2009

Intra-individual variation in children's physical activity patterns: Implications for measurement

Kate Ridley
Flinders University, kate.ridley@flinders.edu.au

Tim Olds
University of South Australia, tim.olds@unisa.edu.au

Beth P. Hands
University of Notre Dame Australia, bhands@nd.edu.au

Dawne Larkin
University of Western Australia, dlarkin@cyllene.uwa.edu.au

Helen Parker
University of Notre Dame Australia, hparker@nd.edu.au

This article was originally published as:
Ridley, K., Olds, T., Hands, B. P., Larkin, D., & Parker, H. (2009). Intra-individual variation in children's physical activity patterns: Implications for measurement. *Journal of Science and Medicine in Sport, 12* (5), 568-572.
http://doi.org/10.1016/j.jsams 2008.09.009

Available online at www.sciencedirect.com
ScienceDirect
Journal of Science and Medicine in Sport xxx (2008) xxx–xxx
Journal of Science and Medicine in Sport
www.elsevier.com/locate/jsams

Intra-individual variation in children's physical activity patterns: Implications for measurement

Kate Ridley [a,*], Tim Olds [b], Beth Hands [c], Dawne Larkin [d], Helen Parker [c]

[a] School of Education, Flinders University, Australia
[b] Sansom Institute, University of South Australia, Australia
[c] School of Health Sciences, University of Notre Dame Australia, Australia
[d] School of Sports Science, Exercise and Health, The University of Western Australia, Australia

Received 5 May 2008; received in revised form 28 August 2008; accepted 11 September 2008

Abstract

Children's physical activity (PA) patterns change from day to day. This intra-individual variability affects precision when measuring key physical activity and sedentary behaviour variables. This paper discusses strategies used to reduce the random error associated with intra-individual variability and demonstrates the implications for assessing PA when varying number of days are sampled. Self-reported data collected on two hundred and ninety eight 13–14-year-olds were used to compare estimates of PA and sedentary behaviour derived from between 1 and 7 days of recall data. Large intra-individual coefficients of variation were calculated for physical activity level (14.5%), moderate-to-vigorous physical activity (83.4%), screen time (60.8%) and sleep (12.2%). While the magnitude of error associated with estimating means decreased as more days were sampled, the paper notes that depending on the nature of the research question being asked, sampling fewer days may yield sufficiently precise estimates. Therefore, researchers should conduct power analyses based on estimated inter- and intra-individual variability and sample size to determine how many days to sample when assessing children's PA patterns.

Keywords: Motor activity; Adolescent; Child; Raine study; Self-report

1. Introduction

Physical activity (PA) researchers are often interested in questions such as: "How does daily activity behaviour contribute to health and well-being?", and "How do activity patterns vary among groups of children?" The ability to answer these questions relies on accurate measurement of true (or typical) activity patterns. Therefore, researchers estimate the 'true mean' of a PA variable, for instance, moderate-to-vigorous physical activity (MVPA). For an individual child, the true mean is the hypothetical average about which the child varies during a period over which the habitual PA pattern is maintained (e.g. a year or school term).

Like all measurements, assessment of PA entails two types of errors: systematic error (fixed or bias) and random error. Systematic error results in a consistent over- or under-estimation in activity levels across the sample, while random error occurs when over- and under-estimations are made for an individual's estimation of their true mean; however across the sample these errors occur randomly.[1] Error may result from measurement validity issues, such as over- or under-reporting of activity, or from large intra-individual variability in daily activity patterns, making it difficult to measure the true mean.[2–5] Therefore, activity measured on a single day is likely to be a poor estimate of activity over a week, a month or a year.

There are many modifying factors of children's PA that can contribute to day-to-day variation in an individual. Depending on the sampling techniques used (i.e. sampling of participants and sampling of monitoring days), these factors can be considered as random or systematic variability. The variation in daily activity patterns caused by weekend/weekday and seasonal effects result in systematic variations, as all children experience weekdays, weekends and changes in season, no matter how a sample is selected.

* Corresponding author.
E-mail address: Kate.Ridley@flinders.edu.au (K. Ridley).

1440-2440/$ – see front matter © 2008 Sports Medicine Australia. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsams.2008.09.009

Organised PA, on the other hand, may result in systematic or random variation, depending on the sampling techniques. Researchers must determine whether the factors influencing the variation in PA across a sample are being presented randomly or systematically. As systematic biases can present significant measurement error, the effect and magnitude of these biases need to be understood in order to effectively address them in study design. For example, if a research question requires estimation of activity across a year, days should be sampled across weekends and weekdays and across seasons.[3] Moreover, if the effectiveness of a school-based intervention is being assessed, researchers need to ensure the proportion of school sport days is similar in the two comparison groups.

Random intra-individual variability across samples affects various statistical analyses in different ways. The consequences of large intra-individual variation will depend on the research question being asked and therefore the statistical approach used to answer the question. When predicting or comparing means from independent samples, random intra-individual variability across a sample will not bias estimated group means.[1,6] The main impact of intra-individual variation in these situations is to increase total variance, therefore decreasing statistical power, making it more difficult to detect a true difference between groups.[1] When regression or correlational analyses are performed (e.g. relating television viewing time to socio-economic status) increased variability may result in an attenuation of the correlation coefficient, i.e. biased toward zero.[1] Nevertheless, if the level of intra-individual variation is known or predicted, various statistical techniques have been developed to estimate the true correlation coefficient after removing the effect of random error.[7] Random intra-individual variability causes more serious problems when distributional analyses (e.g. estimating the proportion of a population meeting PA recommendations) are performed. Large intra-individual variation causes distributions to be inflated and will then change the proportion of individuals located beyond a selected cut-point, leading to potential errors.

One way to reduce error and increase the precision of mean estimates is to increase the number of days being sampled. However, this solution may come at a price of greater respondent burden and reactivity. Many repeated measures also increase the time and financial cost involved in administration and analyses. A number of research studies have posed the question of how many days should be sampled to get an accurate estimation of children's activity levels.[2,5,8] In an attempt to answer this question, the studies calculated an intra-class correlation coefficient (ICC), which assesses the proportion of total variance (i.e. inter- and intra-individual variation) that is accounted for by inter-individual variation.[3] Predictive formulae, such as the Spearman–Brown Prophecy formula,[3] are then used to estimate the number of days required to reach target reliability[5,8,9] based on the established ICC. These studies have reported between 3 and 9 days of PA monitoring are required to reach a reliability coefficient of 0.80, depending on the sample characteristics.[5,8,9]

While this approach offers interesting insights regarding the variability of PA across a variety of samples and measurement instruments, applying any of these target number of days to all studies of PA in children has limitations. First, the ICC is constrained by the sample in which it was collected.[2] This is because the magnitudes of intra- and inter-individual variances in PA are specific to the sample in which they are collected; and the factors that influenced PA in the days that were sampled in the monitoring period. A limitation of using the Spearman–Brown formula in this setting is that formula assumes the ICC remains the same when additional monitoring days are added. However, if days with smaller day-to-day variation in PA are sampled (e.g. only weekdays monitored), the observed intra-individual variability is likely to be smaller, resulting in a higher ICC. Moreover, if a sample containing participants with greater inter-individual variation in PA was used, the ICC would be increased.[2,3]

The sample-specific nature of the ICC has been demonstrated in a number of PA studies. Trost and colleagues[5] found age-related differences in the reliability, calculated as ICCs, with adolescents (grades 7–12) displaying greater daily intra-individual variability in objectively measured PA than children (grades 1–6). Different types of activities and particular segments of the day may also display diverse levels of intra- and inter-individual variability resulting in variations of reliability. Fairclough and colleagues[8] found reliability coefficients of daily activity in children were sex-specific and that specific periods of the day that were more regulated, such as in-school hours, demonstrated superior reliability compared to the entire day. Finally, Janz and colleagues[9] found variations in ICCs depending on what activity measure was used, or how the data were analysed.

In addition, when using the Spearman–Brown formula small changes in the target reliability coefficient can radically alter the number of days "required". For example, using the Spearman–Brown formula[9] and given an ICC of 0.4 (which is not atypical for PA variables in children[9]), a target ICC of 0.7 would require 4 days of monitoring. A target ICC of 0.8 would require 6 days, while 0.9 would require 14 days. Fairclough and colleagues[8] also demonstrated the limitations of the formula, when a negative number of monitoring days required to reach a reliability of 0.80 was derived due to a wide variation in girls' early evening PA scores.

Finally, this approach relies entirely on increasing the number of repeated measures. It does not consider the effect of increasing sample size on statistical power, for estimating true means. In order to increase statistical power, total variance can be decreased either by increasing the number of days measured for each individual (thus reducing intra-individual variance) or by increasing the total number of participants (thus reducing the standard error of the mean).

Therefore, simply asking "How many days of data are needed to estimate typical activity?" is problematic. As Beaton[1] suggests (in relation to dietary data) the question

Please cite this article in press as: Ridley K, et al. Intra-individual variation in children's physical activity patterns: Implications for measurement. *J Sci Med Sport* (2008), doi:10.1016/j.jsams.2008.09.009

Table 1
Subject characteristics and activity summary data.

	Boys ($n = 130$)	Girls ($n = 168$)	All ($n = 298$)
Age	14.0 (0.20)	14.0 (0.37)	14.0 (0.31)
Height (cm)	166.0 (10.5)	163.5 (6.1)	164.5 (8.4)
Mass (kg)	58.3 (13.2)	56.2 (13.8)	57.1 (13.6)
PAL	1.65 (0.20)	1.56 (0.16)	1.60 (0.19)
MVPA (min/day)	114 (54)	93 (50)	102 (53)
Screen time (min/day)	210 (111)	161 (83)	182(99)
Sleep time (min/day)	576 (50)	586 (50)	582 (50)

Note: Data are shown as mean (S.D.). The data for PAL, MVPA, screen time and sleep are 7-day averages. MVPA, moderate-to-vigorous physical activity; PAL, physical activity level (an expression of the multiple of the resting metabolic rate for a day's energy expenditure).

needs to be recast as: "How many days are required to do what with what precision"?

Another approach in assessing the impact of intra-individual variability on the measurement of PA is to consider the impact on precision of estimates when progressively fewer days are sampled. The purpose of this study is to use self-reported PA data collected in the Raine longitudinal survey in Western Australia to compare estimates of PA and sedentary behaviour derived from between 1 and 7 days of recall data, and to discuss the implications for survey design.

2. Methods

The Raine longitudinal survey is a long-term project which started as a pregnancy cohort in which 2979 women attending antenatal clinics at King Edward Memorial Hospital for Women were enrolled between 1989 and 1991. The children have been followed at birth, 1, 2, 3, 6, 8, 10, and now at 14 years of age. The protocol for the original study has been reported elsewhere.[10] The data presented in this paper were collected on a sub-sample of the larger cohort. Table 1 shows the subject characteristics.

PA was measured by a self-report questionnaire, the Multimedia Activity Recall for Children and Adolescents (MARCA). The MARCA is a computerised use-of-time instrument which allows children to recall everything they did from the time they woke up to the time they went to bed on the previous day. The MARCA's validity is comparable to other self-report instruments, with correlations of $\rho = 0.57$ and 0.41 for physical activity level (PAL) and daily minutes of MVPA respectively, in children aged 11 and over, when compared to accelerometry. Test–retest reliability is high with coefficients of $\rho = 0.84$ and 0.83 for PAL and MVPA respectively.[11]

Each participant recalled at least 7 days. Where children had recalled more than 7 days, 7 days were randomly chosen from all the days recalled and used in the analysis. At least 1 weekend day was included for each participant. PAL, an expression of the multiple of the resting metabolic rate for a day energy expenditure, was calculated as a weighted mean MET day^{-1} score from the energy costs assigned to each activity and the reported duration of the activities.[11] MVPA was defined as any activity requiring at least three METs. Screen time included television/video/DVD, computer use and videogames. Sleep time was calculated from reported wake-up and bed times. For the purposes of analysis, the 7-day sample was considered to be the criterion measure of "typical" behaviour.

Recalls of 7 days, usually consecutive and in 95% of cases including 2 weekend days and 5 weekdays, were available from two hundred and ninety eight 13–14-year-old children.

To calculate the estimates of these response variables (PAL, MVPA, screen time and sleep time) from subsets of the 7 recall days, between 1 and 6 days were chosen at random without replacement. From each subset, the following sample statistics were calculated for each of the four response variables:

- the ICC,[12] between the subset averages and the 7-day averages. This statistic shows how well the averages derived from the subset correspond to the 7-day averages across the whole sample and
- the mean absolute difference between the subset averages and the 7-day averages. This statistic shows the typical difference between the average derived from the subset and the 7-day average for any individual. This procedure was performed 100 times for each subset (1–6 days), and the distribution of the derived statistics recorded.

3. Results

The various days of the week were approximately equally represented among the recalled days (ranging from 279 Saturdays to 309 Tuesdays and Thursdays). Exactly half of the days recalled were in autumn or winter, and half in spring or summer. Exactly half of the days recalled were non-school days, reflecting a typical year for an Australian 13–14-year-old when weekends, holidays, sick days and school closure days are taken into account.

There was substantial intra-individual variability in all response variables across 7 days, with mean coefficients of variation (S.D.) of 14.5 (6.2)% for PAL, 83.4 (35.7)% for MVPA, 60.8 (32.6)% for screen time and 12.2 (4.8)% for sleep. The corresponding inter-individual CVs (based on 7-day averages) were 11.7% (PAL), 52.0% (MVPA), 54.3% (screen) and 8.6% (sleep).

Please cite this article in press as: Ridley K, et al. Intra-individual variation in children's physical activity patterns: Implications for measurement. *J Sci Med Sport* (2008), doi:10.1016/j.jsams.2008.09.009

Table 2 shows the ICCs between 1–6-day subsets and the 7-day ("typical activity") averages. ICCs increased curvilinearly as the subset increased from 1 to 6 days. Using a 4-day subset, ICCs were typically about 0.90, showing strong correlations between 4-day subsets and 7-day averages. Table 2 also shows the mean absolute differences between subset averages and the 7-day average for individual participants. Again as expected the differences diminish as the number of days in the subset increases. The mean absolute difference for PAL diminished from 11% of the average PAL for the 1-day subset, to 4% for the 4-day subset, to 2% for the 6-day subset. The corresponding figures were 57%, 21% and 10% for MVPA; 39%, 14% and 7% for screen time; and 9%, 3% and 2% for sleep.

4. Discussion

These results demonstrate the variation in type and magnitude of error that may result from large intra-individual variability in children's daily activity levels. However, depending upon the requirements of the analysis, the instruments used, the nature of the response variables and sample characteristics, samples of fewer days may yield sufficiently precise estimates. A 4-day sample from the population in the present dataset, for example, will yield PALs which are strongly correlated with 7-day averages (ICC = 0.90), show a mean absolute difference for individuals of just 4.4%, and are within 11% of the 7-day average for 95% of individuals.

Surveys of children's PA can be used to answer many different questions. For example, our primary interest may be to estimate population mean values from a sample, perhaps in order to track secular changes or to monitor the effects of an intervention. In other cases, we may be more interested in the proportion of children who achieve certain threshold levels of activity or screen time, in order to monitor compliance with guidelines. In others still, our primary focus may be the relationship between activity patterns and other psychological, social, demographic or use-of-time variables. In some cases, these relationships may be time-specific (e.g. do children who acquire a lot of screen time during an evening go to bed later that night?), in others they may relate to habitual activity (e.g. are singleton children more active than children with siblings?). The precision with which we can answer each of these different types of questions will be impacted differently by the number of days we sample, and will have different potential trade-offs with the number of children we sample. In general, aside from avoiding systematic biases, there is no reason why decisions regarding the number of days to be sampled should not follow the same considerations as usual power analyses. Any attempt to specify a target number of recall days for all PA studies in children is misconceived. There will be occasions when 7 or more days of measurement may be required. This is especially true when monitor-

Table 2
ICCs and mean absolute differences between 1–6-day subset averages and 7-day ("typical activity") averages.

Number of days in subset	PAL		MVPA		Screen time		Sleep	
	ICC	Mean error (METs)	ICC	Mean error (min/day)	ICC	Mean error (min/day)	ICC	Mean error (min/day)
1	0.54 (0.038)	0.18 (0.160)	0.49 (0.027)	58 (50)	0.69 (0.024)	71 (60)	0.50 (0.029)	53 (47)
2	0.73 (0.018)	0.12 (0.004)	0.70 (0.019)	38 (32)	0.85 (0.013)	46 (38)	0.71 (0.018)	36 (28)
3	0.83 (0.014)	0.09 (0.003)	0.81 (0.016)	28 (23)	0.91 (0.007)	34 (28)	0.82 (0.013)	26 (21)
4	0.90 (0.008)	0.07 (0.003)	0.88 (0.010)	21 (17)	0.95 (0.005)	25 (21)	0.89 (0.008)	19 (15)
5	0.94 (0.005)	0.05 (0.003)	0.93 (0.006)	15 (13)	0.97 (0.002)	18 (15)	0.94 (0.005)	14 (11)
6	0.98 (0.002)	0.03 (0.001)	0.97 (0.003)	10 (8)	0.99 (0.001)	12 (10)	0.97 (0.002)	9 (8)

Note: Values are shown as averages of 100 random draws with standard errors in parentheses.

Please cite this article in press as: Ridley K, et al. Intra-individual variation in children's physical activity patterns: Implications for measurement. *J Sci Med Sport* (2008), doi:10.1016/j.jsams.2008.09.009

ing changes within an individual, and there is no option of increasing power by increasing the number of participants.

Estimates of intra- and inter-individual variability are required for power analyses, and the data shown here may be used as estimates for similar populations. MVPA and screen time showed very large intra-individual variability (CVs > 60%) and show strong positive skews, while PAL and sleep time showed lower variability (<15%). This is probably because PAL includes a large number of obligatory activities, such as sleeping and self-care, and sleep habits are highly routinised, at least on school days.

Our results, including ICC calculations and mean error estimations, are bound by the measurement tool used and the sample in which the data were collected. The sample was not a representative sample of Australian children and we cannot assume that the intra- or inter-individual coefficients of variation are generalisable to other populations of children. However, it is important to note that our goal was to use a dataset of PA data collected in children on multiple days to demonstrate variation in the level of precision in estimating a 'true' mean when different numbers of monitoring days are used. Our aim was not to recommend a set target of monitoring days for any one instrument. We conclude that when calculating the number of days required to assess activity patterns, researchers should conduct appropriate power analyses based on the best available data on inter- and intra-individual variability in the key response variables, rather than adopt blanket recommendations.

Practical implications

- Intra-individual variability in the daily physical activity patterns of children differs depending on the characteristics of the children, the PA variable being measured, the days sampled and the instrument used to measure the activity.
- The impact of intra-individual variability on precision of measurement varies depending on whether the variability is systematic or random and what statistical analyses are being undertaken.
- Researchers should consider the levels of inter- and intra-individual variation in the physical activity variables of interest, to assist determining how many days of measurement to undertake.

Acknowledgements

We would like to thank all the adolescent participants and their families. We also thank Nick Sloan, Rosemary Austin, Lee Clohessy, Alex D'Vauz, Monique Robinson and Diana Wood for data collection. The Western Australian Pregnancy Cohort Study is funded by the Raine Medical Research Foundation, the National Health and Medical Research Council (NHMRC) of Australia, the Telstra Foundation and the Western Australian Health Promotion Foundation.

References

1. Beaton GH. Approaches to analysis of dietary data: relationship between planned analyses and choice of methodology. *Am J Clin Nutr* 1994;**59**(1):253S–61S.
2. Baranowski T, De Moor C. How many days was that? Intra-individual variability and physical activity assessment. *Res Q Exerc Sport* 2000;**71**(2):74–8.
3. Levin S, Jacobs DR, Ainsworth BE, et al. Intra-individual variation and estimates of usual physical activity. *Ann Epidemiol* 1999;**9**: 481–8.
4. Matthews CE, Hebert JR, Freedson PS, et al. Sources of variation in daily physical activity levels in the seasonal variation of blood cholesterol study. *Am J Epidemiol* 2001;**153**(10):987–95.
5. Trost SG, Pate RP, Freedson PS, et al. Using objective physical activity measures with youth: how many days of monitoring are needed? *Med Sci Sports Exerc* 2000;**32**(2):426–31.
6. Liu K, Stamler J, Dyer A, et al. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *J Chronic Dis* 1978;**31**(6–7):399–418.
7. Rosner B, Willett WC. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. *Am J Epidemiol* 1988;**127**(2): 377–86.
8. Fairclough SJ, Butcher ZH, Stratton G. Whole-day and segmented day physical activity variability of northwest England school children. *Prev Med* 2007;**44**(5):421–5.
9. Janz KF, Witt J, Mahoney LT. The stability of children's physical activity as measured by accelerometry and self-report. *Med Sci Sports Exerc* 1995;**27**(9):1326–32.
10. Newnham JP, Evans SF, Michael CA, et al. Effects of frequent ultrasound during pregnancy: a randomised controlled trial. *Lancet* 1993;**342**(8876):887–91.
11. Ridley K, Olds T, Hill A. The multimedia activity recall for children and adolescents (MARCA): development and evaluation. *Int J Behav Nutr Phys Act* 2006;**3**(May (10)).
12. Patterson P. Reliability, validity, and methodological response to the assessment of physical activity via self-report. *Res Q Exerc Sport* 2000;**71**(2):15–20.

Please cite this article in press as: Ridley K, et al. Intra-individual variation in children's physical activity patterns: Implications for measurement. *J Sci Med Sport* (2008), doi:10.1016/j.jsams.2008.09.009