Citation for published version:
Barreira, TV, Schuna, JM, Tudor-Locke, C, Chaput, J-P, Church, TS, Fogelholm, M, Hu, G, Kuriyan, R, Kurpad, A, Lambert, EV, Maher, C, Maia, J, Matsudo, V, Olds, T, Onywera, V, Sarmiento, OL, Standage, M, Tremblay, MS, Zhao, P & Katzmarzyk, PT 2015, 'Reliability of accelerometer-determined physical activity and sedentary behavior in school aged children: a 12 country study', International Journal of Obesity, vol. 5, no. S2, pp. S29-S35. https://doi.org/10.1038/ijosup.2015.16

DOI:
10.1038/ijosup.2015.16

Publication date:
2015

Document Version
Peer reviewed version

Original Article

Reliability of accelerometer-determined physical activity and sedentary behavior in school aged children: A 12- country study

Tiago V. Barreira[1,2], John M. Schuna, Jr.[2,3], Catrine Tudor-Locke[2,4], Jean-Philippe Chaput[5], Timothy S. Church[2], Mikael Fogelholm[6], Gang Hu[2], Rebecca Kuriyan[7], Anura Kurpad[7], Estelle V. Lambert[8], Carol Maher[9], José Maia[10], Victor Matsudo[11], Timothy Olds[9], Vincent Onywera[12], Olga L. Sarmiento[13], Martyn Standage[14], Mark S. Tremblay[5], Pei Zhao[15], and Peter T. Katzmarzyk[2], for the ISCOLE Research Group

[1]Department of Exercise Science, University of Syracuse, Syracuse, USA; [2]Pennington Biomedical Research Center, Baton Rouge, USA; [3]Oregon State University, Corvallis, USA; [4]Department of Kinesiology, University of Massachusetts Amherst, Amherst, USA [5]Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada; [6]Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland; [7]St. Johns Research Institute, Bangalore, India; [8]Division of Exercise Science and Sports Medicine, Department of Human Biology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa; [9]Alliance for Research in Exercise Nutrition and Activity (ARENA), School of Health Sciences, University of South Australia, Adelaide, Australia; [10]CIFI[2]D, Faculdade de Desporto, University of Porto, Porto, Portugal; [11]Centro de Estudos do Laboratório de Aptidão Física de São Caetano do Sul (CELAFISCS), Sao Paulo, Brazil; [12]Department of Recreation Management and Exercise Science, Kenyatta University, Nairobi, Kenya; [13]School of Medicine, Universidad de los Andes, Bogota, Colombia; [14]Department for Health, University of Bath, Bath, United Kingdom; [15]Tianjin Women's and Children's Health Center, Tianjin, China

Running Head: Reliability of accelerometer-derived metrics

Address correspondence to: Tiago V Barreira, PhD, Syracuse University, School of Education, 820 Comstock Ave , Women's Building Room 207, Syracuse, New York 13244; Telephone: 315.443.5588; Fax: 315.443.9375; Email: tvbarrei@syr.edu

Abstract

Objective: Focused on accelerometer-determined physical activity and sedentary time metrics in 9-11 year old children, we sought to determine: (1) the number of days that are necessary to achieve reliable estimates (G≥0.8); (2) the proportion of variance attributed to different facets (Participants and Days) of reliability estimates; and (3) the actual reliability of data as collected in The International Study of Childhood Obesity, Lifestyle and Environment (ISCOLE).

Subjects/Methods: The analytical sample consisted of 6025 children (55% girls) from sites in 12 countries. Physical activity and sedentary time metrics measures were assessed for up to 7 consecutive days for 24 h/day with a waist-worn ActiGraph GT3X+. Generalizability theory using R software was used to investigate objectives 1 and 2. Intra-class correlation coefficients (ICC) were computed using SAS PROC GLM to inform objective 3.

Results: The estimated minimum number of days required to achieve a reliability estimate of G≥0.8 ranged from 5-9 for boys and 3-11 for girls for LPA; 5-9 and 3-10 respectively for MVPA; 5-10 and 4-10 for total activity counts; and 7-11 and 6-11 for sedentary time. For all variables investigated, the Participant facet accounted for 30-50% of the variability whereas Days accounted for ≤5%, and the interaction (PxD) accounted for 50-70% of the variability. The actual reliability for boys in ISCOLE ranged from ICCs of 0.78-0.86, 0.73-0.85, and 0.72-0.86 for LPA, MVPA, and total activity counts, respectively, and 0.67-0.79 for sedentary time. The corresponding values for girls were 0.80-0.88, 0.70-0.89, 0.74-0.86, and 0.64-0.80.

Conclusion: It was rare that only 4 days from all participants would be enough to achieve desirable reliability estimates. However, asking participants to wear the device for 7 days and requiring ≥4 days of data to include the participant in the analysis might be an appropriate approach to achieve reliable estimates for most accelerometer-derived metrics.

Key Words: accelerometry, stability, repeatability, validity

Trial Registration: ClinicalTrials.gov: Identifier NCT01722500

Introduction

It is imperative that measurements used in research are both reliable and valid. The degree of reliability represents the stability of the measured value and this will influence the strength of relationships observed between variables and the ability to detect changes in a measured outcome.[1] Accelerometers are now widely used and becoming the standard objective measure of physical activity and sedentary behavior.[2,3] For this reason, it is important for physical activity and sedentary behavior researchers to evaluate both the reliability (i.e., repeatability) of the instrument *per se* and the reliability (i.e., stability) of the measured behavior (physical activity, sedentary behavior, etc.).[4-6] When tested in mechanical shakers, it has been demonstrated that accelerometers can output repeatable and reliable measures.[6-9] The stability of behaviors is commonly evaluated by determining how many days are necessary to achieve a desired reliability level for the variable being investigated.[4,10-13]

Based on prior pediatric accelerometer- and pedometer-based studies, it is currently common practice to ask children to wear accelerometers for 7 consecutive days and then to only use data from those that provide at least 3 or 4 days of valid data.[14-16] We employed a similar strategy in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE), by implementing an *a priori* decision rule to only include participants with ≥4 days of accelerometer data, including at least 1 weekend day.[17] However, there is limited evidence that this protocol is actually sufficient to achieve the commonly desired intra-class correlation coefficient (ICC) of ≥ 0.8. While an ICC of 0.8 is not an absolute criterion, it is commonly used in physical activity research [13,18,19] and it is enough to decrease correlation between variables by 10%.[1] Reliability estimates have been previously derived from studies conducted with small sample sizes, wide age ranges, and sex-combined (boys and girls) results. However, some studies have provided a thorough investigation of the reliability of accelerometer-derived physical activity in youth. For example, Trost et al.[19] tested a total of 381 children and adolescents and concluded that 4-5 days were necessary to achieve ICC ≥ 0.8 for children's accelerometer-derived MVPA measurement while 8-9 days were necessary for adolescents. In the same study, the ICC for 4 days ranged from 0.64 to 0.79 for the different groups. Hinkley et al.[13] investigated the number of days required to achieve different levels of reliability for percentage of time at MVPA in a sample of 799 preschool children, taking into consideration different numbers of hours per day to be included in the analysis and determined that > 4 days were necessary for ICC ≥ 0.8 when 10 h/day was used. In a more recent investigation by Wickel,[20] detailed reliability information for accelerometer-derived MVPA was presented for both complete data (only those with 7 days) and incomplete data (1-7 days of data) using different wear time standards in a sample of 1082 children. For the most commonly accepted wear time definition of a valid day (≥10 h/day), reliability coefficients ranged from 0.77 to 0.84 for complete data and 0.54 to 0.65 for incomplete data. Two of these studies[19,20] suggest that > 4 days might be necessary to achieve target reliability in estimates of MVPA in youth. However, the aforementioned studies did not investigate the reliability of other physical activity metrics, including time spent in different physical activity intensities and sedentary time. In one of the few studies that investigated sedentary behavior, Basterfield et al.[21] studied 291, 6- to 8- years-old children and demonstrated that ≥ 5 days were necessary to achieve the desirable level of reliability for estimating percentage of time spent in sedentary behavior.

ISCOLE was conducted at 12 different sites in countries around the globe representing a wide range of cultures and levels of human development.[17] Thus, ISCOLE provides a unique opportunity to test the reliability of accelerometer-determined activities in a large and culturally diverse sample of children. Focused on accelerometer-determined physical activity and sedentary time metrics in 9-11 year old children, we sought to determine: (1) the number of days that are necessary to achieve reliable estimates (G≥0.8); (2) the proportion of variance attributed to different facets (Participants and Days) of reliability estimates; and (3) the actual reliability of accelerometry data as collected in ISCOLE.

Subjects and Methods

Detailed information about ISCOLE's design, methods and accelerometry procedures have been previously published, including open access publication of a detailed Manual of Operations[17,22] For this reason, only those procedures directly related to this study are presented here. The Institutional Review Board at the Pennington Biomedical Research Center (coordinating center) approved the overarching ISCOLE protocol, and approval was also obtained by each institution with their respective Institutional/Ethical Review Boards. Written informed consent was obtained from parents or legal guardians, and child assent was also obtained as required by local Institutional/Ethical Review Boards before participation in the study.

Study Sample

From the original sample of 7372 participants, 31 were excluded because they did not have BMI data and further 793 did not have valid accelerometer data (standard described below). The final sample included in this analysis consisted of 6548 children aged 9-11 years (55% girls) from sites in 12 countries (Australia, Brazil, Canada, China, Colombia, Finland, India, Kenya, Portugal, South Africa, United Kingdom, and United States). Recruitment was conducted with students nested within schools that were nested within country sites with a goal to enroll a sex-balanced sample of at least 500 children per site. Data was collected when children were attending school and excluded major holidays.

Accelerometer-Derived Activities

Participants were asked to wear an ActiGraph GT3X+ accelerometer (ActiGraph LLC, Pensacola, FL, USA) at the waist on an elasticized belt, and positioned in-line with the right mid-axillary line, for at least 7 consecutive days (plus an initial familiarization day and the morning of the final day). Children were asked to wear the accelerometer 24 h/day (removing only for water-related activities). The accelerometer assessment was conducted when school was in session. Data were collected at a sampling frequency of 80 Hz, and subsequently downloaded using ActiLife Software (version 5.64 or later, ActiGraph LLC). Raw accelerometer data were integrated into 1 s epochs and later re-integrated into 60 and 15 s epochs with the low-frequency extension filter enabled. Since the accelerometer was worn for 24 h/day, it was necessary to identify nocturnal sleep episode time distinct from waking non-wear time, and this was done using a 60 sec epoch and published automated algorithms.[23,24] After exclusion of the nocturnal sleep episode time, non-wear time was determined as any sequence of at least 20 consecutive min of zero activity counts.[25] Once nocturnal sleep episode time and non-wear time were computed, waking wear time and the different activity levels and sedentary time were calculated and identified using the 15 s epoch data. Children were only included in this analysis if they had ≥4 days of monitoring with at least 10 h/day of waking wear time, including at least 1 weekend day. Time spent in light physical activity (LPA), MVPA, and sedentary time were estimated using the Evenson cut-points.[26] In addition, total activity counts were calculated for valid wake wear time.

Statistical Analysis

We performed reliability analyses following the generalizability theory framework using R statistical software. Generalizability theory is an extension of intra-class reliability and ANOVA which is typically divided into two parts, the G-study and the D-study.[10] The G-study is used to quantify the proportion of variance associated with each facet and its interactions. For the G-study, participant (P) and day (D) were considered random facets in a fully crossed design (P x D). Variance components corresponding to P, and D for accelerometer-determined physical activity and sedentary time metrics were estimated using restricted maximum likelihood with the "lme4" package's "lmer" function in R.[27] Using the aforementioned estimated variance components, we subsequently conducted a D-study to calculate Generalizability coefficients (G), which can be interpreted in the same manner that ICC values are interpreted and can be used to compute and extrapolate the minimum number of days required to achieve a reliability estimate of ≥ 0.8.[4,18] Descriptive statistics and ICC were calculated using SAS version 9.4 (SAS

Institute, Cary, NC, USA). A macro that uses PROC GLM was used to calculate ICC(2,1).[28] ICC(2,1) was chosen because the days that the participants were tested were a random selection.

Results

The number of participants from each site, average body mass index (BMI), and the average number of valid days of accelerometry are presented in Table 1. All but one site averaged ≥6 days of valid accelerometer data and only one site had < 200 participants for each sex.

The minimum theoretical number of days necessary to achieve a G ≥ 0.8 is presented in Figures 1-4. For LPA, the estimated number of days required to achieve desirable reliability ranged from 5-9 for boys and 3-11 for girls. For boys' LPA when compared to girls', an equal or higher number of days was required to achieve a G ≥ 0.8 in 9 of the sites. The estimated number of days required to achieve minimal reliability estimates for MVPA ranged from 5-9 for boys and 3-10 for girls. More days were required for boys versus girls in only 3 sites. For total activity counts, the results were very close to the MVPA results: the estimated minimum number of days ranged from 5-10 for boys and 4-10 for girls. For sedentary time, the estimated minimum number of days ranged from 7-11 for boys and 6-11 for girls.

The detailed variance results for boys and girls in each site are presented in Table 2. The Participant (P) facet accounted for 30-50% of the variability, Days (D) accounted for less than 5%, and the interaction (PxD) accounted for 50-70% of the variability. This demonstrates that a large percentage of observed variability was not explained by the Participant and Days facets.

ICC results are presented in Table 3. For boys, ICCs ranged from 0.78-0.86, 0.73-0.85, 0.72-0.86, and 0.67-0.79, for LPA, MVPA, total activity counts, and sedentary time, respectively. For girls, ICCs ranged from 0.80-0.88, 0.70-0.89, 0.74-0.86, and 0.64-0.80, for LPA, MVPA, total activity counts, and sedentary time, respectively. For both boys and girls combined, the ranges were similar and the mean ICCs were 0.83, 0.82, 0.82, and 0.75, for LPA, MVPA, total activity counts, and sedentary time, respectively.

Discussion

The generalizability theory results, specifically the D-study results, demonstrated that there is large variability between sites and between boys and girls in the estimated number of days necessary to achieve the desirable level of reliability for all intensities of physical activity and sedentary time. This was also true for the ICC results. This level of variability with different samples and different metrics has not been previously reported. In addition, it became clear that it is very rare that only 4 days of valid data from all participants would be required to achieve desirable levels of reliability. However, as was implemented in ISCOLE, asking participants to wear the device for 7 days and only including participants with ≥ 4 days of data in the analysis might produce a data set with on average ≥6 days of valid data and be an appropriate approach to achieve reliable estimates for most of the accelerometer-derived activity metrics.

For LPA, 4-8 days were necessary for the desirable reliability values in almost all the sites and only for girls in the Kenya site < 4 days was required (3 days). The G-study variance components for Participants and Days explained around 40% of the variability which still left room for a large amount of unexplained variance. The ICC values were almost all in the desirable level and were the highest among the physical activity measures investigated. This indicates that asking children to wear the devices for 7 days and requiring ≥ 4 days is acceptable to obtain reliable measures of LPA. We are unaware of other studies that have investigated the reliability of LPA measurement in children.

For both MVPA and total activity counts, the results were very similar. Once again in most cases it was estimated that > 4 days were required for a G ≥ 0.8 and as many as 10 days would be required to achieve this level of reliability. The interaction term or unexplained variances from the D-study were as high as 68%, leaving much of the variance unaccounted for. However, almost all ICC values were in the desirable range or very close to it. The ICC values were relatively high even though we included participants with as few as 4 days of valid data; however, the average number of days considered was closer to 7 days. These results are in line with the findings reported by Trost et al.[19], which concluded that 4-5 days were necessary to achieve ICC ≥ 0.8 for children's accelerometer-derived MVPA measurement while 8-9 days were necessary for adolescents. Our results also agree with the results from Hinkley et al.[13] that > 4 days are necessary to achieve an ICC ≥ 0.8 for percentage of time in MVPA in preschool children and the recent study by Wickel,[20] which reported reliability coefficients ranging from 0.77 to 0.84 for participants with 7 days of data and 0.54 to 0.65 for when including participants with 1-7 days of data. Both of those studies [19,20] suggested that ≥ 4 days might be necessary to achieve reliable estimates of MVPA in youth. Basterfield et al.[21] demonstrated that ≥ 6 days were necessary to achieve ≥ 80% reliability for total volume of physical activity and percentage time in MVPA.

For sedentary time specifically, the results from the D-study showed that the minimum amount of days necessary to achieve reliability coefficients ≥ 0.80 was 6 in only one site (and only for girls) and for all other sites it was ≥ 7 days. These results indicate that it is unlikely for researchers to realize high levels of reliability for sedentary time measurement with a period of only 7 days of data collection since it is rare to obtain complete valid data from all participants. In addition, the G-study results displayed large interaction term values indicating that a large percentage of the variability was not explained by the PxD model. The ICC reliability estimates realized for sedentary time fall in line with the generalizability results. An ICC ≥ 0.8 was only realized for girls assessed at one of the ISCOLE sites. This clearly demonstrates that sedentary time is not as reliably measured as physical activity. Although research in this area is scarce, Davies et al. [29] reported that ≥ 5 days were necessary for ICC ≥ 0.8 in a sample of 30 children. Similar results were also reported by Basterfield et al.[21] when estimating the reliability of percent time spent in sedentary behavior for 6-8 year-old children. In another study with 56 children, it was found that ICC values from week-to-week (not between days) ranged from 0.40 to 0.79

during week days and 0.25-0.60 during weekends for different sedentary behavior measures.[30] The need to assess more days to establish a reliable estimate of sedentary time when compared to other physical activity indicators could be because of the levels of day to day variability in sedentary time and/or the possible lack of accuracy of sedentary time measurement by waist-worn accelerometry.[31,32]

While this study was carefully conducted, it is not free of limitations. The samples are not representative of the countries from where they were drawn. Moreover, the samples in most cases were from just one city and sampling was conducted to maximize variability in socio-economic status.[17] Data was collected when children were attending school and it is possible that results might be different when they are not in school. We only tested the reliability of the different physical activity metrics and sedentary time using one set of cut points and therefore it is possible that other cut points might provide different results. However, we chose the cut point that appears to be the most valid for this age group.[33] We used the data for participants with ≥4 valid days (defined as ≥10 h of wake wear), and other *a priori* standards might produce different results as shown by Wickel.[20] However, this is a very common standard for inclusion criteria.[34] This study only investigated a narrow age range and the results cannot be generalized across all children and adolescent age groups as demonstrated in the study by Trost et al.[19] Lastly, we focused on the 0.8 standard for reliability, however 0.7 and 0.9 have also been used in physical activity research.[13,19]

This study highlights the variability in reliability estimates for accelerometer-derived variables. The same accelerometer and nearly identical protocol (slight variations in the compliance checks and incentives to wear were allowed as indicated by local customs) was used at all sites. However, a large amount of variability was observed in the number of days required to achieve a reliability estimate of $G \geq 0.8$. A large variability was also seen for the ICC values for all accelerometer-derived variables investigated. This study also highlights the importance of reporting reliability estimates achieved in the study even if *a priori* decisions guide analyses. The range in the ICC estimate between sites for the same variable was usually within 0.10 but was as large as 0.19. This difference can have a dramatic influence in correlate estimates[1] with other studied variables like BMI for example. It is possible that differences in the strength of the correlations between physical activity, sedentary time, and BMI across studies could be solely attributed to this difference in reliability estimates; however, this is not known because of lack of reliability reports.

Conclusion

In summary, we reported that there is a large amount of variability in the reliability estimates of different accelerometer-derived variables among different samples. In addition, it became clear that it is rare that only 4 days for all participants is required to achieve desirable levels of reliability estimates. However, asking participants to wear the accelerometer for 7 days and requiring ≥ 4 days of data to include the participants in the analysis might be an appropriate approach to achieve reliable estimates for most of the accelerometer-derived activities.

Acknowledgements:

We wish to thank the ISCOLE External Advisory Board and the ISCOLE participants and their families who made this study possible. A membership list of the ISCOLE Research Group and External Advisory Board is included in Katzmarzyk, Lambert and Church. An Introduction to the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). *Int J Obes Suppl*. (This Issue).

ISCOLE was funded by The Coca-Cola Company. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

MF has received a research grant from Fazer Finland and has received an honorarium for speaking for Merck. AK has been a member of the Advisory Boards of Dupont and McCain Foods. RK has received a research grant from Abbott Nutrition Research and Development. VM is a member of the Scientific Advisory Board of Actigraph and has received an honorarium for speaking for The Coca-Cola Company. TO has received an honorarium for speaking for The Coca-Cola Company. The authors reported no other potential conflicts of interest.

References

1. Allen MJ, Yen WM. *Introduction to Measurement Theory*. Waveland Press, Inc: Long Groove, IL, 2002.
2. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, et al. Methods of measurement in epidemiology: Sedentary behaviour. *Int J Epidemiol* 2012; **41**: 1460-1471.
3. Pedisic Z, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. *Br J Sports Med* 2015; **49**: 219-223.
4. Welk GJ, Schaben JA, Morrow JR, Jr. Reliability of accelerometry-based activity monitors: a generalizability study. *Med Sci Sports Exerc* 2004; **36**: 1637-1645.
5. McClain JJ, Sisson SB, Tudor-Locke C. Actigraph accelerometer interinstrument reliability during free-living in adults. *Med Sci Sports Exerc* 2007; **39**: 1509-1514.
6. Esliger DW, Tremblay MS. Technical reliability assessment of three accelerometer models in a mechanical setup. *Med Sci Sports Exerc* 2006; **38**: 2173-2181.
7. 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.
8. Silva P, Mota J, Esliger D, Welk G. Technical reliability assessment of the Actigraph GT1M accelerometer. *Meas Phys Educ Exerc Sci* 2010; **14**: 79-91.
9. Ried-Larsen M, Brond J, Brage S, Hansen B, Grydeland M, Andersen L, et al. Mechanical and free living comparisons of four generations of the Actigraph activity monitor. *International Journal of Behavioral Nutrition and Physical Activity* 2012; **9**: 113.
10. Kang M, Bjornson K, Barreira TV, Ragan BG, Song K. The minimum number of days required to establish reliable physical activity estimates in children aged 2-15 years. *Physiol Meas* 2014; **35**: 2229-2237.

11. Kang M, Bassett DR, Barreira TV, Tudor-Locke C, Ainsworth B, Reis JP, et al. How many days are enough? A study of 365 days of pedometer monitoring. *Res Q Exerc Sport* 2009; **80**: 445-453.

12. Baranowski T, Masse LC, Ragan B, Welk G. How many days was that? We're still not sure, but we're asking the question better! *Med Sci Sports Exerc* 2008; **40**: S544-549.

13. Hinkley T, O'Connell E, Okely AD, Crawford D, Hesketh K, Salmon J. Assessing volume of accelerometry data for reliability in preschool children. *Med Sci Sports Exerc* 2012; **44**: 2436-2441.

14. Deforche B, De Bourdeaudhuij I, D'Hondt E, Cardon G. Objectively measured physical activity, physical activity related personality and body mass index in 6- to 10-yr-old children: a cross-sectional study. *Int J Behav Nutr Phys Act* 2009; **6**: 25.

15. Mattocks C, Ness A, Leary S, Tilling K, Blair SN, Shield J, et al. Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision. *J Phys Act Health* 2008; **5 Suppl 1**: S98-111.

16. Taylor RW, Murdoch L, Carter P, Gerrard DF, Williams SM, Taylor BJ. Longitudinal study of physical activity and inactivity in preschoolers: the FLAME study. *Med Sci Sports Exerc* 2009; **41**: 96-102.

17. Katzmarzyk PT, Barreira TV, Broyles ST, Champagne CM, Chaput J-P, Fogelholm M, et al. The International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE): Design and methods. *BMC Public Health* 2013; **13**: 900.

18. Wickel EE, Welk GJ. Applying generalizability theory to estimate habitual activity levels. *Med Sci Sports Exerc* 2010; **42**: 1528-1534.

19. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? *Med Sci Sports Exerc* 2000; **32**: 426-431.

20. Wickel EE. Reporting the reliability of accelerometer data with and without missing values. *PLoS One* 2014; **9**: e114402.

21. Basterfield L, Adamson AJ, Pearce MS, Reilly JJ. Stability of habitual physical activity and sedentary behavior monitoring by accelerometry in 6- to 8-year-olds. *J Phys Act Health* 2011; **8**: 543-547.

22. Tudor-Locke C, Barreira TV, Schuna JM, Jr., Mire EF, Chaput JP, Fogelholm M, et al. Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). *Int J Behav Nutr Phys Act* 2015; **12**: 172.

23. Tudor-Locke C, Barreira TV, Schuna JM, Jr., Mire EF, Katzmarzyk PT. Fully automated waist-worn accelerometer algorithm for detecting children's sleep-period time separate from 24-h physical activity or sedentary behaviors. *Appl Physiol Nutr Metab* 2014; **39**: 53-57.

24. Barreira TV, Schuna JM, Jr., Mire EF, Katzmarzyk PT, Chaput JP, Leduc G, et al. Identifying children's nocturnal sleep using 24-h waist accelerometry. *Med Sci Sports* 2015; **47(5)**: 937-943.

25. Mark AE, Janssen I. Dose-response relation between physical activity and blood pressure in youth. *Med Sci Sports Exerc* 2008; **40**: 1007-1012.

26. 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.

27. Bates D, Maechler M, Bolker B, Walker S. lme4: Linear mixed-effects models using Eigen and S4. 2014; http://CRAN.R-project.org/package=lme4.

28. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; **86**: 420-428.

29. Davies G, Reilly JJ, McGowan AJ, Dall PM, Granat MH, Paton JY. Validity, practical utility, and reliability of the activPAL in preschool children. *Med Sci Sports Exerc* 2012; **44**: 761-768.

30. Hinckson EA, Hopkins WG, Aminian S, Ross K. Week-to-week differences of children's habitual activity and postural allocation as measured by the ActivPAL monitor. *Gait Posture* 2013.

31. Hart TL, Ainsworth BE, Tudor-Locke C. Objective and subjective measures of sedentary behavior and physical activity. *Med Sci Sports Exerc* 2010.

32. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc* 2011; **43**: 1561-1567.

33. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc* 2011; **43**: 1360-1368.

34. Tudor-Locke C, Camhi SM, Troiano RP. A catalog of rules, variables, and definitions applied to accelerometer data in the National Health and Nutrition Examination Survey, 2003-2006. *Prev Chronic Dis* 2012; **9**: E113.

Figure Legends

Figure 1. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived light physical activity (LPA).

Figure 2. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived moderate-to-vigorous physical activity (MVPA).

Figure 3. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived total activity counts.

Figure 4. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived sedentary time.

Table 1. Descriptive characteristics of study sample

Site	Boys			Girls		
	N	BMI (kg/m^2)	Number of valid days	N	BMI (kg/m^2)	Number of valid days
Australia (Adelaide)	225	18.6 ± 2.9	6.7 ± 0.6	266	19.1 ± 3.5	6.6 ± 0.7
Brazil (Sao Paulo)	242	19.9 ± 4.7	6.7 ± 0.6	252	19.5 ± 4.2	6.7 ± 0.7
Canada (Ottawa)	217	18.5 ± 3.4	6.8 ± 0.6	306	18.2 ± 3.3	6.7 ± 0.6
China (Tianjin)	261	19.8 ± 4.4	6.8 ± 0.5	240	17.9 ± 3.6	6.9 ± 0.4
Colombia (Bogota)	422	17.8 ± 2.6	6.7 ± 0.7	435	17.4 ± 2.4	6.6 ± 0.7
Finland (Helsinki, Espoo & Vantaa)	235	17.6 ± 2.5	5.8 ± 0.6	269	17.9 ± 2.6	5.8 ± 0.6
India (Bangalore)	254	17.7 ± 3.4	6.6 ± 0.7	299	18.2 ± 3.3	6.3 ± 0.7
Kenya (Nairobi)	233	17.1 ± 2.8	6.0 ± 0.9	269	17.3 ± 3.3	6.0 ± 0.9
Portugal (Porto)	305	19.5 ± 3.5	6.7 ± 0.6	381	19.4 ± 3.4	6.8 ± 0.5
South Africa (Cape Town)	184	17.7 ± 3.2	6.7 ± 0.6	284	18.1 ± 3.8	6.6 ± 0.7
United Kingdom (Bath & NE Somerset)	211	18.2 ± 2.7	6.6 ± 0.7	267	18.7 ± 3.2	6.6 ± 0.8
United States (Baton Rouge)	203	18.7 ± 3.6	6.4 ± 0.8	288	19.0 ± 4.1	6.4 ± 0.7

Note: BMI = body mass index

Table 2. Percent variance of different components of reliability estimates of accelerometer-determined metrics

		Boys				Girls			
Site	Factor	LPA	MVPA	Activity Counts	Sedentary Time	LPA	MVPA	Activity Counts	Sedentary Time
Australia (Adelaide)	Participant	45.1	32.4	31.6	33.3	42.8	35	35.6	28.1
	Day	3.7	6.1	1.7	2.0	4.6	4.7	2.2	3.3
	Residual	51.1	61.5	66.7	64.7	52.5	60.3	62.2	68.6
Brazil (Sao Paulo)	Participant	38.9	40.2	41.0	28.0	42.1	36.7	36.4	34.7
	Day	1.4	1.8	1.2	0.2	1.5	2.8	2.8	3.1
	Residual	59.7	58.0	57.8	71.8	56.4	60.5	60.8	62.2
Canada (Ottawa)	Participant	38.1	29.3	27.3	35.5	40.4	36	34.4	34.3
	Day	6.6	8.7	4.8	3.5	4.8	3.6	2.5	2.2
	Residual	55.3	62.0	67.9	61.0	54.8	60.4	63.1	63.5
China (Tianjin)	Participant	46.5	35.0	35.8	34.1	52.7	38.8	42.0	38.2
	Day	2.7	2.9	3.2	10.2	2.1	2.7	3.8	13.9
	Residual	50.8	62.0	61.0	55.8	45.2	58.5	54.2	47.9
Colombia (Bogota)	Participant	39.1	36.7	37.4	34.8	42.7	39.6	43.1	31.2
	Day	3.8	2.5	0.7	4.3	1.8	2.2	0.9	9.5
	Residual	57	60.8	61.9	60.9	55.5	58.2	56.1	59.3
Finland (Helsinki, Espoo & Vantaa)	PID	38.9	35.5	36.2	31.0	41.5	30.1	35.3	31.4
	Day	1.8	2.7	0.2	0.7	3.4	5.5	0.9	2.6
	Residual	59.3	61.8	63.6	68.3	55.1	64.5	63.8	66
India (Bangalore)	Participant	37.8	43.7	40.7	37.4	42.8	47.6	47.8	34.9
	Day	3.6	1.4	0.8	1.6	0.5	1.1	0.2	9.7
	Residual	58.6	54.9	58.5	61.0	56.7	51.3	52	55.3
Kenya (Nairobi)	Participant	38.0	47.1	42.8	29.1	41.6	57.2	49.5	24.1
	Day	0.9	2.2	2.5	9.7	0.5	0.9	2.1	12.5
	Residual	61.1	50.6	54.7	61.2	57.8	41.9	48.4	63.4
Portugal (Porto)	Participant	46.1	31.4	31.5	30.5	44	26.4	28.6	27.9
	Day	1.5	7.6	6.3	0.4	1.8	7.1	7.9	2.7
	Residual	52.3	61.0	62.2	69.1	54.2	66.4	63.5	69.3

South Africa (Cape Town)	Participant	47.4	46.4	47.3	35.1	51.7	49.7	48.5	30.6
	Day	2.5	0.8	0.3	6.8	2.2	0.6	1.2	14.6
	Residual	50.2	52.9	52.3	58.2	46.1	49.7	50.4	54.8
United Kingdom (Bath & NE Somerset)	Participant	39.1	32.1	32.5	29.3	45.1	32.7	32.4	28.4
	Day	5.5	4.4	0.7	0.9	4.4	4.5	0.7	3.6
	Residual	55.4	63.5	66.8	69.8	50.5	62.7	66.9	67.9
United States (Baton Rouge)	Participant	43.3	30.6	30.6	24.3	44.1	35.5	33.3	27.1
	Day	2.7	7.9	4.9	14.7	5.3	4.1	3.3	16.8
	Residual	54.0	61.6	64.5	61.0	50.6	60.4	63.4	56.1

Note: LPA - light physical activity; MVPA - Moderate-to-vigorous physical activity

Table 3. Intra-class correlation coefficients (ICC) values for different accelerometer-derived metrics.

	Boys				Girls				Total Sample			
Site	LPA	MVPA	Activity Counts	Sedentary Time	LPA	MVPA	Activity Counts	Sedentary Time	LPA	MVPA	Activity Counts	Sedentary Time
Australia (Adelaide)	0.84	0.76	0.76	0.77	0.83	0.78	0.79	0.78	0.84	0.80	0.79	0.75
Brazil (Sao Paulo)	0.81	0.82	0.82	0.72	0.83	0.79	0.79	0.78	0.82	0.85	0.84	0.75
Canada (Ottawa)	0.81	0.73	0.72	0.79	0.82	0.79	0.78	0.78	0.82	0.79	0.77	0.78
China (Tianjin)	0.86	0.79	0.79	0.77	0.88	0.81	0.83	0.80	0.88	0.81	0.83	0.8
Colombia (Bogota)	0.81	0.79	0.80	0.78	0.83	0.81	0.83	0.74	0.82	0.82	0.83	0.76
Finland (Helsinki, Espoo & Vantaa)	0.78	0.75	0.77	0.72	0.80	0.70	0.76	0.72	0.80	0.78	0.79	0.72
India (Bangalore)	0.80	0.84	0.82	0.79	0.83	0.85	0.85	0.77	0.89	0.82	0.88	0.81
Kenya (Nairobi)	0.79	0.85	0.82	0.69	0.82	0.89	0.86	0.64	0.8	0.88	0.85	0.67
Portugal (Porto)	0.86	0.76	0.76	0.75	0.84	0.72	0.74	0.73	0.85	0.80	0.80	0.75
South Africa (Cape Town)	0.86	0.85	0.86	0.78	0.87	0.87	0.86	0.74	0.87	0.87	0.87	0.75
United Kingdom (Bath & NE Somerset)	0.81	0.76	0.76	0.73	0.85	0.76	0.76	0.73	0.83	0.79	0.78	0.73
United States (Baton Rouge)	0.83	0.75	0.75	0.67	0.84	0.78	0.76	0.71	0.79	0.83	0.77	0.68
Mean	0.82	0.79	0.79	0.75	0.84	0.80	0.80	0.74	0.83	0.82	0.82	0.75

Note: LPA - light physical activity; MVPA - Moderate-to-vigorous physical activity

Figure 1. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived light physical activity (LPA).

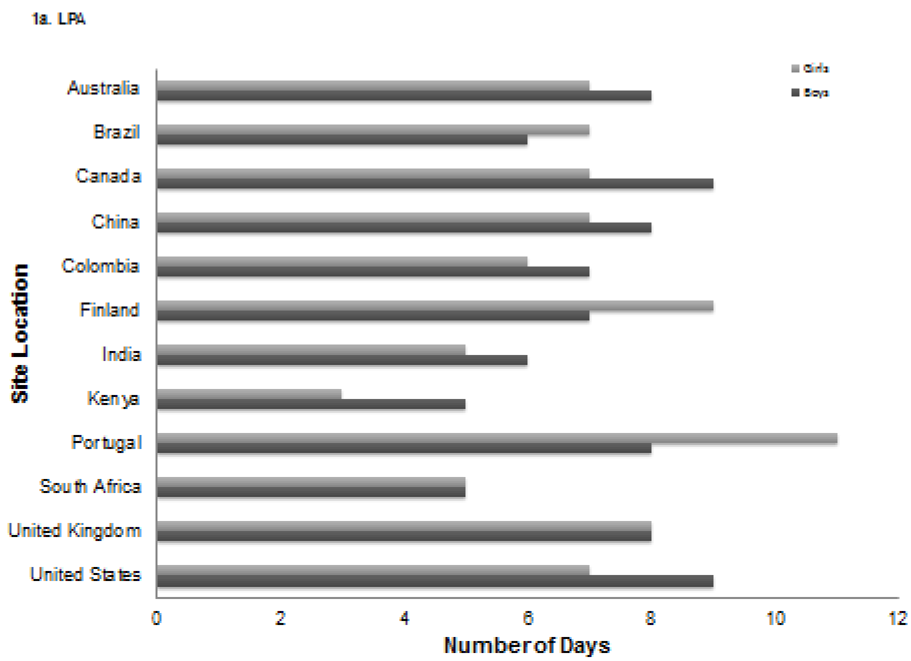

Figure 2. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived moderate-to-vigorous physical activity (MVPA).

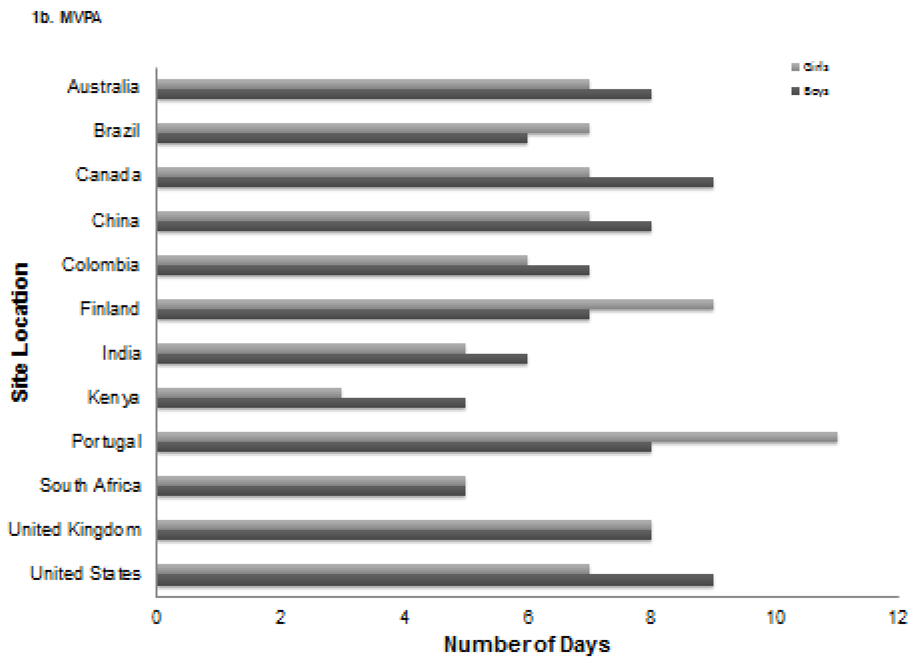

Figure 3. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived total activity counts.

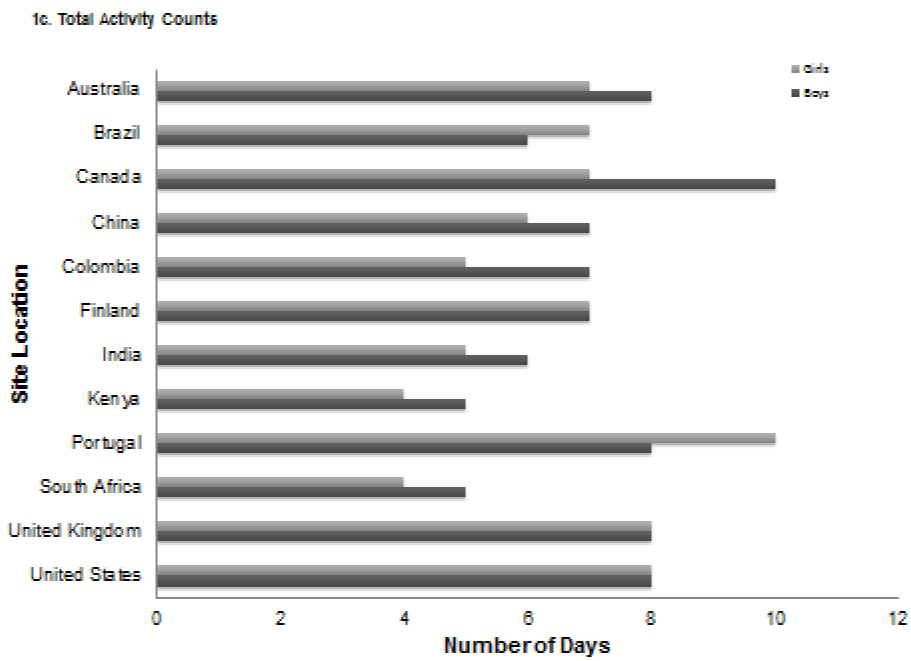

Figure 4. Number of days necessary to achieve a reliability coefficient G ≥ 0.80 for accelerometry-derived sedentary time.