Title: *Wear compliance and activity in children wearing wrist and hip mounted accelerometers*

Stuart J. Fairclough[1,2], Robert Noonan[3], Alex V. Rowlands[4,5], Vincent van Hees[6], Zoe Knowles[3], Lynne M. Boddy[3]

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

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

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

[4]Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, UK

[5]NIHR Leicester-Loughborough Diet, Lifestyle and Physical Activity Biomedical Research Unit, UK

[6]MoveLab, Physical Activity and Exercise Research, Institute of Cellular Medicine, Newcastle University, Newcastle, UK

Corresponding author: Professor Stuart Fairclough

Department of Sport and Physical Activity,

Edge Hill University,

Ormskirk,

Lancs

L39 4QP, UK

Tel: 0044 1695 584143

Fax: 0044 1695 579997

Email: stuart.fairclough@edgehill.ac.uk

ABSTRACT

Purpose. This study aimed to (i) explore children's compliance to wearing wrist and hip-mounted accelerometers, (ii) compare children's physical activity (PA) derived from wrist and hip raw accelerations, and (iii) examine differences in raw and counts PA measured by hip-worn accelerometry.

Methods. One hundred and twenty nine 9-10 y old children wore a wrist-mounted GENEActiv accelerometer (GAwrist) and a hip-mounted ActiGraph GT3X+ accelerometer (AGhip) for 7 d. Both devices measured raw accelerations and the AGhip also provided counts-based data.

Results. More children wore the GAwrist than the AGhip regardless of wear time criteria applied ($p<.001$ - .035). Raw data signal vector magnitude (SVM; r = .68), moderate PA (MPA; r = .81), vigorous PA (VPA; r = .85), and moderate-to-vigorous PA (MVPA; r = .83) were strongly associated between devices ($p<.001$). GAwrist SVM (p = .001), MPA (p = .037), VPA (p = .002), and MVPA (p = .016) were significantly greater than AGhip. According to GAwrist raw data, 86.9% of children engaged in at least 60 min MVPA·d^{-1}, compared to 19% for AGhip. ActiGraph MPA (raw) was 42.00 ± 1.61 min·d^{-1} compared to 35.05 ± 0.99 min·d^{-1} (counts) (p=.02). Actigraph VPA was 7.59 ± 0.46 min·d^{-1} (raw) and 37.06 ± 1.85 min·d^{-1} (counts; p=.19).

Conclusion. In children accelerometer wrist placement promotes superior compliance than the hip. Raw accelerations were significantly higher for GAwrist compared to AGhip, possibly due to placement location and technical differences between devices. AGhip PA calculated from raw accelerations and counts differed substantially, demonstrating that PA outcomes derived from cutpoints for raw output and counts cannot be directly compared.

Keywords: raw accelerations, wear time, physical activity, GENEActiv, ActiGraph GT3X+

INTRODUCTION

Accelerometry is the most widely used objective method of assessing children's free-living physical activity (PA) (1). Accelerometers allow accelerations to be quantified, and in the context of PA research the accelerometer outcome is related to a measure of energy expenditure (12) or PA behaviour (17). Traditionally, accelerometers have been worn on the hip as this location is thought to provide the most accurate estimations of energy expenditure and activity intensity (26). Recently there has been an increased use of wrist-worn devices, which it has been argued, promote better compliance to device wear. In the NHANES 2011-12 data collection cycle using wrist-worn accelerometers, median wear time duration was 21-22 hours per day, which was up to 100% longer than in previous cycles using hip-worn devices (28). Compared to hip-worn accelerometers, those worn on the wrist may be perceived as less burdensome to research participants, thus promoting wear-time compliance (21, 39). Variable compliance to accelerometer monitoring protocols influences the application of minimum wear time criteria (i.e., number of minutes wear that constitutes a 'valid' day of measurement and the minimum number of days required for a reliable estimate of PA levels), which are subject to variation in researcher decisions about how 'non-wear' time is defined (33). Better compliance gives greater confidence that PA data are representative of actual daily PA due to the association between duration of monitoring and reliability of PA data (16). Presently though, there is limited evidence of the extent of improved compliance in children wearing accelerometers on the wrist.

The growing popularity of the wrist as the accelerometer placement site warrants comparisons with PA data derived from devices worn on the hip, which has traditionally been the most commonly used site. Recently, PA intensity cutpoints derived from raw acceleration output have been developed in the same study for the GENEActiv (Activinsights, Cambs, UK) and ActiGraph GT3X+ (ActiGraph, Pensacola, FL) accelerometers, which are designed

for wear both on the wrist and hip (12). Using these protocol-specific cutpoints together may help improve our understanding of how concurrent estimates of PA intensity from the wrist and hip sites compare. This move towards raw acceleration signal processing is a recent advance in accelerometer-based PA monitoring, which has traditionally used accelerometer output reduced to 'counts'. Direct comparison of PA outcomes derived from different devices has not previously been possible due to differences in proprietary algorithms used to collect, process, filter, and scale raw signal data to produce the device-specific counts (3, 40). This lack of equivalency between devices and therefore comparability between studies using different devices, has led to the emergence of accelerometers such as the GENEActiv range and ActiGraph GT3X+ and GT9X. These devices are capable of collecting and recording raw, unfiltered accelerations which can then be subject to researcher-driven data processing procedures (40). Basing PA data on raw accelerations provides an opportunity to improve comparability between studies using different devices, and promote transparency and consistency of post-data collection analytical processes (12). Presently though, limited published research is available describing children's free-living PA derived from raw accelerometer data. One study involving 47, 1st to 5th grade children wearing GENEActiv accelerometers on the wrist reported mean daily MVPA and VPA of 308.2 min and 32.7 min, respectively (31). In a sample of 58 Australian 10-12 year olds, MVPA derived from GENEActiv raw data was 67.8 min·d^{-1} (hip) and 98.2 min·d^{-1} (wrist) with VPA recorded as 11.1 min·d^{-1} (hip) and 16.7 min·d^{-1} (wrist) (28). These studies however, calculated the signal vector magnitude values differently (i.e., averaging vs. summing raw accelerations per epoch), and used different PA intensity cutpoints (23, 31), which makes direct comparison of findings challenging. Another important issue is that historical accelerometer data used counts and extensive validation work has been conducted on counts-based accelerometer data (9, 17, 24, 35). Although the 'cutpoint conundrum' exists, there has been some consensus in

recent years for using the cutpoints of Evenson et al. (9), which have convincing evidence of validity in children (34). These cutpoints therefore provide a basis for free-living comparison with more contemporary cutpoints based on raw accelerations (12, 23, 31).

As the field moves more towards utilisation of raw data processing and the availability of wrist-worn devices increases, studies reporting the comparability of PA outcomes based on raw accelerations and counts from both wrist and hip are warranted. Therefore, the aims of this study were (i) to explore children's compliance to wearing wrist and hip-mounted accelerometers during free-living, (ii) to compare children's PA derived from raw acceleration signals of wrist and hip worn accelerometers, and (iii) to examine differences in PA estimated from raw data with that from counts data measured by a hip worn accelerometer.

METHODS

Participants. The participants were 129 Year 5 (9-10 y) children (79 girls) from six primary schools in Liverpool, England. Following ethical approval from the University Research Ethics Committee, all Year 5 children (n = 326) in participating schools were invited to participate. They received a pack which contained parent and child information sheets, consent and assent forms, and a medical screening form. Written informed consent and assent was received from parents and their children, respectively before children could participate in the study.

Anthropometrics. Stature and sitting stature were assessed to the nearest 0.1cm using a portable stadiometer (Leicester Height Measure, Seca, Birmingham, UK). Body mass was assessed to the nearest 0.1kg (Seca, Birmingham, UK). Body mass index (BMI) was calculated for each participant with BMI z-scores also assigned (4). Age and sex specific BMI cut points were used to classify children as normal weight or overweight/obese (5).

Gender-specific regression equations (20) were used to predict children's age from peak height velocity (APHV), which is a proxy measure of biological maturation. All measurements were taken by the second author and a research assistant using standard procedures.

Socio-economic status. Neighbourhood-level socio-economic status (SES) was calculated using the 2010 Indices of Multiple Deprivation (IMD) (7). The IMD is a UK Government produced measure comprising seven areas of deprivation (income, employment, health, education, housing, environment, and crime). Deprivation scores were generated using the National Statistics Postcode Directory database from parent reported home postcodes. Higher SES was represented by lower IMD scores.

Physical Activity. Free-living PA was assessed using the GENEActiv Original triaxial accelerometer (Activinsights, Cambs, UK) worn on the non-dominant wrist (GAwrist) and the ActiGraph GT3X+ triaxial accelerometer (ActiGraph, Pensacola, FL) worn on the right hip (AGhip). The GENEActiv can be worn on the wrist, upper arm, hip, chest, ankle, and thigh, has a dynamic range of ± 8 g, and is a valid measure of PA in children (12, 23, 31). The GENEActiv was selected because it measures raw accelerations and is typically worn on the wrist (30, 31, 38, 39). ActiGraph accelerometers have been used in PA research for around 20 years and have been validated on several occasions with children (9, 19, 24, 35). The GT3X+ model has a dynamic range of ± 6 g, and can be worn on the hip, ankle, wrist, and thigh. The ActiGraph was selected as it is the most commonly used accelerometer in children's PA research, and though it is being worn on the wrist in the most recent NHANES data collection cycles (28), traditionally it has been worn on the hip (26). The GT3X+ has the capability to generate raw acceleration and count data to enable straightforward backwards interpretation of data in either format. Both devices were initialised to record raw accelerations at a frequency of 100 Hz, and participants were asked to wear the monitors at

all times for 7 consecutive days except when sleeping and engaging in water based activities (e.g., bathing, swimming). Data collection took place during the regular school term from January to May 2014 so activities were representative of usual free-living activities. After 7 days 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, and to AGD format for analysis of counts data. 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 autocalibrated the raw triaxial accelerometer signals (37) and converted them 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 has previously been referred to as the Euclidean norm minus one (ENMO) (38). Raw data were further reduced by calculating the average SVM values per 1-s epoch expressed in $mg \cdot s^{-1}$ over each of the 7 monitored days.

AGhip and GAwrist raw data wear times 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 (38). A time window was classified as nonwear 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 (30). This approach has been applied previously in studies using both devices worn at the wrist and hip (27, 28, 38). For ActiGraph counts data, non-wear is conventionally determined from accumulated pre-determined time periods of consecutive zero counts. To address study

aim 3, and in keeping with previous work (10, 25), the 1-s epoch AGhip counts data non-wear time was defined as at least 20 min periods of consecutive zero counts (2).

Raw acceleration outcome variables for AGhip and GAwrist were average gravity-based SVM (mg), and min of MPA, VPA, and MVPA which were calculated using device and location-specific cutpoints based on the ENMO metric (12). These were 142.6 mg (MPA) and 464.6 mg (VPA) for AGhip, and 191.6 mg (MPA) and 695.8 mg (VPA) for GAwrist (12). Comparing PA values based on ENMO-derived SVM was important as this metric was applied to ActiGraph GT3X+ and GENEActiv data in the same calibration study (12). For analysis of raw acceleration and counts-based PA levels, inclusion criteria were at least 10 $h \cdot day^{-1}$ wear time for at least three days, including a minimum of one weekend day. This resulted in analytical samples of 84 participants for the GAwrist vs. AGhip raw data analyses, and 65 participants for the AGhip raw vs. counts data analyses. Outcome variables for AGhip counts data were min of MPA, VPA, and MVPA which were classified according to empirical cutpoints (9) that have demonstrated acceptable classification accuracy across a range of intensities in children (34). Presently, no published sedentary time cutpoints exist for GAwrist and AGhip raw accelerations calculated using the ENMO approach. For this reason we did not investigate differences in sedentary time and light intensity PA.

Analysis. Kolmogorov-Smirnov tests confirmed that raw PA outcome data for the overall week and week days were normally distributed but that weekend GAwrist SVM and VPA, weekend AGhip SVM, MVPA, and VPA, and AGhip counts data had skewed distributions ($p<.05$). Following log (SVM, MVPA), square root (VPA), and reciprocal (AGhip counts MPA, VPA, MVPA) transformations, data were normalized and included for analyses. All transformed data were back-transformed for presentation purposes. To analyse compliance

(study aim 1), mean daily valid wear time and number of valid days were calculated for GAwrist and AGhip raw data. Paired samples McNemar's tests and t-tests assessed compliance and wear time differences against differing wear time criteria. To address study aim 2, partial Pearson correlation analyses assessed raw data relationships between devices for SVM, MPA, VPA, and MVPA, while controlling for the effects of wear time. Bland-Altman plots were constructed to assess agreement between device raw data outputs, and repeated measures ANCOVAs compared raw data PA outcomes between AGhip and GAwrist for the whole week, week days, and weekend days. For aim 3, repeated measures ANCOVAs examined differences between whole week reciprocal transformed MPA, VPA, and MVPA derived from AGhip raw and from counts data. In each ANCOVA adjustment was made for device wear time and sex. Statistical significance was set to $p<.05$. All analyses were conducted using IBM SPSS Statistics version 22 (IBM, Armonk, NY).

RESULTS

Descriptive characteristics of the participants are displayed in Table 1. Around three-quarters of the children were of healthy weight which is typical for Liverpool but somewhat lower than the English national average. Boys and girls were similarly aged but girls were more advanced than boys in regards to somatic maturation. IMD scores indicated that participants resided in some of the lowest SES neighbourhoods in England.

TABLE 1 HERE

Raw data device compliance

AGhip and GAwrist data were available for 115 and 128 children, respectively. Instances of device malfunction (n=1), software errors (n=5), and accelerometer non-wear (n=8) accounted for the modest data attrition. The percentage of children that wore each device for between 6 and 12 $h \cdot d^{-1}$ on 1 to 7 d is presented in the Supplemental Digital Content (see

223 Table, Supplemental Digital Content 1). Over 95% of children wore the AGhip and GAwrist
224 for at least 12 h on a single day. Irrespective of the number of monitoring days, the
225 percentage of children wearing both devices decreased with hours of wear, and this drop-off
226 was more prominent for the AGhip. For example, the difference in the proportion of children
227 wearing the AGhip for 6 h over 3 days and those wearing it for 12 h over 3 d was -18.3%,
228 compared to -5.8% for the GAwrist. Ten h wear time over at least 2 d has been demonstrated
229 to provide reliable estimates of PA in population studies of older primary school aged
230 children (25). Taking 10 h wear time as the criterion for a valid day, the decrease in children
231 wearing the AGhip for between 1 and 7 d was 80.5%, in comparison to 62.0% for the
232 GAwrist. A similar trend was observed when the inclusion of at least one weekend day was
233 considered. With inclusion criteria of a minimum of 10 h wear on at least 3 weekdays plus a
234 minimum of one weekend day, GAwrist non-compliance (16.4%) was lower than for the
235 AGhip (25.2%).
236
237 When the number of children classified as 'included' as defined by commonly used wear
238 time criteria (25) were analysed, significantly more children achieved wear time criteria when
239 wearing the GAwrist than the AGhip for at least 9 $h \cdot d^{-1}$ (p=.002) and 10 $h \cdot d^{-1}$ (p=.035) on any
240 4 d of the week (Table 2). When a weekend day was included in the criteria this level of
241 compliance was achieved by significantly more children wearing the GAwrist than the AGhip
242 for either 9 $h \cdot d^{-1}$ or 10 $h \cdot d^{-1}$ over 2, 3, and 4 week days (p=.001-.002). Average daily wear
243 time across the different wear time criteria ranged from 15.57 to 15.82 $h \cdot d^{-1}$ for the GAwrist,
244 and 14.18 to 14.21 $h \cdot d^{-1}$ for the AGhip. GAwrist daily wear time was significantly higher
245 than for the AGhip, regardless of wear time criteria applied (p<.001). Children wore the
246 GAwrist for significantly more days than the AGhip. When a valid day was defined as at
247 least 9 h wear, the GAwrist was worn for 5.8 d out of 7 d compared to 5.1 d for the AGhip

(p<.001), and for 5.6 d versus 4.9 d when 10 h wear was the criterion (p<.001). During weekdays the GAwrist was worn for 4.2 d (9 h) and 4.1 (10 h) in comparison to 3.8 d (p<.001) and 3.7 d (p<.001) respectively, for the AGhip. The GAwrist was also worn most at weekends when valid day minimum wear was set to 9 and 10 h (GAwrist: 1.6 d and 1.5 d, respectively; AGhip: 1.3 d and 1.2 d, respectively; p<.001).

TABLE 2 HERE

Raw data physical activity levels

Significant partial correlations between raw data PA outcomes confirmed that after adjustment for wear time, SVM (r = .68), MPA (r = .81), VPA (r = .85), and MVPA (r = .83) were moderately to strongly associated between devices (p<.001). Bland-Altman plots are presented in Figure 1A-D and show that the extent of differences in SVM, MPA, VPA, and MVPA between GAwrist and AGhip increased linearly with children's levels of PA engagement. Correlation coefficients between the mean of the measures and the bias were r = .75 (SVM), r = .64 (MPA), r = .75 (VPA), and r = .69 (MVPA), indicating that the 95% limits should be treated with caution.

FIGURE 1A-D HERE

Comparisons of PA levels between devices are presented in Table 3. Wear time and sex-adjusted SVM values during the whole week, weekdays, and weekend days were significantly higher for the GAwrist than the AGhip (p=.001). MPA recorded by the GAwrist on weekdays, weekend days, and over the whole week was 45.2% (p=.07), 41.1% (p=0.1), and 44.2% (p=.04) greater respectively, than values derived from the AGhip. GAwrist VPA was also significantly higher than AGhip at the different times of the week (p=.02 - .001), with the greatest difference of 54.7% occurring at weekends. MVPA was 43.3-45.7% greater for the GAwrist than the AGhip across the whole week, week days, and weekend days.

According to the GAwrist raw data, 86.9% of children engaged in at least 60 min MVPA·d^{-1}, compared to 19% according to AGhip-derived MVPA.

TABLE 3 HERE

Physical activity levels from AGhip raw and counts data

Analyses of raw and counts data for AGhip revealed that children's adjusted whole week MPA (raw) was 42.00 ± 1.61 min·d^{-1} compared to 35.05 ± 0.99 min·d^{-1} (counts) (p=.02), a difference of 16.5% (Figure 2). Adjusted VPA differed by 79.5% between counts (37.06 ± 1.85 min·d^{-1}) and raw data (7.59 ± 0.46 min·d^{-1}; p=.19). These combined MPA and VPA differences were reflected in overall MVPA (72.11 ± 2.60 min·d^{-1} [counts] vs. 49.59 ± 2.01 min·d^{-1} [raw]; p=.57). The recommended 60 min·d^{-1} of MVPA was achieved by 20.2% and 67.7% of children with valid raw and counts data, respectively.

FIGURE 2 HERE

DISCUSSION

In 2009 experts in PA measurement recommended that researchers' estimations of PA should in future be based on raw acceleration data rather than proprietary movement counts (11). Since then raw accelerometer data have been reported more frequently, but still much less often than counts data. This study adds to the raw accelerometer data evidence base, as it is the first to examine children's compliance to wrist and hip-worn devices, between-device differences in PA intensities derived from raw accelerations, and differences in hip-mounted ActiGraph GT3X+ raw acceleration versus counts-based estimates of free-living PA.

Accelerometer compliance

More children wore the GAwrist than AGhip irrespective of the wear time inclusion criteria applied or time of week observed. Using the wrist as the accelerometer placement site may promote better device compliance, as illustrated by the improved wear time reported in the

297 2011-12 NHANES data collection cycle (28). There is though a paucity of research
298 investigating children's compliance to wrist and hip-worn accelerometers worn in parallel.
299 While it has been suggested that children (32) and adults (39) prefer the wrist as the device
300 placement site, such preferences may be partly dependent upon specific device features (e.g.,
301 feedback on activity (32)) and monitor-specific wear instructions (e.g., removal of hip-worn
302 devices during sleep and water-based activities (39)). This latter point is exemplified by a
303 recent examination of hip-worn ActiGraph data from 9-11 y olds across 12 countries, which
304 reported how a 24 h accelerometer wear protocol resulted in an average wear time of 22.6 h
305 (36). Thus, asking children to only remove devices for water-based activities elicits much
306 greater total wear times than are typically observed in waking time protocols. Waking wear
307 time though was 14.7 $h \cdot d^{-1}$ (36) which was similar to the AGhip values and less than the
308 GAwrist values observed in our study. These findings confirm the combined influences of
309 wear location and protocol on accelerometer wear compliance. To our knowledge no
310 previous studies have examined children's compliance to wearing wrist and hip-mounted
311 accelerometers concurrently. Our findings confirm that children's perceived acceptability of
312 and preference for wrist-worn devices (32), reflect actual wear when children were asked to
313 use two devices under the same conditions. Where feasible, future youth PA studies should
314 employ wrist-worn accelerometry to increase the likelihood of longer wear time which would
315 result in more representative and reliable estimates of PA (16). Wrist-worn devices may not
316 only result in superior compliance, but according to recent evidence, may also provide better
317 estimates of children's energy expenditure compared to hip mounted accelerometers (6). For
318 wrist-worn accelerometry to become widely adopted however, more needs to be known about
319 the comparability of children's PA levels derived from raw accelerations, with historical
320 counts-based data.

321 *PA derived from raw acceleration signals of wrist and hip worn accelerometers*

Correlations between wrist-worn GENEActiv and hip-worn ActiGraph free-living raw accelerations have not previously been reported in children. We observed moderate to strong partial correlations between AGhip and GAwrist (r = .68-.85) which were lower than the recently reported correlation of r = .93 between hip worn GENEActiv and ActiGraph GT3X+ average accelerations (27). Our findings indicate that both devices measured children's free-living accelerations which explained almost 70% of the shared variance in MVPA. Notwithstanding these strong associations, there were considerable differences between devices in average SVM and the derived outcomes (time spent in MPA, VPA, and MVPA). GAwrist values were consistently higher than those from the AGhip, particularly at higher intensities. These differences were most extreme for SVM values (~60%) which were calculated for both devices using identical data processing methods. In the only previous study to compare children's raw GAwrist and AGhip data using the ENMO data processing approach, GAwrist SVM was significantly higher for a range of moderate-to-vigorous activities performed during a controlled device calibration protocol (i.e., fast walking, stepping, running, and circuit training) (12). Moreover, in agreement with our MPA and VPA results, greater relative differences between AGhip and GAwrist SVM values were observed as activity intensity increased (12). Similar differences between devices worn at the same site have previously been reported in adults as well as children regardless of analytical approaches used to generate raw accelerations (15, 27, 28). During vigorous ambulatory activities such as fast running, higher accelerations at the wrist relative to the hip may be observed due to greater shoulder muscle activity, compared to during walking and slow running, when arm swing and resultant wrist accelerations are more passive (29). Moreover, wrist accelerations will be disproportionately greater than those of the hip for certain types of movements that may occur regularly during children's free-living activity (e.g., some sports, computer gaming, homework), and for example among children who gesticulate vigorously

(28). This 'decoupling' of wrist and hip accelerations may also occur in reverse (e.g., walking with hands in pockets) and is likely population-specific (28). We did not record the children's activity modes but it may be feasible that their daily activities involved a disproportionate volume of 'pro-wrist' decoupling of wrist and hip accelerations, which contributed to higher GAwrist values.

Although device placement location is arguably the most obvious reason why PA outcomes differed to the extent that they did, the strong inter-device associations between outcomes suggest that placement was not the only reason. Raw acceleration data from each device were used to generate the PA outcomes, but data cannot be considered equivalent (40), as raw accelerations for the GENEActiv have been observed to be greater than those for the ActiGraph GT3X+ when worn at the same site in controlled and free-living conditions (15, 27, 29). For example, during mechanical shaker testing GENEActiv peak accelerations were up to 7.4% greater than ActiGraph GT3X+ with differences increasing in line with shaker acceleration magnitude (15). Similarly, average GENEActiv high-pass filtered accelerations were recently observed to be over 10% greater than ActiGraph GT3X+ accelerations when both devices were worn at the hip during children's free-living activities (27). Technical differences between devices, such as the micro-electro-mechanical sensors used and their dynamic ranges, reference voltage, analogue-to-digital conversion rate, and ActiGraph's proprietary data filtering processes (14, 15, 27), are the likely explanations of the differences in each device's acceleration outputs.

Comparison of raw and counts PA data measured by a hip-mounted accelerometer

Systematic differences in AGhip PA outcomes from raw and counts data were not observed. Raw data MPA values were 15.9% higher than counts data, but raw data VPA values were

79.6% lower than counts data. To our knowledge, no previous study has compared hip-mounted ActiGraph GT3X+ raw and counts data output in children. The closest comparison is provided by Rowlands and colleagues who compared ActiGraph GT3X+ counts data using the cutpoints of Evenson et al. (9) with GENEActiv raw data, with both devices worn at the hip (28). The comparison is based on the very strong associations between devices for MVPA measured at the hip (r=.93) (28). Rowlands et al.'s findings mirrored ours whereby raw data MPA was greater than counts data (56.7 vs. 32.3 min·d^{-1}), but was lower for VPA (11.1 vs. 30.0 min·d^{-1}) (28). The magnitude of the differences though differed somewhat, which may relate to the different raw data processing procedures and raw acceleration cutpoints (23) applied between our study and that of Rowlands and colleagues (28). It is likely that comparable raw acceleration values reported by Rowlands et al. would have been higher than those observed in our study, due to differences in raw acceleration data processing (i.e., converting acceleration negative values to their absolute, summing acceleration values per 1-s epoch) (8, 12, 23). Moreover, the PA intensity cutpoints used in both studies were derived from different calibration protocols (12, 23), which may be a more influential factor on PA outcomes than placement site or device type (28). While some inferences about output differences can be made on the basis of raw acceleration data processing, the proprietary nature of the ActiGraph GT3X+ algorithm to convert raw acceleration into counts makes similar suppositions difficult. These findings demonstrate that raw acceleration and counts data cannot be directly compared because insufficient information is available about how counts are generated. This reinforces the calls of others (13, 18, 22) for transparent raw accelerometer data processing to become the norm so as to progress the field towards equivalency of data output and better scope for comparability of findings between studies using different devices.

A strength of this study is that it is the first to assess children's free-living PA derived from raw wrist and hip accelerations using the GENEActiv and ActiGraph GT3X+ accelerometers, respectively. Further, for the first time, children's compliance to wearing these devices concurrently over a 7-d monitoring protocol has been reported. Wearing the accelerometers in parallel standardizes possible confounding variables such as the type of PA performed during the monitoring period (39). Raw acceleration data were processed and analysed using the same open source procedures which adds transparency and consistency to the data. The study sample was though limited to 9-10 y olds in a low socioeconomic area of England and our findings should be interpreted and applied with this in mind as free-living PA routines may be different for other age groups and for children from geographic locations. A further limitation is that data were collected during school term times and so may not be representative of PA during extended non-school time such as school holidays and vacations. We also did not report time spent being sedentary or in light intensity PA. Children's sedentary time and light PA are associated with various health outcomes but presently, raw acceleration thresholds for GENEActiv and ActiGraph GT3X+ based on the ENMO metric do not exist, and so we were limited to reporting MPA, VPA, and MVPA.

During free-living activity children had significantly better compliance to wearing the GAwrist than AGhip. The recognised association between duration of monitoring and reliability of PA data means that better compliance gives researchers and research users greater confidence in the PA data reported. The superior compliance of the GAwrist confirms that the wrist is a feasible accelerometer placement location in children. Raw acceleration values derived using the same data processing procedures were significantly higher for GAwrist compared to AGhip. It is unclear why these disparities occurred but it was likely a combination of the effects of placement location and technical differences between the

GENEActiv and ActiGraph GT3X+. To address this, it has been recently suggested that differences in acceleration magnitude between GENEActiv and ActiGraph GT3X could be addressed by the application of an appropriate conversion factor to make values interchangeable between devices (27). For this approach to be effective standardized data processing procedures would need to be applied to the raw acceleration data collected. AGhip PA levels calculated from raw accelerations and counts differed substantially, particularly in respect of VPA. These findings demonstrate that regardless of device placement location raw output and counts cannot be directly compared because of the lack of information about the ActiGraph proprietary filtering algorithm applied to generate counts. Raw acceleration data processing potentially enables greater transparency, and comparability between studies using the same data processing methods, though comparisons to counts-based data are limited. From a health promotion perspective, current PA guidelines are mainly based on self-report questionnaires and to a lesser extent, data from hip mounted accelerometer counts. As the use of raw acceleration data increases, examination of activity-health relationships using raw data from wrist mounted devices is warranted. We used the ENMO metric to calculate SVM but presently no SVM thresholds for children's light PA and sedentary time exist using this method. Future work should include development of these thresholds which may help enhance our understanding of the influence of device type and placement location on children's free-living raw accelerations and associated health outcomes.

ACKNOWLEDGMENTS

This study was funded by Liverpool John Moores University. We thank the teachers and children for their participation.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest. The results of the present study do not constitute endorsement by ACSM.

447 REFERENCES

448 1. Cain KL, Sallis JF, Conway TL, Van Dyck D, Calhoon L. Using accelerometers in
449 youth physical activity studies: A review of methods. *J Phys Act Health*.
450 2013;10(3):437-50.
451 2. Catellier DJ, Hannan PJ, Murray DM et al. Imputation of missing data when
452 measuring physical activity by accelerometry. *Med Sci Sports Exerc*. 2005;37(11
453 (Suppl)):S555-S62.
454 3. Chen KY, Bassett JR DR. The technology of accelerometry-based activity monitors:
455 Current and future. *Med Sci Sports Exerc*. 2006;37(11 (Suppl)):S490-S500.
456 4. Cole T, Freeman J, Preece M. Body mass index reference curves for the UK, 1990.
457 *Arch Dis Child*. 1995;73(1):25 - 9.
458 5. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for
459 child overweight and obesity worldwide: international survey. *BMJ*. 2000;320:1240-
460 4.
461 6. Crouter SE, Flynn JI, Bassett DR, Jr. Estimating physical activity in youth using a
462 wrist accelerometer. *Med Sci Sports Exerc*. 2015;47(5):944-51.
463 7. Department for Communities and Local Government. *The English Indices of*
464 *Deprivation 2010*. Wetherby: Communities and Local Government Publications;
465 2011. p. 2-20.
466 8. Esliger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the
467 GENEA Accelerometer. *Med Sci Sports Exerc*. 2011;43(6):1085-93.
468 9. Evenson K, Catellier D, Gill K, Ondrak K, McMurray R. Calibration of two objective
469 measures of physical activity for children. *J Sports Sci*. 2008;26:1557 - 65.

10. Fairclough SJ, Boddy LM, Mackintosh KA, Valencia-Peris A, Ramirez-Rico E. Weekday and weekend sedentary time and physical activity in differentially active children. *J Sci Med Sport*. 2014; doi: 10.1016/j.jsams.2014.06.005.
11. Freedson P, Bowles HR, Troiano R, Haskell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S1-4.
12. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. Age-group comparibility of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc*. 2014;46(9):1816-24.
13. Intille SS, Lester J, Sallis JF, Duncan G. New horizons in sensor development. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S24-31.
14. John D, Miller R, Kozey-Keadle S, Caldwell G, Freedson PS. Biomechanical examination of the 'plateau phenomenon' in ActiGraph vertical activity counts. *Physiol Meas*. 2012;33(2):219.
15. John D, Sasaki J, Staudenmayer J, Mavilia M, Freedson P. Comparison of Raw Acceleration from the GENEA and ActiGraph™ GT3X+ Activity Monitors. *Sensors*. 2013;13(11):14754-63.
16. Levin S, Jacobs DR, Ainsworth BE, Richardson MT, Leon AS. Intra-individual variation and estimates of usual physical activity. *Ann Epidemiol*. 1999;9:481-8.
17. Mackintosh KA, Fairclough SJ, Stratton G, Ridgers ND. A calibration protocol for population-specific accelerometer cut-points in children. *PLoS ONE*. 2012;7(5):e36919.
18. Matthews CE, Hagstromer M, Pober DM, Bowles HR. Best practices for using physical activity monitors in population-based research. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S68-76.

19. Mattocks C, Leary SD, Ness AR et al. Calibration of an accelerometer during free living activities in children. *Int J Pediatr Obes*. 2007;2:218-26.
20. Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. *Med Sci Sports Exerc*. 2002;34(4):689-94.
21. Nyberg GA, Nordenfelt AM, Ekelund U, Marcus C. Physical activity patterns measured by accelerometry in 6- to 10-yr-old children. *Med Sci Sports Exerc*. 2009;41(10):1842-8.
22. Peach D, Hoomissen JV, Callender H L. Exploring the ActiLife ® filtration algorithm: converting raw acceleration data to counts. *Physiol Meas.* 2014;35(12):2359.
23. Phillips LR, Parfitt G, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport*. 2012;16(2):124-8.
24. Puyau M, Adolph A, Vohra F, Butte N. Validation and calibration of physical activity monitors in children. *Obes Res*. 2002;10(3):150-7.
25. Rich C, Geraci M, Griffiths L, Sera F, Dezateux C, Cortina-Borja M. Quality control methods in accelerometer data processing: defining minimum wear time. *PloS One*. 2013;8(6):e67206.
26. Rosenberger ME, Haskell WL, Albinali F, Mota S, Nawyn J, Intille S. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Med Sci Sports Exerc*. 2013;45(5):964-75
27. Rowlands AV, Fraysse F, Catt M et al. Comparability of measured acceleration from accelerometry-based activity monitors. *Med Sci Sports Exerc*. 2015;47:201-10.

518 28. Rowlands AV, Rennie K, Kozarski R et al. Children's Physical Activity Assessed
519 With Wrist- and Hip-Worn Accelerometers. *Med Sci Sports Exerc*. 2014;46(12):2308-
520 16.
521 29. Rowlands AV, Stiles VH. Accelerometer counts and raw acceleration output in
522 relation to mechanical loading. *J Biomech*. 2012;45(3):448-54.
523 30. Sabia S, van Hees VT, Shipley MJ et al. Association between questionnaire and
524 accelerometer-assessed physical activity: the role of sociodemographic factors. *Am J*
525 *Epidemiol*. 2014;179(6):781-90.
526 31. Schaefer CA, Nigg CR, Hill JO, Brink LA, Browning RC. Establishing and
527 evaluating wrist cutpoints for the GENEActiv accelerometer in youth. *Med Sci Sports*
528 *Exerc*. 2014;46(4):826-33.
529 32. Schaefer SE, Van Loan M, German JB. A Feasibility Study of Wearable Activity
530 Monitors for Pre-Adolescent School-Age Children. *Prev Chronic Dis*. 2014;11:E85.
531 33. Sirard JR, Slater ME. Compliance with wearing physical activity accelerometers in
532 high school students. *J Phys Act Health*. 2009;6(Suppl 1):S148-S55.
533 34. Trost S, Loprinzi P, Moore R, Pfeiffer K. Comparison of accelerometer cut points for
534 predicting activity intensity in youth. *Med Sci Sports Exerc*. 2011;43:1360 - 8.
535 35. Trost SG, Ward DS, Moorehead SM, Watson PD, Riner W, Burke JR. Validity of the
536 Computer Science and Application (CSA) activity monitor in children. *Med Sci*
537 *Sports Exerc*. 1998;30(4):629-33.
538 36. Tudor-Locke C, Barreira TV, Schuna JM et al. Improving wear time compliance with
539 a 24-hour waist-worn accelerometer protocol in the International Study of Childhood
540 Obesity, Lifestyle and the Environment (ISCOLE). *Int J Behav Nutr Phys Activity*.
541 2015;12(1):1-9.

37. van Hees VT, Fang Z, Langford J et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. *J Appl Physiol*. 2014;117(7):738-44.
38. van Hees VT, Gorzelniak L, Dean León EC et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. *PLoS ONE*. 2013;8(4):e61691.
39. van Hees VT, Renstrom F, Wright A et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *PloS One*. 2011;6(7):e22922.
40. Welk GJ, McClain J, Ainsworth BE. Protocols for evaluating equivalency of accelerometry-based activity monitors. *Med Sci Sports Exerc*. 2012;44(1 Suppl 1):S39-49.

FIGURE CAPTIONS

Figure 1A-D. Bland-Altman plots displaying agreement between AGhip and GAwrist derived (A) SVM, (B) MPA, (C) VPA, and (D) MVPA. Note. The observed positive bias indicates that GAwrist values were higher than AGhip. Horizontal lines represent mean bias and 95% limits of agreement.

Figure 2. Whole week MPA and VPA according to AGhip counts and raw data (n = 65)

* AGhip raw MPA > AGhip counts MPA, p=.02

LIST OF SUPPLEMENTAL DIGITAL CONTENT

Supplemental Digital Content 1. Table showing percentage of children available for analyses according to daily wear time and number of wear days.pdf