ACCEPTED 26.10.2016 IN THE JOURNAL OF SPORTS SCIENCES

Title: Comparison of children's free-living physical activity derived from wrist and hip raw accelerations during the segmented week.

Robert J. Noonan[1], Lynne M. Boddy[1], Youngwon Kim[2], Zoe R. Knowles[1], Stuart J. Fairclough[3, 4]

[1]Physical Activity Exchange, Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK

[2]MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK

[3]Department of Sport and Physical Activity, Edge Hill University, Ormskirk, UK

[4]Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland

Corresponding author details: Robert J. Noonan

The Physical Activity Exchange,

Research Institute for Sport & Exercise Sciences,

Liverpool John Moores University,

62 Great Crosshall Street,

Liverpool, L3 2AT, UK

Email: R.J.Noonan@2010.ljmu.ac.uk

Tel: +44 (0)151 231 4275

Acknowledgements

The authors would like to thank the children, schools, teachers and parents involved in this study. This study was funded by Liverpool John Moores University.

Abstract

This study assessed children's physical activity (PA) levels derived from wrist-worn GENEActiv and hip-worn ActiGraph GT3X+ accelerometers and examined the comparability of PA levels between the two devices throughout the segmented week. One hundred twenty nine 9-10 year old children (79 girls) wore a GENEActiv (GAwrist) and ActiGraph GT3X+ (AGhip) accelerometer on the left wrist and right hip respectively for seven days. Mean minutes of light PA (LPA) and moderate-to-vigorous PA (MVPA) per weekday (whole-day, before-school, school and after-school) and weekend day (whole-day, morning and afternoon-evening) segments were calculated, and expressed as percentage of segment time. Repeated measures ANOVA examined differences in LPA and MVPA between GAwrist and AGhip for each time segment. Bland–Altman plots assessed between-device agreement for LPA and MVPA for whole weekday and whole weekend day segments. Correlations between GAwrist and AGhip were weak for LPA (r=0.18-0.28), but strong for MVPA (r=0.80-0.86). LPA and MVPA levels during all weekday and weekend day segments were significantly higher for GAwrist than AGhip (p<0.001). The largest inter-device percent difference of 26% was observed in LPA during the school day segment. Our data suggest that correction factors are needed to improve raw PA level comparability between GAwrist and AGhip.

Introduction

Accelerometers provide valid and reliable assessments of physical activity (PA) at varying intensities in children (Butte, Ekelund, & Westerterp, 2012; de Vries et al., 2009), and are the most widely used objective measure of child PA (Cain, Sallis, Conway, Van Dyck, & Calhoon, 2013). One of the advantages of using accelerometers is their ability to capture PA variability within and between days. Accelerometer device output is traditionally expressed as an arbitrary 'count' value which is then related to specific PA intensity thresholds. Due to differences in how raw data are processed, filtered, and scaled, count data cannot be directly compared across studies using different accelerometer devices (Welk, McClain, & Ainsworth, 2012). However, the latest versions of accelerometers, including

GENEActiv and ActiGraph GT3X+ can provide raw, unfiltered acceleration data. Compared to traditional count-based approaches, raw acceleration data offers greater control over data reduction, potentially allowing comparisons to be made more easily between studies using different accelerometer brands (Fairclough et al., 2016; Hildebrand, Van Hees, Hansen, & Ekelund, 2014).

Aside from the challenge of comparing PA levels between device brands, another challenge is the comparability of PA levels between devices placed at different body locations. Traditionally, accelerometers are worn at the hip to capture whole-body movement, but compliance to device wear is typically low (Fairclough et al., 2016). In an attempt to improve device wear there has been an increased use of wrist-worn accelerometers, including the GENEActiv. Compared to hip-worn accelerometers, wrist-worn accelerometers are more sensitive to upper body movement (e.g. climbing, throwing) but less sensitive to sedentary activities (Ellis et al., 2014; Ellis, Kerr, Godbole, Staudenmayer, & Lanckriet, 2016; Kim, Lee, Peters, Gaesser, & Welk, 2014). This may limit the comparison of findings between studies using wrist and hip-worn accelerometers. Given the increased use of the wrist-worn GENEActiv (da Silva et al., 2014; Edwardson et al., 2015; Keane, Kearney, Perry, Browne, & Harrington, 2014; Wake et al., 2014), and the wealth of existing international data obtained from hip-worn ActiGraph accelerometers (Cooper et al., 2015; Corder et al., 2016; Sherar et al., 2011) it is important to understand whether PA estimates derived from GAwrist and AGhip are comparable.

Fairclough et al. (2016) compared children's whole-day MPA and VPA derived from the GAwrist and AGhip and found that mean PA levels for both intensities were significantly higher for the GAwrist than the AGhip. However, the comparability of PA levels between the GAwrist and AGhip at the lower end of the intensity spectrum is less well understood. Moreover, the agreement between the GAwrist and AGhip may fluctuate in response to variability in PA levels both within and between days (Brooke, Corder, Atkin, & van Sluijs, 2014; Fairclough, Beighle, Erwin, & Ridgers, 2012). However, studies comparing GAwrist and AGhip data have been limited to reporting PA estimates (Fairclough et al., 2016; Rowlands et al., 2014), and raw accelerations across the whole day (Rowlands et al., 2015). Therefore, little is known about their comparability across specific time-segments. For that reason, the

aim of this study was to assess children's PA levels derived from GAwrist and AGhip raw acceleration data, and examine the comparability of PA levels between the two devices throughout the segmented week.

Methods

Participants and settings

The participants were 129 children (79 girls) aged 9-10 years (age: 10.1 ± 0.3 y (mean ± SD)) from six schools in Liverpool, England. After ethical approval from the university research ethics committee (13/SPS/048), all year 5 children (n = 326) in participating schools were invited to participate and received parent and child information sheets, and consent and assent forms, to take home to parents and return upon completion. Written informed consent and assent were received from parents and their children, respectively, before children could participate in the study. Data collection took place between January and May 2014.

Procedure and measurements

Each child wore a GENEActiv (GAwrist; Activinsights, Cambs, UK) and ActiGraph GT3X+ (AGhip; ActiGraph, Pensacola, FL) accelerometer on their left wrist and right hip, respectively, for seven consecutive days. The GAwrist was selected because it measures raw accelerations, is typically worn on the wrist, and has demonstrated reliability and validity in child populations (Phillips, Parfitt, & Rowlands, 2013). ActiGraph accelerometers are the most commonly used accelerometer in child PA research (Cain et al., 2013). The GT3X+ model was selected because it is traditionally worn on the hip (Rosenberger et al., 2013), has the capability to generate raw acceleration, and has been validated for use with children (Hanggia, Phillips, & Rowlands, 2013; Robusto & Trost, 2012). Children were instructed to wear both monitors concurrently during all waking hours except when engaged in water-based activities. Verbal and written instructions for care and placement of the monitors were given to children. Prior to testing, monitors were synchronised with Greenwich Mean Time (GMT) and

programmed to record data at 100 Hz. Data collection took place during the regular school term so activities were representative of usual free-living activities.

Data analysis

GAwrist data were downloaded using GENEActiv v.2.2 software (Activinsights, Cambs, UK) and saved in raw format as binary files. AGhip data were downloaded using ActiLife v. 6.11.4 (ActiGraph, Pensacola, FL) and saved in raw format as GT3X files. These were subsequently converted to CSV format to facilitate raw data processing. GAwrist and AGhip raw data files were then processed in R (http://cran.r-project.org) using the GGIR package (version 1.1-4) which converted raw triaxial acceleration values into one omnidirectional measure of acceleration, termed the signal vector magnitude (SVM). SVM was calculated from raw accelerations from the three axes minus 1*g* which represents the value of gravity (i.e., SVM = $\sqrt{(x^2 + y^2 + z^2)} - 1$), after which negative values were rounded to zero. This metric is referred to as the Euclidean norm minus one (ENMO) (van Hees et al., 2013). Raw data were further reduced by calculating the average SVM values per 1-s epoch expressed in *mg* over each of the 7 monitored days. Wear time periods for raw data from GAwrist and AGhip were estimated on the basis of the standard deviation and value range of each axis, calculated for 60 min moving windows with 15 min increments (van Hees et al., 2013). A time window was classified as non-wear time if, for at least 2 out of the 3 axes, the standard deviation was less than 13.0 *mg* or if the value range was less than 50 *mg* (van Hees et al., 2013). A valid day was classified as 10 hours or more of device wear. At a minimum, children were required to have worn both devices on the same 3 days including 1 weekend day to be included in the analyses. (Mattocks et al., 2008).

We used device specific prediction equations provided by Hildebrand et al. (2014) to identify ENMO cut-points for classifying LPA and MVPA (Hildebrand et al., 2014). It has recently been reported that in youth 2 METs and 4 METs had higher classification accuracy for differentiating sedentary time (from LPA) and MVPA (from LPA), respectively, compared with 1.5 METs and 3 METs (Saint-Maurice, Kim, Welk, & Gaesser, 2016). Therefore, the Hildebrand equations were solved for 2 METs and 4 METs resulting in LPA and MPVA cut-points of 23.5*mg* and 359.7*mg*, respectively, for GAwrist, and

35.2*mg* and 249.9*mg*, respectively, for AGhip. For example, the GAwrist LPA *mg* cut-point threshold was calculated as follows: *mg* = ((2METs x 6 mL $O_2 \cdot kg^{-1} \cdot min^{-1}$) – 11.16)/0.0357 = 23.5*mg*.

Once converted to minutes of LPA and MVPA, data were sorted into hourly segments from 06:30 until 23:59 on weekdays and weekend days using Stata (STATA/SE Version 12; StataCorp LP, College Station, TX) code developed by the third author. Sleep time was defined as midnight until 06:30. These hourly values were then used to construct whole-day and segmented day minutes of LPA and MVPA. During weekdays the following time segments were used: before-school (06.30 to 08:59), during school (09:00 to 15:29), and after-school (15:30 to 23:59). For weekend days the segments were: morning (06:30 to 11:59) and afternoon-evening (12:00 to 23:59). Variables were calculated by summing minutes spent in each activity threshold during each discrete time segment. Mean minutes of GAwrist and AGhip LPA and MVPA data for each segment were divided by total segment time, multiplied by 100, and expressed as percentage of total segment time.

The primary outcome variables were percentage segment time for LPA and MVPA. Repeated measures ANOVAs examined between segment differences for each device (e.g., GAwrist LPA whole weekday vs GAwrist LPA whole weekend day), and between device differences for each segment (e.g., GAwrist LPA whole weekday vs AGhip LPA whole weekend day). Pearson correlation analyses examined associations between the two devices for percentage of time spent in LPA and MVPA during whole-day weekday and weekend day. Bland–Altman plots were constructed to assess between-device agreement of LPA and MVPA for whole weekday and whole weekend day segments. All analyses were conducted using IBM SPSS Statistics v.23 (IBM, Armonk, NY) and Microsoft Excel 2010 (Microsoft, Redmond, WA). For all analyses, statistical significance was set at 0.05.

Results

AGhip and GAwrist data were available for 115 and 128 children, respectively. Participants not meeting the wear time criteria for either monitor were excluded from analyses. This reduced the sample to 107 (67 girls) for the GAwrist and 83 (51 girls) for the AGhip. Children without 3 valid days for both

monitors were then excluded from the analysis, resulting in a final analytical sample of 77 (48 girls) participants. There were no significant differences for any of the measured variables between children included in analyses and those excluded. Means and 95% confidence intervals (CI) for PA outcomes on weekdays and weekend days for GAwrist and AGhip are presented in Table 1. Whole weekday PA outcomes were higher than mean whole weekend day PA outcomes (p<0.05). PA outcomes were higher during the school segment compared to all other weekday segments (p<0.001). On weekend days children were more active in the afternoon-evening compared to the morning (p<0.01).

GAwrist PA levels were significantly higher than AGhip PA levels during all weekday and weekend day segments (p<0.001; Table 1) but varied between time segments and PA intensities. On weekdays the largest inter-device differences in PA levels occurred during the school segment (LPA 26.7%; MVPA 1.8%; p<0.001), and the smallest inter-device differences occurred in the before school segment (LPA 10.3%; MVPA 0.5%; p<0.001). On weekend days the largest inter-device differences occurred in the afternoon-evening (LPA 17.7%; MVPA 1.6%; p<0.001), and the smallest inter-device differences occurred in the morning (LPA 10.3%, MVPA 0.8%; p<0.001). For all intensities the magnitude of inter-device differences was largest at weekends compared to weekdays.

Significant correlations between whole weekday (r=0.80) and whole weekend day (r=0.86) MVPA levels confirmed that MVPA was strongly associated between devices (p<0.001). Correlations between the devices were weak for LPA during whole weekdays (r=0.28 p<0.01) and whole weekend days (r=0.18; p=0.11). Bland–Altman plots (Figure 1) show the extent of differences in LPA and MVPA between GAwrist and AGhip during whole weekdays and weekend days.

[TABLE 1 NEAR HERE]

[FIGURE 1 NEAR HERE]

Discussion

This is the first study to compare children's LPA and MVPA assessed with GAwrist and AGhip across

distinct time windows in a week. Another novel aspect of this study is the use of raw data processing techniques, which theoretically enables direct comparisons of activity outcomes obtained from different accelerometer brands. Overall, we observed weak correlations between AGhip and GAwrist for LPA (r=0.18-0.28), but strong correlations for MVPA (r=0.80-0.86). The strong correlations observed for MVPA are similar to those reported by Fairclough et al. (2016). They are though slightly lower than the reported correlation of r=0.93 between hip-worn GENEActiv and ActiGraph GT3X+ mean accelerations (Rowlands et al., 2015). Despite these strong associations, we found that GAwrist derived PA levels were consistently higher than those derived from the AGhip for all outcome variables and across various time segments. These findings suggest that child PA surveillance is strongly influenced by device brand and body placement.

LPA and MVPA levels during all weekday and weekend day segments were significantly higher for the GAwrist than those for the AGhip (p<0.001). Previous research comparing whole-day accelerometer output from wrist-worn GENEActiv and hip-worn ActiGraph in children reported similar findings (Fairclough et al., 2016; Hildebrand et al., 2014). Fairclough et al. (2016) reported a 68% difference in the number of children achieving at least 60 minutes of MVPA per day using the GENEActiv compared to ActiGraph GT3X+. Similarly, Rowlands et al. (2015) found that average daily accelerations from the wrist-worn GENEA were between 12%–13% higher than the ActiGraph GT3X+. Another recent study found that the ActiGraph GT3X+ worn on the wrist produced higher average step counts per day compared to the ActiGraph GT3X+ at the hip in free-living environments, but fewer steps during laboratory treadmill testing (Tudor-Locke, Barreira, & Schuna, 2015). These contrasting differences in step outputs between research settings are likely consequential of the restrictive nature of treadmill walking which minimises free swinging of the arms relative to free-living.

A unique element of this study is the comparison of PA levels between GAwrist and AGhip across different time segments. We found that differences in PA levels between the two devices varied in magnitude between intensity levels. As the intensity level increased, the magnitude of the difference in PA levels between the GAwrist and AGhip decreased. The largest differences in PA levels were seen

in LPA. . Mean GAwrist LPA was over 100% higher than that for the AGhip in all segments with the exception of the before school segment.

During free-living children typically engage in a range of seated activities that involve a high level of arm movement but limited movement at the hip (Kim et al., 2014). Unsurprisingly, during such activities, disproportionate levels of acceleration will be observed at the wrist relative to the hip. This is reflected by the high inter-device difference in LPA during the school day segment. LPA accounted for 42.6% and 15.6% of school segment time for the GAwrist and AGhip, respectively, a difference of over 26%. The profound difference in LPA observed during the school day likely reflects these disjointed wrist and hip movement patterns when children characteristically spend a large proportion of the day seated at a desk reading, writing, or using a computer which all involve some element of wrist movement. Greater accelerations will also be observed at the wrist relative to the hip during mixed static/dynamic movements (e.g., playing catch), and high intensity activities such as running and jumping that naturally incur a medium to high level of shoulder and upper body rotation (Ellis et al., 2014, 2016; Kim et al., 2014). However, the level of decoupling (i.e., greater acceleration capture at one wear site relative to the other) during such activities is likely dependent on individual biomechanics (i.e., level of arm swing), and thus will be population specific (Rowlands & Stiles, 2012; Tudor-Locke et al., 2015).

The weaker correlations and larger inter-device differences observed for LPA compared to MVPA suggests that in children of this age, pro-wrist "decoupling", is more dominant during LPA. In contrast, earlier studies observed greater decoupling as the magnitude of acceleration increased. However, these studies did not examine accelerations at intensities lower than 3 METs (Fairclough et al., 2016; Hildebrand et al., 2014). Children's free-living accelerations were over 10% greater for the GENEActiv compared to the ActiGraph in a recent study when both devices were worn at the hip (Rowlands et al., 2015). This suggests that additional factors other than monitor placement may have also contributed to the observed differences in GAwrist and AGhip PA levels. Similarly, John, Sasaki, and Staudenmayer (2013) found that GENEActiv peak accelerations were up to 7.4% greater than ActiGraph peak

accelerations during mechanical shaker testing. Irrespective of placement location, potential factors that may cause inter-monitor differences in raw acceleration between the GAwrist and AGhip include differences in microelectromechanical sensors, dynamic ranges and proprietary filtering processes used to minimise signal distortion during initial analogue-to-digital conversion (John & Freedson, 2012; John et al., 2013). Therefore, the current generation of accelerometry-based monitors may not be directly compared with each other even at the raw acceleration level, due to the discrepancies in how the raw data are collected and filtered. Further research and/or discussions are required to achieve the "true" harmonization of raw data collected from different types of devices.

A common outcome in child PA research is time spent in MVPA which is used to identify the number of children meeting the PA guidelines (i.e. at least 60 min of MVPA per day) (Chief Medical Officers, 2011). To complicate comparisons further between GAwrist and AGhip, accelerometer data are commonly analysed using a broad range of intensity thresholds leading to widely varying estimates of MVPA within and between studies (Guinhouya, Samouda, & de Beaufort, 2013; Routen, Upton, Edwards, & Peters, 2012). For example, Schaefer, Nace, and Browning (2014) found that estimates of wrist derived MVPA decreased by 27% (from 308 to 225 minutes) when the MVPA cut-point threshold was increased from 3 METs to 4 METs. The difference in MVPA levels between GAwrist and AGhip within this study and between other studies highlights the influence of device and wear location on MVPA prevalence, and the challenge of comparing MVPA data between studies using different intensity thresholds and devices worn at different body locations. Rowlands et al. (2015) found that applying a population specific correction factor to the GAwrist data removed the significant difference in accelerations between GAwrist and AGhip data. This method may therefore be an appropriate way of improving the comparability of findings between studies using different device brands and placement locations in the future.

This is the first study to examine the comparability of GAwrist and AGhip derived LPA and MVPA throughout the segmented week. The study observed differential agreement between GAwrist and AGhip. Agreement differed according to PA intensity and time of day, with the greatest difference

occurring in LPA during school hours. Future studies should therefore be cautious when comparing PA data derived from GAwrist and AGhip, especially studies investigating children's school day PA and segmented days. PA levels were derived from raw acceleration data and were processed and analysed using the same open-source procedures, which adds transparency and consistency to the data. However, the results of this study were performed in a relatively small sample of children living in a highly deprived area of England, which limits the generalisability of findings to other locations and populations. Device wear time was greater for the GAwrist compared to the AGhip which may have contributed to the observed differences in PA levels. The inclusion criteria used in this study for whole-day device wear is consistent with recommendations and common practices, but we did not apply wear time criteria to specific time segments (e.g., before-school). This may have biased the PA outcomes for individual segments depending on segment wear time.

Conclusion

In conclusion, PA levels from the GAwrist and AGhip are not comparable under free-living conditions. PA levels derived using raw data processing procedures were significantly higher for GAwrist compared with those for AGhip during all time segments. The magnitude of these differences was greatest during school hours and in LPA. Comparisons of raw data assessed by different monitors worn at the wrist and hip in children should therefore be undertaken with caution. We recommend the development of PA level correction factors to aid comparison of findings between studies using the GAwrist and AGhip.

References

Brooke, H. L., Corder, K., Atkin, A. J., & van Sluijs, E. M. F. (2014). A Systematic Literature Review with Meta-Analyses of Within and Between-Day Differences in Objectively Measured Physical Activity in School-Aged Children. *Sports Medicine*, *44*, 1427–1438.

Butte, N. F., Ekelund, U., & Westerterp, K. R. (2012). Assessing Physical Activity Using Wearable Monitors: Measures of Physical Activity. *Medicine & Science in Sports & Exercise*, *44*(1 Suppl), S5–S12.

Cain, K. L., Sallis, J. F., Conway, T. L., Van Dyck, D., & Calhoon, L. (2013). Using Accelerometers in Youth Physical Activity Studies: A Review of Methods. *Journal of Physical Activity and Health*, *10*, 437-450.

Chief Medical Officers. (2011). *Start Active. Stay Active. A report on physical activity for health from the four home countries.* London: Department of Health.

Corder, K., Sharp, S. J., Atkin, A. J., Andersen, L. B., Cardon, G., Page, A., … Sardinha, L. B., Ekelund, U., & Van Sluijs EM, on behalf of the International Children's Accelerometry Database (ICAD) Collaborators. (2016). Age-related patterns of vigorous-intensity physical activity in youth: the International Children's Accelerometry Database. *Preventive Medicine Reports*, *4*, 17-22.

Cooper, A. R., Goodman, A., Page, A. S., Sherar, L. B., Esliger, D. W., van Sluijs, E. M., … Ekelund, U. (2015). Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). *International Journal of Behavioral Nutrition and Physical Activity*, *12*, 113.

da Silva, I. C. M., van Hees, V. T., Ramires, V. V., Knuth, A. G., Bielemann, R. M., Ekelund, U., … Hallal, P. C. (2014). Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. *International Journal of Epidemiology*, *43*(6), 1959–1968.

de Vries, S. I., van Hirtum, W. J. E. M., Bekker, I., Hopman-Rock, M., Hirasing, R. A., & van Mechellen, W. (2009). Validity and Reproducibility of Motion Sensors in Youth: A Systematic Update. *Medicine & Science in Sports & Exercise*, *41*(4), pp. 818–827.

Edwardson, C. L., Harrington, D. M., Yates, T., Bodicoat, D. H., Khunti, K., Gorely, T., … Davies, M. J. (2015). A cluster randomised controlled trial to investigate the effectiveness and cost effectiveness

of the ‘Girls Active’ intervention: a study protocol. *BMC Public Health*, 15, 526.

Ellis, K., Kerr, J., Godbole, S., Lanckriet, G., Wing, D., & Marshall, S. (2014). A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiological Measurement*, *35*, 2191–2203.

Ellis, K., Kerr, J., Godbole, S., Staudenmayer, J., & Lanckriet, G. (2016). Hip and Wrist Accelerometer Algorithms for Free-Living Behavior Classification. *Medicine & Science in Sports & Exercise*, *48*(5), 933–940.

Fairclough, S. J., Beighle, A., Erwin, H., & Ridgers, N. D. (2012). School day segmented physical activity patterns of high and low active children. *BMC Public Health*, *12*, 406.

Fairclough, S. J., Noonan, R. J., Rowlands, A. V., van Hees, V., Knowles, Z., & Boddy, L. M. (2016). Wear compliance and activity in children wearing wrist and hip mounted accelerometers. *Medicine and Science in Sports and Exercise*, *48*(2), 245–253.

Guinhouya, B. C., Samouda, H., & de Beaufort, C. (2013). Level of physical activity among children and adolescents in Europe: a review of physical activity assessed objectively by accelerometry. *Public Health*, *127*, 301-311.

Hanggia, J. M., Phillips, L. R. S., & Rowlands, A. V. (2013). Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *Journal of Science and Medicine in Sport*, *16*(1), 40–44.

Hildebrand, M., Van Hees, V. T., Hansen, B. H., & Ekelund, U. (2014). Age-group comparibility of raw accelerometer output from wrist- and hip worn monitors. *Medicine and Science in Sports and Exercise*, *46*(9), 1816–24.

John, D., & Freedson, P. (2012). ActiGraph and Actical Physical Activity Monitors: A Peek under the Hood. *Medicine & Science in Sports & Exercise*, *44*(1 Suppl), S86–S89.

John, D., Sasaki, J., Staudenmayer, J., Mavilia, M., & Freedson, P. (2013). Comparison of raw acceleration from the GENEA and ActiGraph GT3X+ activity monitors. *Sensors*, *13*(11), 14754–63.

Keane, E., Kearney, P. M., Perry, I. J., Browne, G. M., & Harrington, J. M. (2014). Diet, physical activity, lifestyle behaviours, and prevalence of childhood obesity in Irish children: the Cork Children's Lifestyle Study Protocol. *JMIR Research Protocols*, *3*(3), e44.

Kim, Y., Lee, J-M., Peters, B. P., Gaesser, G. A., & Welk, G. J. (2014). Examination of Different Accelerometer Cut-Points for Assessing Sedentary Behaviors in Children. *PLoS ONE*, *9*(4), e90630.

Lohman, T. G., Roche, A. F., & Martorell, R. (1988). *Anthropometric Standardization Reference Manual*. Champaign, IL: Human Kinetics.

Mattocks, C., Ness, A., Leary, S., Tilling, K., Blair, S. N., Shield, J., … Riddoch, C. (2008). Use of Accelerometers in a Large Field-Based Study of Children: Protocols, Design Issues, and Effects on Precision. *Journal of Physical Activity & Health*, *5*(Suppl 1), S98-S111.

Phillips, L. R. S., Parfitt, G., & Rowlands, A. V. (2013). Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *Journal of Science and Medicine in Sport*, *16*(2), 124–128.

Robusto, K. M., & Trost, S. G. (2012). Comparison of three generations of ActiGraph™ activity monitors in children and adolescents, *Journal of Sports Sciences*, *30*(13), 1429-1435.

Rosenberger, M. E., Haskell, W. L., Albinali, F., Mota, S., Nawyn, J., and Intille, S. (2013). Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Medicine & Science in Sports & Exercise*, *45*(5), 964–75.

Routen, A. C., Upton, D., Edwards, M. G., & Peters, D. (2012). Discrepancies in accelerometer-measured physical activity in children due to cut-point non-equivalence and placement site, *Journal of Sports Sciences*, *30*(12), 1303-1310.

345 Rowlands, A. V., Fraysse, F., Catt, M., Stiles, V. H., Stanley, R. M., Eston, R. G., & Olds, T. S. (2015).
346 Comparability of measured acceleration from accelerometry-based activity monitors. *Medicine &*
347 *Science in Sports & Exercise*, *47*(1), 201-210.

348 Rowlands, A. V., Rennie, K., Kozarski, R., Stanley, R. M., Eston, R. G., Parfitt, G. C., & Olds, T. S.
349 (2014). Children's physical activity assessed with wrist- and hip-worn accelerometers. *Medicine &*
350 *Science in Sports & Exercise*, *46*(12), 2308-2316.

351 Rowlands, A. V, & Stiles, V. H. (2012). Accelerometer counts and raw acceleration output in relation
352 to mechanical loading. *Journal of Biomechanics*, *45*, 448-454.

353 Saint-Maurice, P. F., Kim, Y., Welk, G. J., & Gaesser, G. A. (2016). Kids are not little adults: what
354 MET threshold captures sedentary behavior in children? *European Journal of Applied Physiology*,
355 *116*(1), 29-38.

356 Schaefer, C. A., Nace, H., & Browning R. (2014). Establishing Wrist-Based Cutpoints for the Actical
357 Accelerometer in Elementary School-Aged Children. *Journal of Physical Activity & Health*, *11*, 604-
358 613.

359 Sherar, L., Griew, P., Esliger, D., Cooper, A., Ekelund, U., Judge, K., & Riddoch C. (2011).
360 International children's accelerometry database (ICAD): Design and methods. *BMC Public Health*, *11*,
361 485.

362 Tudor-Locke, C., Barreira, T. V., & Schuna, Jr. J. M. (2015). Comparison of step outputs for waist and
363 wrist accelerometer attachment sites. *Medicine & Science in Sports & Exercise*, *47*(4), 839–842.

364 van Hees, V. T., Gorzelniak, L., Leon, E. C. D, Eder, M., Pias, M., Taherian, S., … Brage, S. (2013).
365 Separating movement and gravity components in an acceleration signal and implications for the
366 assessment of human daily physical activity. *PLoS One*, *8*(4), e61691.

367 Wake, M., Clifford, S., York, E., Mensah, F., Gold, L., Burgner, D., & Davies, S. (2014). Introducing

368 Growing Up in Australia's Child Health CheckPoint: A physical health and biomarkers module for the
369 Longitudinal Study of Australian Children. *Family Matters*, *95*, 15-23.

370 Welk, G. J., McClain, J., & Ainsworth, B. E. (2013). Protocols for evaluating equivalency of
371 accelerometry-based activity monitors. *Medicine & Science in Sports & Exercise*, *44*(Suppl 1), S39-49.

Table 1 Physical activity outcomes for GAwrist and AGhip for weekday and weekend day segments

	GAwrist				AGhip				GAwrist - AGhip segment difference
	Mean minutes	95% CI	% segment time	95% CI	Mean minutes	95% CI	% segment time	95% CI	%
LPA									
Whole week	306.8	291.3 - 322.3	29.2	27.9 - 30.9	128.8	118.3 - 139.5	12.3	11.3 - 13.4	16.9***
Whole weekday	329.9	316.6 - 343.3	31.5+++	30.3 - 33.0	134.7	124.4 - 145.1	12.9+	11.8 - 13.9	18.6***
Before school	34.8	31.6 - 38.0	23.3	21.1 - 25.6	18.3	16.0 - 20.5	12.3	10.8 - 13.8	11.0***
During school	165.5	160.1 - 173.0	42.6‡‡‡	40.6 - 44.5	61.7	56.4 - 67.4	15.9‡‡‡	14.6 - 17.3	26.7***
After school	129.6	121.7 - 137.5	25.5	23.7 - 27.2	54.7	49.3 - 60.1	10.8	9.8 - 11.9	14.7***
Whole weekend day	283.6	265.9 - 301.3	27.1	25.4 - 28.8	122.9	112.1 - 133.8	11.7	10.7 - 12.8	15.4***
Morning	58.3	48.4 - 68.1	17.7	14.7 - 20.7	24.5	20.5 - 28.6	7.4	6.2 - 8.7	10.3***
Afternoon-evening	225.4	213.0 - 37.8	31.4†††	29.6 - 33.1	98.3	88.5 - 108.2	13.7†††	12.3 - 15.1	17.7***
MVPA									
Whole week	30.0	27.3 - 32.8	2.9	2.6 - 3.2	16.5	14.5 - 18.5	1.6	1.4 - 1.8	1.3***
Whole weekday	31.9	29.7 - 34.2	3.0+	2.8 - 3.3	18.7	17.2 - 20.2	1.8+++	1.7 - 2.0	1.2***
Before school	2.4	2.0 - 2.8	1.6	1.4 - 1.9	1.7	1.4 - 2.0	1.1	0.9 - 1.3	0.5***
During school	16.7	15.3 - 18.0	4.3‡‡‡	3.9 - 4.7	9.8	8.8 - 10.7	2.5‡‡‡	2.2 - 2.7	1.8***
After school	12.7	11.4 - 14.1	2.5	2.2 - 2.8	7.2	6.2 - 8.3	1.4	1.2 - 1.6	1.1***
Whole weekend day	28.1	24.8 - 31.4	2.7	2.3 - 3.0	14.2	11.8 - 16.7	1.4	1.1 - 1.6	1.3***
Morning	5.9	4.1 - 7.6	1.8	1.3 - 2.3	3.3	2.2 - 4.5	1.0	0.7 - 1.4	0.8***
Afternoon-evening	22.2	19.1 - 25.2	3.1†††	2.7 - 3.5	10.9	8.8 - 12.9	1.5††	1.2 - 1.8	1.6***

Significantly different between GAwrist % segment and AGhip % segment at ***p<0.001. Significantly different between GAwrist % weekday and % weekend day at +p<0.05, +++p<0.001. Significantly different between AGhip % weekday and % weekend day at +p<0.05, +++p<0.001. Significantly different between GAwrist % before school – % during school – % after school at ‡‡‡p<0.001. Significantly different between AGhip % before school – % during school – % after school at ‡‡‡p<0.001.Significantly different between GAwrist % weekend morning and % afternoon-evening at †††p<0.001. Significantly different between AGhip % weekend morning and % afternoon-evening at †† p<0.01, †††p<0.001.

Figure caption

Figure 1 Bland–Altman plots displaying agreement between GAwrist and AGhip derived whole weekday and whole weekend day LPA and MVPA. Note that the observed positive bias indicates that GAwrist values were higher than AGhip values. Horizontal lines represent mean bias and 95% limits of agreement.