Cross validation of Actigraph derived accelerometer cut-points for assessment of sedentary behaviour and physical activity in children aged 8-11 years

Michael J. Duncan[a]*, Emma L. J. Eyre[a], Val Cox[a], Clare M. P. Roscoe[b], Mark A. Faghy[b], Jason Tallis[a], Alexandra Dobell[a]

Short title: cross-validation of actigraph cut-points

[a]School of Life Sciences, Coventry University, Coventry UK

[b] School of Human Sciences, University of Derby, Derby, UK

*Address for correspondence

Prof Michael J. Duncan

School of Life Sciences

Coventry University

Coventry, UK

E: michael.duncan@coventry.ac.uk

Abstract

Aim: To cross-validate previously calibrated Actigraph cut points in children. **Methods:** Twenty eight children (50%boys) aged between 8 and 11 years of age (9.4 ± 1.4 years) performed a series of 5 minute bouts of activity reflective of different levels of PA from sedentary behaviour (SB) to moderate to vigorous physical activity (MVPA); $\dot{V}O_2$ was assessed using breath by breath indirect calorimetry and activity was assessed using Actigraph accelerometers worn on the hip and non-dominant wrist. The $\dot{V}O_2$ values were then converted into age-specific METs (measured METs) and coded as SB, light PA and MVPA. Accelerometer data was analysed using previously calibrated cut-points at different epochs i.e. 5, 15 30 and 60 seconds. **Results:** Receiver Operating Characteristic (ROC) curve analysis indicated that there was excellent discrimination of SB using the Evenson et al (15sec), Romanzini (15sec), Treuth et al (30sec), Freedson et al (60sec), Treuth et al (60sec) and Troiano et al (60sec) cut points. ROC analysis indicated poor discrimination for LPA irrespective of the cut-points used. Good discrimination of MVPA was evident for all existing cut-points using a 60sec epoch. **Conclusion:** There is considerable variation in the performance of existing cut-points for assessment of SB, LPA and MPA in children.

Keywords: Cut-Points; Indirect Calorimetry; Energy Expenditure; Motor Skills

Key Notes

- Accelerometer cut-points derived to assess physical activity and sedentary behavior in children have rarely been cross-validated.
- We cross-validated existing hip and wrist based cut-points for the Actigraph accelerometer in a sample of 8-11 year old children.
- In this population, there is considerable variation in the performance of existing Actigraph cut-points for assessment of sedentary behaviour, light physical activity and moderate intensity physical activity in children.

Introduction

Assessing physical activity (PA) is key in population monitoring of children's health status as well as to determine the efficacy of exercise related interventions. Accelerometery, as a tool to assess PA is becoming widespread in children (1), because it is suggested to be more reliable and valid than alternative methods such as self-report (2). As accelerometry provides a device-based assessment of time spent in PA and sedentary behaviour (SB) (3), there have been considerable efforts made to derive accelerometer cut-points to more accurately estimate children's PA (4, 5, 6).

Developments in the field of PA assessment over the last decade has however resulted in the publication of numerous cut-points which are derived over varying epoch lengths. This has resulted in a lack of clarity for research and practitioners involved in child health to compare across studies. A recent systematic review (7) concluded that there was a need for researchers to employ age and criteria specific cut-points to better quantify PA. This is particularly important for children where the nature of PA tends to be sporadic and multidirectional (2). Stature may also have an influence in device-based assessment of PA given that taller stature, and subsequent longer levers will result in different volume of acceleration for the same intensity of PA as an individual with shorter stature. Specific accelerometer cut-points, either using the vertical axis (VA) or vector magnitude (VM), derived from the widely used Actigraph accelerometer (e.g., 8, 9, 10), using different placement locations (wrist vs hip), epoch lengths (from 5-60secs) and calibration protocols have been presented in the literature. This can cause confusion during PA assessment as there are differences between the cut-points suggested in these studies and the majority of existing cut-points have not been cross validated (7). Cross validation using an independent sample, which the cut-points are initially derived from is key to establishing cut-point accuracy (10). The current study sought to address this issue by cross-validating previously calibrated Actigraph cut points in children aged 8-11 years.

Methods

Participants

Twenty eight healthy, Caucasian, children (50% boys) aged between 8 and 11 years of age (9.4 ± 1.4 years) took part in this study following institutional ethics approval, parental written informed consent and child assent. Mean ± SD of height, mass and body mass index (BMI), was 140.5 ± 10.3cm, 34.6 ± 8.6 kg and 17.6 ± 2.5 kg/m^2 respectively.

Procedures

All testing took place in the morning (9am-12pm). Two triaxial accelerometers (Actigraph GT3X, ActiGraph inc, Pensacola, Florida, USA) were fitted to each participant at the non dominant wrist and non dominant hip prior to children enaging in a series of activities with different levels of PA. These activities comprised: lying supine, playing with Lego whilst seated, slow and medium paced walking, a medium paced run, overarm throwing and catching (a standard size tennis ball), passing a size 3 football with the instep and cycling (Lode Corival Paediatric, Lode BV, Groningen, Netherlands). These were performed as per prior work by Duncan et al (6) and in order as per prior work by Phillips et al. (5). All participants were fully familiarised with the treadmill used (Woodway Inc, Wisconsin, USA) prior to data collection. Each activity was performed for 5 minutes followed by a 5-minute rest in between. Previous protocols (6, 8) were used as guidelines where walking and running speeds were set at $3kmph^{-1}$, $4.5kmph^{-1}$, and $6.5kmph^{-1}$ to represent slow, medium pace walking and running respectively. Cadence for overarm throwing and catching and passing a football was based on those reported by previous authors (6) to ensure one complete action (e.g. a throw) was completed every 3 seconds (6). Throughout the testing procedure, breath by breath gas analysis (MetaMax 3B, Cortex Biophysik GmbH, Leipzig, Germany) was used to assess $\dot{V}O_2$ and $\dot{V}CO_2$. Children wore a junior face mask (Hans Rudolph, Kansas, USA) and the MetaMax was calibrated prior to commencing testing with gases of known concentration. The MetaMax system has previously been established as a valid and reliable system for the assessment of ventilator parameters at rest and during exercise in children (11, 12, 13), including children with low body mass, as is the case in the current study, where low ventilation rate can cause issues when assessing respiratory values using metabolic carts. Prior to testing of each participant accelerometers and Metamax were synchronised to Greenwich Mean Time enabling accurate synching of accelerometer and breath by breath data for each activity, as per prior research (5). During analysis manual (visual) checks were used to confirm the synch given the 5-minute gap between each activity which provided a clear indication of start/stop of each of the different activities.

Data processing

Once the exercise protocol was completed and accelerometer and indirect calorimetry data had been saved to computer, the first and last minute of each bout were discarded leaving a 3-minute period for analysis. This is in accordance with prior studies (5, 6) to ensure that MET values for each activity bout were at the required intensity. This process ensured that data were at steady state (14) and, although discarding only the first minute especially at high intensities may seem short, there were no differences in $\dot{V}O_2$ between second and third minutes of analysis. when checked there was Raw accelerometer data was recorded at a frequency of 100Hz, as recommended by Migueles et al. (7). Data from accelerometers were measured in counts, a variable that is linearly related to the amount of movement and acceleration during an epoch (7). Data were saved and analysed using both the summed vector magnitude (VM) and vertical axis (VA) counts and at different epoch lengths using the Actigraph post processing software (Actilife Version 6, ActiGraph Inc Pensacola, FL, USA) to match previously validated accelerometer cut-points for children aged 8-11 years old, as identified by Migueles et al. (7). Low frequency extension function was enabled in the analysis of Actigrpah data files.

Data was downloaded for 5-second epochs only for the wrist using both the summed vector magnitude (VM) and vertical axis (VA) counts in line with the Chandler et al (15) cut points. The data for the hip was downloaded with just the VA counts and epochs of 15, 30 and 60 seconds in line with cut points previously validated by Puyau et al. (8), Evenson et al, (9) Romanzini et al (16) (15sec), Treuth et al (17) (30sec), Andersen et al (18), Freedsen et al (10), Treuth et al (17), Troiano et al (19) and Vanhelst et al (20) (60sec).

To determine METs we followed procedures previously employed by Harrell et al (21) where we divided the measured $\dot{V}O_2$ of the observed activity for each participant by that participant's resting $\dot{V}O_2$ to provide age-specific MET values (21). The MET values were subsequntly coded into one of four intensity categories (sedentary < 1.5 METs), light (1.5-2.99 METs), moderate (3-5.99 METs) and Vigorous (>6 METs). On inspection, none of the activities undertaken by the participants resulted in MET values that were classed as vigorous in intensity (>6 METs). Data were then subsequently recoded into 3 intensity categories reflecting sedentary (SB), light (LPA) and moderate PA (MPA).

Statistical Analysis

Accuracy of assessment of SB, light PA and MVPA using previously validated Actigraph cut-points against breath by breath derived METs values was determined using Receiver Operating Characteristic (ROC) curve analysis was undertaken (22). The accelerometer counts were coded into sedentary, light PA and MPA using previously mentioned validated cut-points for the wrist (11) and hip (8, 9, 12 (15sec), Treuth et al., (17) (30sec), Andersen et al., (18), Freedsen et al., (10), Treuth et al (17), Troiano et al., (19) and Vanhelst et al, (20) (60sec)). Counts were coded into binary indicator variables (0 or 1) based on intensity (sedentary versus >sedentary, less than moderate versus moderate and vigorous) in order for ROC curve analysis to be conducted as described previously (22). We sought to compare how well the existing cut points for children could classify the intensity of the activities compared to the intensity determined by indirect calorimetry and thus provide cross validation of their cut-points. Two sets of values were determined for each set of existing cut-points, one including cycling and one without cycling, similar to previous work using the GENEActiv accelerometer (6). The area under the curve (AUC) was calculated for each analysis as a measure of diagnostic accuracy with AUC values of; ≥ 0.90 considered excellent, 0.80–0.89 good, 0.70–0.79 fair, and < 0.70 poor as per other research (4, 5, 6). ROC curve analysis was conducted as described previously and calculation of sensitivity (Se) and specificity (Sp) was also performed (23). McNemar's tests for paired proportions were then used to examine any differences in the sensitivity and specificity between sets of cut-points at each epoch. The Statistical Package for Social Sciences (SPSS, version 24) was used for analysis. In order to examine any differences between ROC curves at each epoch the De Long test was employed (24) with analysis performed in R (25) using the pROC package (26).

Results

Results from ROC analysis including AUC and sensitivity and specificity values are presented in Table 1. These results indicate that existing cut-points were most effective in discriminating SB with excellent discrimination demonstrated for the Evenson et al (9) (15sec epoch), Romanzini (16) (15sec epoch), Treuth et al (17)

(30sec epoch), Freedson et al (10) (60sec epoch), Treuth et al (17) (60sec epoch) and Troiano et al (19) (60sec epoch) cut points when cycling was not included in the data analysis. Excellent discrimination of SB was also evident for the Treuth et al (17) (60sec epoch) and Troiano et al (19) (60 sec epoch) cut-points when cycling was included. When LPA was considered no cut-points performed particularly well. ROC analysis indicated poor discrimination for LPA irrespective of the cut-points used or whether cycling was or was not included in the analysis. For MPA there was excellent discrimination for the Treuth et al (17) (30sec epoch) and Puyau et al (8) (30sec epoch) when cycling was not included. Good discrimination of MPA was evident for all existing cut-points using a 60sec epoch irrespective of whether cycling was included in the analysis or not. Likewise, the Treuth et al (17) (30sec cycling included), Romanzini et al (16) (15sec epoch without cycling), Evenson et al (9) (15sec epoch without cycling), Romanzini et al (16) (15sec epoch with cycling), Evenson et al (9) (15sec epoch with cycling) and Puyau et al (8) (15 sec epoch with cycling) cut-points all had AUC values that indicated good discrimination for MPA. The Chandler et al (11) (5 sec epoch for VA and VM) both represented poor discrimination for MPA. For SB Chandler (11) VM data presented higher specificity than sensitivity indicating that the Chandler SB cut-points were better at capturing data that was not sedentary as such but the lower sensitivity would be indicative of a higher false negative rate. Such data need to be considered alongside the AUC values reported. The Treuth et al (30sec) cut-points (17) (without cycling included) which had the highest AUC values for all three of SB, LPA and MPA as well as the highest values for sensitivity and specificity.

Table 1 Here

Results from McNemar's tests indicated significant differences (all P>.01) between cut-points values at every epoch with the exception of the Treuth (17) and Troiano (19) cut-points for SB, Andersen (18) and Van Helst (20) and Troiano (19) and Van Helst (20) cut-points for SB at 60s. The outcomes from McNemar's tests remained the same irrespective of whether cycling was or was not included in the analysis.

When comparing ROC curves, results from De Long's tests indicated significant differences (all P>.001) between all cut-points at every epoch with the exception of the following: Andersen (18) and Freedson (10), Andersen (18) and Van Helst (20), Freedson (10) and Van Helst (20) and Treuth (17) and Troiano (19) cut-points for LPA at 60s epoch. There was also no significant difference in ROC curves for MPA at 60s epochs between the Andersen and Troiano (19), Andersen (18) and Van Helst (20), Freedson (10) and Van Helst (20) and Troiano (19) and Van Helst (20) cut points. Nor was there any significant difference between Chandler (11) VA and VM cut-points for MPA at 5s ecpoch. Similar to the McNemar tests, the outcomes of the analysis remained the same irrespective of cycling being included in the analysis or not.

Discussion

The current study cross validates a range of previously calibrated Actigraph cut points, using epochs from 5-60secs, in children aged 8-11 years. Such information is needed to aid researchers in making decisions as to which accelerometer cut-points may be best suited in assessing SB, LPA and MPA in pediatric populations. To date, although multiple accelerometer cut-points have been developed, few cross validation studies have been performed. The current study, therefore, provides independent, cross-validation data on a range of existing Actigraph based cut-points used to assess PA in children. However, the results of the present study highlight some of the difficulties associated with use of particular cut-points to assess SB, LPA and MPA in children.

While the cut-points may present ability to discriminate the different intensities of activity in children, results from McNemar and De Long's tests demosntarte considerable differences between the different cut-points at the same epoch. The result sof the present study therefore agree with assertions made by Migueles et al (7) that existing cut-point based methods need careful consideration when using device-based assessment of movement behaviours in children.

The inclusion of cycling with accelerometer calibration protocols has been a point of debate. Cycling activity results in minimal movement at the hip and wrist, compared to other more ambulatory activities which often results in misclassification of cycling activity by accelerometers worn at these locations (27). However, cycling is recognised as an important health enhancing activity (27) and, as such it is important that accelerometer cut points can classify this activity. For this reason, in the present study, data cross-validating existing actigraph cut-points with and without cycling were included. When cycling data was included in the analysis, discrimination of activity intensity was poorer and sensitivity and specificity values lower than when cycling data was not included in the analysis, similar to prior work (6, 27). The inclusion of activities reflective of different fundamental movement skills within the exercise protocol is also important to note.

The data presented here are based on activities conducted in a laboratory setting. Such laboratory based data are needed to cross validate cut-points in a manner equivalent to how the cut-points were initially calibrated. While the equipment used in the current study can be employed in outdoor settings, it was not logistically possible for us to do this with additional activities in addition to the current protocol. A useful next step for researchers is to examine the validity of these cut-points during free living PA in children. In the current study, participants were children who engaged with grassroots football and the results presented here are therefore indicative of children who had ‘good’ motor competence and were all within ‘healthy’ BMI based weight status categories.

This study extends the literature in the area of PA assessment by cross validating existing Actigraph accelerometer cut-points in children aged 8-11 years of age. The results of the current study suggest that, in this population, there is considerable variation in the performance of existing cut-points for assessment of SB, LPA and MPA in children. Researchers therefore need to be mindful of such variation when making decisions relating to accelerometer cut-point use in pediatric populations. The results presented here also highlight the limitations of device-based assessment of movement using a cut-point approach.

Conflicts of Interest: None

Funding: None

References

1. Vale S, Trost SG, Duncan MJ, Mota J. Step based physical activity guidelines for preschool children. *Prev Med* 2015;70:78–82.
2. Rowlands AV, Eston RG. The measurement and interpretation of children’s physical activity. *J Sports Sci* 2007;6:270-273.
3. Crouter SE, Flynn Oody J, Bassett Jr DR. Estimating physical activity in youth using an ankle accelerometer. *J. Sports Sci* 2018;36:2265-2271.

4. Duncan MJ, Wilson S, Tallis J, Eyre E. Validation of the Phillips et al. GENEActiv accelerometer wrist cut-points in children aged 5–8 years old. *Eur J Pediatr* 2016;175:2019-2021.
5. Phillips LR, Parfitt G, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport* 2014;16:124-128.
6. Duncan MJ, Roscoe CMP, Faghy M, Tallis J, Eyre ELJ. Estimating Physical Activity in Children Aged 8-11 Years Using Accelerometry: Contributions From Fundamental Movement Skills and Different Accelerometer Placements. *Front Physiol* 2019;10:242.
7. Migueles JH, Cadenas-Sanchez C, Ekelund U, Nystrom CD, Mora-Gonzalez J, Lof M, Labayen I, Ruiz JR, Ortega FB. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports Med* 2017;47:1821-1845.
8. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and Calibration of Physical Activity Monitors in Children. *Obes Res* 2002;10:150-157.
9. 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: 1557-1565.
10. Freedson P, Pober D, Janz KF. Calibration of Accelerometer Output for Children. *Med Sci Sports Exerc* 2005;37:S523-S530.
11. Medbo JI, Mamen A, Welde B, von Heimburg E, Stokke R. Examination of the Metamax I and II oxygen analysers during exercise studies in the laboratory. *Scand J Clin Lab Invest* 2002;62:585–98.

12. Macfarlane DJ, Wong P. Validity, reliability and stability of the portable Cortex Metamax 3B gas analysis system. Eur J Appl Physiol 2012;112:2539-2547.
13. Paap E, van der Net J, Helders PJ, Takken T. Physiologic response of the six-minute walk test in children with juvenile idiopathic arthritis. Arthritis Rheum 2005; 13: 351-356.
14. Rowland TW. *Developmental Exercise Physiology*. Champaign, IL: Human Kinetics; 1995.
15. Chandler JL, Brazendale K, Beets MW, Mealing BA. Calibration of two objective measures of physical activity for children. *Pediat Obes* 2015;11:120-127.
16. Romanzini M, Petroski EL, Ohara D, et al. Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *Eur J Sport Sci* 2014;14:91–99.
17. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc* 2004;36:1259–66.
18. Andersen LB, Harro M, Sardinha LB, et al. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). *Sports Med* 2006;368:299–304.
19. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008;40:181–8.
20. Vanhelst J, Be´ghin L, Turck D, et al. New validated thresholds for various intensities of physical activity in adolescents using the ActiGraph accelerometer. *Int J Rehabil Res* 2011;34:175–7.
21. Harrell JS, McMurray RG, Baggett CD, Pennell ML, Pearce PF, Bangdiwala SI. Energy Costs of Physical Activities in Children and Adolescents. *Med Sci Sports Exerc* 2005;37:329-336.

22. Jago R, Zakeri I, Baranowski T, Watson K. Decision boundaries and receiver operating characteristic curves: new methods for determining accelerometer cutpoints. *J Sports Sci* 2007;25:937-944.

23. Esliger DW, Rowlands AV, Hurst Tl, Catt M, Murray P, Eston RG. Validation of the GENEA Accelerometer. *Med Sci Sports Exerc* 2011;43:1085-1093.

24. DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988, 44 (3): 837-845.

25. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. 2013.

26. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Muller M. pROC : an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12;77.

27. Welch W, Bassett DR, Thompson DL, Freedson PS, Staudenmayer JW, John D, Steeves JA, Conger SA. Classification accuracy of the wrist-worn GENEA accelerometer. *Med Sci Sports Exerc* 2013;45:2012–2019.

Table 1. Area under the curve (AUC), sensitivity and specificity values for existing accelerometer cut-points at different epochs with and without cycling included in the analysis.

		AUC	Sensitivity	Specificity
5 sec epoch VA without cycling				
Chandler (11) (Wrist)	SED	0.726	0.879	0.501
	LPA	0.622	0.832	0.428
	MVPA	0.654	0.938	0.336
5 sec epoch VM without cycling				
Chandler (11)(Wrist)	SED	0.712	0.594	0.727
	LPA	0.614	0.635	0.601
	MVPA	0.666	0.967	0.347
15 sec epoch with cycling				
Puyau (8)(Hip)	SED	0.785	0.968	0.545
	LPA	0.634	0.953	0.443
	MVPA	0.861	0.888	0.712
Evenson (9)(Hip)	SED	0.877	0.915	0.787
	LPA	0.634	0.951	0.445
	MVPA	0.820	0.939	0.617
Romanzini (16)(Hip)	SED	0.850	0.921	0.739
	LPA	0.626	0.953	0.442
	MVPA	0.830	0.964	0.613
15 sec epoch without cycling				
Puyau (8)(Hip)	SED	0.849	0.967	0.610
	LPA	0.671	0.950	0.485
	MVPA	0.906	0.888	0.784
Evenson (9)(Hip)	SED	0.940	0.908	0.860
	LPA	0.643	0.945	0.471
	MVPA	0.873	0.938	0.683
Romanzini (16)(Hip)	SED	0.925	0.915	0.824
	LPA	0.647	0.948	0.476
	MVPA	0.882	0.964	0.678
30 sec epoch with cycling				
Treuth (17)(Hip)	SED	0.886	0.914	0.797
	LPA	0.666	0.954	0.509
	MVPA	0.860	0.999	0.589
30 sec epoch without cycling				
Treuth (17)(Hip)	SED	0.944	0.907	0.866
	LPA	0.681	0.948	0.539
	MVPA	0.905	0.999	0.648
60 sec epoch with cycling				

Andersen (18)(Hip)	SED	0.805	0.928	0.603
	LPA	0.603	0.982	0.367
	MVPA	0.809	0.883	0.649
Freedsen (10)(Hip)	SED	0.877	0.971	0.711
	LPA	0.584	0.826	0.442
	MVPA	0.805	0.928	0.603
Treuth (17)(Hip)	SED	0.903	0.901	0.819
	LPA	0.684	0.960	0.529
	MVPA	0.854	0.968	0.611
Troiano (19)(Hip)	SED	0.903	0.901	0.819
	LPA	0.640	0.952	0.444
	MVPA	0.808	0.882	0.648
Van Helst (20)(Hip)	SED	0.813	0.918	0.633
	LPA	0.603	0.995	0.355
	MVPA	0.806	0.873	0.649
		60 sec epoch without cycling		
Andersen (18)(Hip)	SED	0.877	0.924	0.675
	LPA	0.623	1	0.375
	MVPA	0.867	0.883	0.726
Freedsen (10)(Hip)	SED	0.938	0.892	0.852
	LPA	0.554	0.792	0.474
	MVPA	0.877	0.924	0.675
Treuth (17)(Hip)	SED	0.948	0.891	0.872
	LPA	0.691	0.955	0.552
	MVPA	0.899	0.968	0.674
Troiano (15)(Hip)	SED	0.948	0.891	0.872
	LPA	0.629	0.944	0.461
	MVPA	0.866	0.882	0.724
Van Helst (20)(Hip)	SED	0.889	0.985	0.638
	LPA	0.623	0.970	0.408
	MVPA	0.865	0.873	0.726