validation was conducted during a field-based protocol in a sample independent of the calibration group [52, 53]. Furthermore, although participation rates are generally low for health-related research involving children with intellectual disabilities, the high recruitment rate within this study is similar to that reported in previous research involving typically developing children.

Not without limitations, the semi-structured protocol resulted in between-session differences (Table 4); therefore, the data collected may not be fully representative of the study sample as a whole. More generally, a limitation with cross-validation is that it only estimates how valid the cut points will be in a sample similar to the calibration sample [54]. As cut points are generally age-specific, investigation into the effect of age on validity is important [34]. However, as the age range of the participants in the present study was relatively small, it was not possible to make inferences regarding the validity of these cut in children with intellectual disabilities who are younger or older than the included sample. Furthermore, as specific data were not collected regarding the aetiology of participant's intellectual disabilities, it was not possible to investigate disability-related factors that could affect the generalisation of these cut points.

Conclusions

This study was the first to calibrate and cross-validate population-specific accelerometer cut points for the estimation of physical activity intensity in children with intellectual disabilities, thus addressing a substantial gap in measurement research relating to this population. Overall, the cut points developed in this study show high sensitivity and specificity for the estimation of physical activity and sedentary behaviours in children with intellectual disabilities, with most demonstrating high classification agreement. With the trend in this study that lower physical activity cut points and higher sedentary cut points are required, in comparison with typically developing children, possible causes for these differences have been discussed. To further increase our knowledge on the validity of the developed cut points, additional field-based and longitudinal validation research needs to be conducted. Moving forward, it is important to consider whether future research may benefit from taking an additional step back to basics and increase our knowledge relating to the biomechanics of how children with intellectual disabilities move and further investigate physiological differences with typically developing children. This will help our understanding of population-specific factors which may have influenced calibration and will help inform the next phases of improving the validity of objectively measured physical activity in children with intellectual disabilities.

Supporting Information

S1 Fig. ROC curves and optimal cut points for sedentary, moderate, and vigorous intensity for vertical axis and vector magnitude counts.
(DOCX)

Acknowledgments

We thank all the children who participated, and the teachers and schools who accommodated recruitment and data collection. We also thank Leanne Harris and Louise Bleazard for their assistance during data collection.

Author Contributions

Conceptualization: AMM CAM VP.

Data curation: AMM CAM VP.

Formal analysis: AMM CAM VP.

Funding acquisition: CAM VP.

Investigation: AMM CAM VP.

Methodology: AMM CAM VP.

Project administration: AMM.

Software: AMM.

Supervision: AMM CAM VP.

Validation: AMM CAM VP.

Visualization: AMM CAM VP.

Writing – original draft: AMM.

Writing – review & editing: AMM CAM VP.

References

1. Ahn S, Fedewa AL. A meta-analysis of the relationship between children's physical activity and mental health. J Pediatr Psychol. 2011; 36:385–397. doi: 10.1093/jpepsy/jsq107 PMID: 21227908

2. Biddle SJH, Asare M. Physical activity and mental health in children and adolescents: a review of reviews. Brit J Sport Med. 2011; 45:886–895.

3. Janssen I, LeBlanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phy. 2010; 7:1–16.

4. Telama R, Yang X, Viikari J, Välimäki I, Wanne O, Raitakari O. Physical activity from childhood to adulthood: a 21-year tracking study. Am J Prev Med. 2005; 28:267–273. doi: 10.1016/j.amepre.2004.12.003 PMID: 15766614

5. Telama R. Tracking of physical activity from childhood to adulthood: a review. Obesity Facts. 2009; 2:187–195. doi: 10.1159/000222244 PMID: 20054224

6. Einarsson IO, Olafsson A, Hinriksdóttir G, Jóhannsson E, Daly D, Arngrímsson SA. Differences in physical activity among youth with and without intellectual disability. Med Sci Sport Exer. 2015; 47:411–418.

7. Maiano C. Prevalence and risk factors of overweight and obesity among children and adolescents with intellectual disabilities. Obesity Reviews. 2010; 12:189–197.

8. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Brit Med J. 2008; 337.

9. Sallis JF, Owen N, Fotheringham MJ. Behavioral epidemiology: a systematic framework to classify phases of research on health promotion and disease prevention. Ann Behav Med. 2000; 22:294–298. PMID: 11253440

10. Matthews CE. Use of self-report instruments to assess physical activity. In: Welk G, editor. Physical Activity Assessments for Health-Related Research. Champaign, IL: Human Kinetics; 2002. p. 107–121.

11. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med 2001; 31:439–454. PMID: 11394563

12. Baynard T, Pitetti KH, Guerra M, Unnithan VB, Fernhall B. Age-related changes in aerobic capacity in individuals with mental retardation: a 20-yr review. Med Sci Sport Exer. 2008; 40:1984–1989.

13. Johnston TE, Moore SE, Quinn LT, Smith BT. Energy cost of walking in children with cerebral palsy: relation to the Gross Motor Function Classification System. Dev Med Child Neurol. 2004; 46:34–38. PMID: 14974645

14. Thomas SS, Buckon CE, Russman BS, Sussman MD, Aiona MD. A comparison of the changes in the energy cost of walking between children with cerebral palsy and able-bodied peers over one year. J Pediatr Rehabil Med. 2011; 4:225–233. doi: 10.3233/PRM-2011-0178 PMID: 22207099

15. Chen KY, Bassett DR. The technology of accelerometry-based monitors: current and future. Med Sci Sport Exer. 2005; 37:S490–S500.

16. Kim Y, Beets MW, Welk GJ. Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but. . .: a systematic review. J Sci Med Sport. 2012; 15:311–321. doi: 10.1016/j.jsams.2011.12.001 PMID: 22306372

17. Freedson P, Pober D, Janz KF. Calibration of accelerometer output for children. Med Sci Sport Exer. 2005; 37:S523–S530.

18. Frey GC, Stanish HI, Temple VA. Physical activity of youth with intellectual disability: review and research agenda. Adapt Phys Act Q. 2008; 25:95–117.

19. Hinckson EA, Curtis A. Measuring physical activity in children and youth living with intellectual disabilities: a systematic review. Res Dev Disabil. 2013; 34:72–86. doi: 10.1016/j.ridd.2012.07.022 PMID: 22940161

20. McGarty AM, Penpraze V, Melville CA. Accelerometer use during field-based physical activity research in children and adolescents with intellectual disabilities: a systematic review. Res Dev Disabil. 2014; 35:973–981. doi: 10.1016/j.ridd.2014.02.009 PMID: 24629542

21. Ridley K, Ainsworth BE, Olds TS. Development of a compendium of energy expenditures for youth. Int J Behav Nutr Phy. 2008; 5:45.

22. McKenzie TL, Sallis JF, Nader PR. SOFIT: system for observing fitness instruction time. J Teach Phys Educ. 1991; 11:195–205.

23. Faison-Hodge J, Porretta DL. Physical activity levels of students with mental retardation and students without disabilities. Adapt Phys Act Q. 2004; 21:139–152.

24. Rowe PJ, Schuldheisz JM, Van der Mars H. Measuring physical activity in physical education: validation of the SOFIT direct observation instrument for use with first to eighth grade students. Pediatr Exerc Sci. 1997; 9:136–149.

25. Rowe PJ, van Der Mars H, Schuldheisz J, Fox S. Measuring students' physical activity levels: validating SOFIT for use with high-school students. J Teach Phys Educ. 2004; 23:235–251.

26. McKenzie TL. System for observing fitness instruction time: generic description and procedures manual. San Diego: San Diego State University; 2009.

27. Krzanowski WJ, Hand DJ. ROC curves for continuous data. London: Taylor & Francis; 2009.

28. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem. 1993; 39:561–577. PMID: 8472349

29. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005; 37:360–363. PMID: 15883903

30. Reilly JJ, Penpraze V, Hislop J, Davies G, Grant S, Paton JY. Objective measurement of physical activity and sedentary behaviour: review with new data. Arch Dis Child. 2008; 93:614–619. doi: 10.1136/adc.2007.133272 PMID: 18305072

31. Tudor-Locke C, Barreira TV, Schuna JM, Katzmarzyk PT. Unique contributions of ISCOLE to the advancement of accelerometry in large studies. Int J Obesity Suppl. 2015; 5:S53–S58.

32. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sport Sci. 2008; 26:1557–1565.

33. Janssen X, Cliff D, Okely T, Hinkley T, Reilly JJ, Jones R, et al. Comparison of ActiGraph cut-points for predicting physical activity intensity in preschool children. J Sci Med Sport. 2013; 15:S66.

34. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. Med Sci Sport Exer. 2011; 43:1360–1368.

35. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. Obes Research. 2002; 10:150–157.

36. Treuth MS, Schmitz K, Catellier DJ, McMurray RG, Murray DM, Almeida MJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. Med Sci Sport Exer. 2004; 36:1259–1266.

37. Mattocks C, Leary S, Ness A, Deere K, Saunders J, Tilling K, et al. Calibration of an accelerometer during free-living activities. Int J Pediatr Obes. 2007; 2:218–226. doi: 10.1080/17477160701408809 PMID: 17852552

38. Phillips AC, Holland AJ. Assessment of objectively measured physical activity levels in individuals with intellectual disabilities with and without Down's syndrome. PLos ONE. 2011; 6:e28618. doi: 10.1371/journal.pone.0028618 PMID: 22205957

39. de Vries SI, Bakker I, Hopman-Rock M, Hirasing RA, van Mechelen W. Clinimetric review of motion sensors in children and adolescents. J Clin Epidemiol. 2006; 59:670–680. doi: 10.1016/j.jclinepi.2005.11.020 PMID: 16765269

40. Rowlands AV. Accelerometer assessment of physical activity in children: an update. Pediatr Exerc Sci. 2007; 19:252–266. PMID: 18019585

41. Santos-Lozano A, Santín-Medeiros F, Cardon G, Torres-Luque G, Bailón R, Bergmeir C, Ruiz JR, Lucia A, Garatachea N. Actigraph GT3X: validation and determination of physical activity intensity cut points. Int J Sports Med. 2013 Nov; 34(11):975–82. doi: 10.1055/s-0033-1337945 PMID: 23700330

42. Jimmy G, Seiler R, Mäder U. Development and validation of GT3X accelerometer cut-off points in 5-to 9-year-old children based on indirect calorimetry measurements. Swiss Sport Exerc Med. 2013; 61:37–43.

43. Romanzini M, Petroski EL, Ohara D, Dourado AC, Reichert FF. Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. Eur J Sport Sci. 2014; 14:91–99. doi: 10.1080/17461391.2012.732614 PMID: 24533499

44. Ried-Larsen M, Brønd JC, Brage S, Hansen BH, Grydeland M, Andersen LB, et al. Mechanical and free living comparisons of four generations of the Actigraph activity monitor. Int J Behav Nutr Phy. 2012; 9:113.

45. Robusto KM, Trost SG. Comparison of three generations of ActiGraph activity monitors in children and adolescents. J Sport Sci. 2012; 30:1429–1435.

46. Hänggi JM, Phillips LR, Rowlands AV. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. J Sci Med Sport. 2013; 16:40–44. doi: 10.1016/j.jsams.2012.05.012 PMID: 22749938

47. Mackintosh KA, Fairclough SJ, Stratton G, Ridgers ND. A calibration protocol for population-specific accelerometer cut-points in children. PLoS ONE. 2012; 7:e36919. doi: 10.1371/journal.pone.0036919 PMID: 22590635

48. Pulsford RM, Cortina-Borja BM, Rich C, Kinnafick FE, Dezateux C, Griffiths LJ. Actigraph accelerometer-defined boundaries for sedentary behaviour and physical activity intensities in 7 year old children. PLoS One. 2011; 6:e21882.

49. Spruijt-Metz D, Belcher B, Anderson D, Lane CJ, Chou CP, Salter-Venzon D, et al. A high-sugar/low-fiber meal compared with a low-sugar/high-fiber meal leads to higher leptin and physical activity levels in overweight Latina females. J Am Diet Assoc. 2009; 109:1058–1063. doi: 10.1016/j.jada.2009.03.013 PMID: 19465188

50. De Decker E, De Craemer M, Santos-Lozano A, Van Cauwenberghe E, De Bourdeaudhuij I, Cardon G. Validity of the ActivPALTM and the ActiGraph monitors in pre-schoolers. Med Sci Sport Exer. 2013; 45:2002–2011.

51. Lante K, Reece J, Walkley J. Energy expended by adults with and without intellectual disabilities during activities of daily living. Res Dev Disabil. 2010; 31:1380–1389. doi: 10.1016/j.ridd.2010.06.022 PMID: 20685074

52. Bassett DR, Rowlands A, Trost SG. Calibration and validation of wearable monitors. Med Sci Sport Exer. 2012; 44:S32–S38.

53. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med Sci Sport Exer. 2005; 37:S501–S511.

54. Staudenmayer J, Zhu W, Catellier DJ. Statistical considerations in the analysis of accelerometry-based activity monitor data. Med Sci Sport Exer. 2012; 44:S61–S67.