This is a peer-reviewed, accepted author manuscript of the following article: Rowlands, A. V., Harrington, D. M., Bodicoat, D. H., Davies, M. J., Sherar, L. B., Gorely, T., Khunti, K., & Edwardson, C. L. (2018). Compliance of adolescent girls to repeated deployments of wrist-worn accelerometers. Medicine and Science in Sports and Exercise, 50(7), 1508-1517. https://doi.org/10.1249/MSS.0000000000001588

Compliance of adolescent girls to repeated deployments of wrist-worn accelerometers

Short title: Compliance to wrist-worn accelerometry

Rowlands, A.V.[1,2,3], Harrington, D.M.[1,2], Bodicoat, D.H[1,2,6], Davies, M.J.[1,2], Sherar, L.B.[4], Gorely, T.[5], Khunti, K.[1,2,6], Edwardson, C.L.[1,2]

1. Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, UK
2. NIHR Leicester Biomedical Research Centre, UK
3. Alliance for Research in Exercise, Nutrition and Activity (ARENA), Sansom Institute for Health Research, Division of Health Sciences, University of South Australia, Adelaide, Australia
4. School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough UK
5. Department of Nursing, School of Health, Social Care and Life Sciences, University of the Highlands and Islands, Inverness, UK
6. NIHR Collaboration for Leadership in Applied Health Research and Care East Midlands, Leicester General Hospital, UK

Corresponding author: Alex Rowlands, Diabetes Research Centre, University of Leicester, Leicester General Hospital, Leicester, LE5 4PW, UK. alex.rowlands@le.ac.uk. Tel: +44 116 258 8595

Abstract

Purpose: To determine the cross-sectional and cumulative compliance of adolescent girls to accelerometer wear at three deployment points and to identify variables associated with compliance. **Methods:** Girls from 20 secondary schools were recruited: 10 schools were participating in the 'Girls Active' intervention and 10 were control schools. Physical activity was measured using the GENEActiv accelerometer worn on the non-dominant wrist 24 hours/day for up to 7-days at baseline, 7-months and 14-months. Demographic and anthropometric characteristics were recorded. **Results:** Seven valid days (≥16 hours) of accelerometer wear were obtained from 83%, 77% and 68% of girls at baseline (n = 1734), 7-months (n = 1381) and 14-months (n = 1326), respectively. 68% provided 7-valid days for both baseline and 7-months, 59% for baseline and 14-months and 52% for all three deployment points. Estimates of physical activity level from 3-days of measurement could be considered equivalent to a 7-day measure (i.e. they fell within a ±5% equivalence zone). Cross-sectionally, 3-valid days were obtained from at least 91% of girls; cumulatively, this was obtained from ≥88% of girls across any two deployment points and 84% of girls across all three deployment points. When controlling for clustering at school level and other potential predictors, physical activity level, being South Asian, being in the intervention group and prior compliance were positively associated with monitor wear. **Conclusion:** Compliance reduced across deployment points, with the reduction increasing as the deployment points got further apart. High prior compliance and high physical activity level were associated with the most additional wear-time.

Key words: physical activity; GENEActiv; Girls Active; longitudinal; intervention; predictors of compliance; 24 h wear

Introduction

Accurate and objective measurement of physical behaviours (physical activity, sedentary time and sleep) is important for population surveillance, intervention studies aiming to change behaviours and cohort studies. Furthermore, as sleep, sedentary behaviour and physical activity are linked to the vast majority of chronic diseases and other health outcomes (1), they are potential confounders in any research project analysing associations between exposures and health, thus require accurate measurement. Increasingly, wearable accelerometers are used to obtain an objective measure of people's physical activity. Until recently, studies predominantly used hip-worn monitors during waking hours only (e.g. 2, 3). The emergence of high-resolution raw acceleration wrist-worn accelerometers, which are acceptable for continuous wear (4), has facilitated measurement of sleep and sedentary time as well as physical activity across the entire 24 hour cycle.

Highly sophisticated technology and data-analysis techniques become irrelevant if the wearable monitor does not provide valid measures of physical behaviours or is not worn (5). Both wrist-worn and hip-worn accelerometers can provide valid measures of physical activity, sedentary behavior (6, 7) and sleep (8, 9). However, with hip-worn monitors compliance is often low resulting in missing data (10, 11), which can lead to measurement error and to selection bias due to exclusion of participants with insufficient data (12). Poor compliance could be related to the wear-site, the look and comfort of the device, and/or the necessity for frequent removal of the monitor, e.g. for sleeping, water-based activities, or changing clothing, and therefore the requirement to remember to replace the monitor (5). Greater compliance has been attained with wrist-worn accelerometers; e.g. in the U.S. National Health and Nutrition Examination Survey (NHANES) 2003-2006, 40-70% of participants wore a hip-worn ActiGraph for 10 h/day (hours per

day) of wear for 6+ days, whereas when NHANES 2011-2014 switched to a wrist-worn ActiGraph 70-80% of participants wore the monitors for 21-22 h/day for 6+ days (13). Similarly, in UK Biobank >80% of over 100 000 participants wore the wrist-worn Axivity for at least 150 hours, equivalent to more than six 24 hour days (14).

It is possible that the requirement for 24/7 wear as opposed to a waking wear only protocol contributes to high compliance, as well as the wear-site. The International Study of Childhood Obesity Lifestyle and the Environment (ISCOLE) used a 24 h wear protocol with the hip-worn ActiGraph and consistently captured more than 22 h of wear time over a 24 h day in children, aged 9-11 years, across all 12 ISCOLE countries (15). However, when children (4) or adolescents (11) were requested to wear wrist and hip-worn monitors concurrently, compliance to a wrist-worn monitor was greater than for a hip-worn monitor for both waking day (4) and 24 h per day protocols (11).

Evidence for factors that are associated with compliance largely come from cross-sectional studies with hip-worn monitors and are inconclusive. For example, studies have reported that: overweight and obese children are both less compliant (16) and more compliant (17); age is not associated with compliance (17) and lower in older than younger children (18, 19, 20); monitor wear-time is higher in South Asians and black African-Caribbean children than White Europeans (21) and non-white children are less likely to return reliable accelerometer data than white children (16). The definition used for non-wear time appears to impact on apparent associations with compliance. This may be due to some groups, e.g. older and overweight/obese children, accumulating longer prolonged bouts of sedentary time that are more likely to be misclassified as non-wear, depending on the non-wear definition applied (12). Compliance is also reported to be worse in children who self-report doing less leisure time physical activity (16) or exercise (16), and in economically disadvantaged groups (16, 22).

A consistent finding is that compliance to hip-worn accelerometers tends to drop off with a requirement for repeated deployments (e.g. in longitudinal studies) (10, 20, 23, 24). Achieving high compliance over repeated deployments is crucial when evaluating interventions or for any study evaluating changes in patterns of physical activity and/or their associations with health outcomes. Notably, following concurrent 24/7 wear of wrist and hip monitors, adolescents have reported that they would be more willing to wear a monitor on the wrist than the hip on a further occasion (11). A recent study showed children's cumulative compliance to 24 h wear of wrist-worn accelerometers at two time-points was high at over 80% (25). To our knowledge there are no studies that have assessed compliance to 24 h wear of wrist-worn accelerometers over more than two deployments, or the factors associated with cumulative compliance to 24 h wear of wrist-worn accelerometers. It is important to determine the degree to which compliance drops off over repeated deployments and which factors, modifiable or not, are associated with compliance, to aid with study planning.

Recently, we used wrist-worn accelerometers to evaluate the effectiveness of the 'Girls Active' school-based physical activity programme for changing behaviours, primarily moderate-to-vigorous physical activity (MVPA) (26). The protocol requested girls, aged 11-14 y at baseline, to wear a sports watch style wrist-worn accelerometer 24 h/ day for seven days at three deployment points: baseline, 7-months and 14-months.

The main aim of this analysis was to determine the cross-sectional and cumulative compliance to accelerometer wear at three deployment points in adolescent girls participating in the evaluation of Girls Active (26). We also aimed to determine: a) the number of days of accelerometer wear required to obtain a measure of physical activity that could be considered equivalent to a 7-day measure; and b) whether group allocation (intervention/control), SES (socio-economic status),

ethnicity, BMI (body mass index) z-score, biological maturity status, physical activity level, day of accelerometer deployment or prior compliance were associated with compliance.

Methods

Data for this study were obtained from the evaluation of the Youth Sports Trust's Girls Active school-based physical activity programme (26). The evaluation included 20 schools in and on the boundary of Leicestershire and Rutland, with 10 schools randomised to receive the Girls Active intervention and 10 schools randomised to continue with usual practice. Approximately 90 girls, aged 11-14 y, were invited to participate from each school. They were provided with an information pack containing a letter of invitation, separate participant and parent/guardian information sheets, and an opt out consent form which they were given two weeks to return. Data were collected at baseline, 7-months and 14-months. Ethical approval for the evaluation was obtained from the University of Leicester's College of Medicine and Biological Sciences Research Ethics representative. This paper focuses on monitor compliance only; the Girls Active evaluation is described in detail elsewhere (26).

In brief, data were collected in measurement sessions run during the school day at baseline, 7-months and 14-months. Following the baseline measurement session, schools were randomised to the intervention or the control condition stratified by school size (pupil median: <850, ≥850) and percentage of black and minority ethnicity (BME) pupils (median: <20%, ≥20%). Following each measurement session girls were requested to wear a GENEActiv accelerometer (ActivInsights Ltd., Cambridgeshire, UK) on their non-dominant wrist for 24 hours/day for 7-days and told they would be given a £5 voucher if they wore the device for enough time during the week (duration not specified).

Height, sitting height and mass were measured at each measurement session using standardised procedures. Body mass index was calculated and expressed in z-scores according to reference curves for the UK (27). Age (in months) was calculated from date of birth to date of measurement, ethnicity was self-reported and later collapsed into categories of White European (WE), South Asian (SA) or other, and socioeconomic status (SES) was estimated using the index of multiple deprivation (IMD) from self-reported postcode. The number of years a girl was from her age at peak height velocity (APHV), an indicator of biological maturity, was predicted from baseline measures of height, sitting height, leg length and age. Girls with an APHV within one standard deviation of the mean were classed as 'average maturing' or 'on time'; those with an APHV greater than one standard deviation below the mean as 'early maturers'; and those with an APHV greater than one standard deviation above the mean as 'late maturers' (28).

Accelerometer processing

The GENEActivs were initialised to collect data at 100 Hz and uploaded using GENEActiv PC software version 2.9. The GENEActiv .bin files were analysed with R-package GGIR version 1.2-2 (http://cran.r-project.org) (29, 30). Signal processing in GGIR includes autocalibration using local gravity as a reference (30); detection of sustained abnormally high values; detection of non-wear; and calculation of the average magnitude of dynamic acceleration corrected for gravity (Euclidean Norm minus 1 *g*, ENMO) averaged over 5 s epochs and expressed in milli-gravitational units (m*g*).

Files were excluded from all analyses if post-calibration error was greater than 0.02 *g* (31) and individual days were classified as invalid and excluded if wear-time was less than 16 h (32). Detection of non-wear has been described in detail previously (See 'Procedure for non-wear

detection' in supplementary document to van Hees et al. (29)). Briefly, non-wear is estimated based on the standard deviation and value range of each axis, calculated for 60 min windows with 15-min sliding window. The window is classified as non-wear if, for at least 2 out of the 3 axes the SD (standard deviation) is less than 13 m*g* or the value range is less than 50 m*g*.

Physical Activity was expressed as average acceleration (m*g*) and time accumulated in moderate-to-vigorous physical activity per day (MVPA), defined as time accumulated above an acceleration of 200 m*g* (33).

Analyses

Descriptive statistics were calculated for relevant demographic and physical variables using mean (standard deviation) for continuous variables and percentage for categorical variables. Ethnicity, SES, APHV and maturity status were calculated from baseline data. All other variables were calculated at each time-point.

Compliance

Compliance was calculated for a minimum of 1-, 2-, 3-, 4-, 5-, 6-, and 7-valid days, and also for a minimum of 4-valid days where at least one day was a weekend day as has been recommended (34). Cross-sectional compliance was calculated as the number of girls achieving a given minimum number of valid days ($\geq$16 h) at each time-point (32). Cumulative compliance was calculated as the number of girls achieving a given minimum number of valid days at repeated deployment points: baseline and 7-months; baseline and 14-months; baseline, 7- and 14-months. Data loss due to zero

valid days or technical problems (file length <2 days, failed calibration) was also calculated cross-sectionally for each time-point and longitudinally for each set of deployment points.

Days of measurement required

We used pairwise 95% equivalence tests to determine the number of days of measurement required to obtain physical activity outcomes that could be considered equivalent to a 7-day measure. If the 95% CI (confidence interval) for the mean physical activity assessed from *n* days fell within a ±5% equivalence zone of the 7-day measure (35) the measure was deemed equivalent. This analysis was carried out for 1-, 2-, 3-, 4-, 5- and 6 days at each of the three deployment points for average acceleration and for MVPA and, by definition, included only participants with 7-valid days.

Associations with compliance

Potential predictors of compliance considered were: age (continuous in years), SES (continuous), BMI z-score (continuous), physical activity level (continuous in m*g*), randomisation (control/intervention), maturity status (early maturers, average maturers, late maturers) ethnicity (WE, SA, other), first 24 h day of accelerometer measurement (Fri/Sat/Mon/Tues/Wed/Thurs; the first full day of measurement was never a Sunday) and prior compliance (continuous, number of valid days at baseline and/or 7-months). When considering predictors of compliance, all accelerometer files that were less than 6-days in duration were removed. This was to ensure that all included participants had the potential for maximum compliance, i.e. there was no fault of the monitor or artificially short measurement period (Baseline N = 34, 7-months N = 38, 14-months N = 25). These files were not removed from the estimates of cross-sectional and longitudinal

compliance to ensure that these estimates reflected a 'real world' compliance result from accelerometer deployment. Generalised estimating equations (GEE) were used to determine predictors of compliance controlling for clustering at the school level (Model 1) and, in addition, controlling for all other predictors except prior compliance (Model 2). Finally, Model 3 extended Model 2 to assess the impact of prior compliance on compliance at 7-months and 14-months. Continuous variables were centred prior to entry into the GEE.

Results

At baseline, 7-months and 14-months 1734, 1381 and 1326 girls, respectively, were requested to wear the GENEActiv. The decreasing sample size reflects dropout/loss to follow up from the evaluation. From the variables calculated at baseline (data not in Table), the sample was predominantly WE (76.7%), with 11.6% SA and 11.6% other ethnicity. Approximately two thirds of the sample were normal weight, 20% overweight, 10% obese and 4% underweight at each time-point. Based on APHV, approximately 15% of the sample were classified as early maturing, 15% as late maturing and 70% of the sample 'on time'. A summary of the other descriptive characteristics at baseline and follow-up is shown in Table 1.

Cross-sectional compliance

Figure 1a shows the percentage of girls by the number of valid days wear at each of the three deployment points. The total height of each column represents the total participating at that time-point (100%) and the height of the black column represents the percentage of girls with the specified number of minimum days. Exact percentages are shown in Supplementary Digital Content (SDC) 1. Cross-sectionally, around 80% of girls provided 7-days of valid wear at baseline

and at 7-months; this dropped to just under 70% of girls providing 7-days of valid wear at 14-months. Almost all girls (≥96%) provided at least one valid day of wear, irrespective of time-point.

Over 90% of the baseline, 7-month and 14-month samples provided a minimum of 6-, 4- and 3-valid days of data, respectively. At baseline, 7-months and 14-months 94%, 90% and 84%, respectively, provided at least 3-valid weekdays and 1-valid weekend day (SDC1). The mean (SD) number of hours of wear per valid day was consistent across deployment points at ~23.5 ± ~1.5 h (Table 1).

Data loss due to zero valid days or technical problems was very low at 0.6% and 0.9%, respectively, at baseline, 1.1% and 1.2%, respectively, at 7-months and 1.7% and 2.2%, respectively, at 14-months. Two girls (0.2%) declined to wear the monitor at 14-months, but there was no refusal to wear at baseline or 7-months.

Cumulative compliance

Figure 1b shows the percentage of girls by the number of valid days wear for repeated deployment points: baseline and 7-months (n = 1381); baseline and 14-months (n = 1326); baseline, 7-months and 14-months (n = 1227). Again, the total height of each column represents the total participating at that set of deployment points (100%) and the height of the black column represents the percentage of girls with the specified number of minimum days. The left set of columns show the percentage of girls who had a minimum of one day's wear at both baseline and 7-months through to those who had 7-days wear at both baseline and 7-months. The middle set of columns show the percentage of girls who had a minimum of 1-day's wear at both baseline and 14-months through to those who had 7-days wear at both baseline and 14-months. The right set

of columns show the percentage of girls who had a minimum of 1-day's wear at all three deployment points through to those who had 7-days wear at all three deployment points. Exact percentages are shown in SDC 1.

All percentages are expressed as a percentage of girls participating in the evaluation at all of the deployment points concerned. Note, the cumulative sample size for all three deployment points was lower than for baseline and 14-months due to missing data at 7-months, e.g. one complete school dropped out at 7-months, but returned at 14-months.

68%, 59% and 52% of girls provided a minimum of 7-valid days for baseline and 7-month deployment points, baseline and 14-month deployment points, and all three deployment points, respectively. This increased to 81%, 74% and 67%, respectively, when considering a minimum of 6-valid days. Applying wear criteria of 3-valid weekdays and 1-valid weekend day lead to inclusion of 87%, 81% and 76% of girls for both baseline and 7-months, both baseline and 14-months, and all three deployment points, respectively (SDC1). The vast majority of girls (91%) provided at least one valid day of wear for all three deployment points.

Data loss due to zero valid days or technical problems across sets of deployment points was 2.9% and 2.3%, respectively, for baseline and seven months, 3.2% and 2.6%, respectively, for baseline and 14-months and 5.0% and 4.0% for all three deployment points.

Days of measurement required

Irrespective of time-point, physical activity was over-estimated, relative to the 7-day measure, if only 1- or 2- days of measurement were used (Figure 2: average acceleration; Figure 3: MVPA) .

Estimates of physical activity level from 3- to 6- days of measurement could be considered equivalent to the 7-day measure (i.e. the 95% CI for the mean of the 3-, 4-, 5- and 6- day measure fell within ±5% of the mean of the 7-day measure). The magnitude of these differences between the means was <0.6 mg for average acceleration, and <1 min for MVPA.

Associations with compliance

The number of valid days by deployment point, intervention/control group and categorical variables is shown in SDC2. Irrespective of categorical variable the mean number of valid days was ≥6, with the exception of the control group at 14-months (mean (SD) = 5.9 (1.9) valid days).

SDC3 shows the regression models predicting compliance, defined as the total number of hours the monitor was worn across the week, adjusted for clustering at the school level only. Table 2 shows the regression models further adjusted for all common predictors and results in Table 3 further adjust the models predicting hours of monitor wear at 7-months and 14-months for prior compliance (number of valid days of wear at prior deployment points).

Age was negatively associated with hours of monitor wear in the unadjusted model (SDC3). When controlling for all common predictors (Table 2), physical activity level was consistently and positively associated with hours of monitor wear with the association tending to get stronger as the study progressed. An increase of 1 m*g* in average acceleration was associated with an additional 0.2, 0.3 and 0.4 h monitor wear at successive deployment points, respectively. The SD and size of the 95% CI for average acceleration is approximately 8 and 30 m*g*, respectively. Therefore, the difference in compliance associated with activity levels at the lower and upper

limits of the CI would be approximately 6 h (baseline) to 12 h (14 months), or up to half a day of monitor wear over the week.

At baseline and 14-months, but not 7-months, being in the intervention group was associated with 4 h (baseline) to 7 h (14 months) more monitor wear. The hours of monitor wear were higher (2-4 h) in SA than WE, increasing from 2 h longer at baseline to 4 h longer at 14-months. Starting the first full day of accelerometer wear on a Friday (i.e. deployment on Thursday) led to the highest hours of monitor wear, approximately 6 h longer than a Saturday (i.e. deployment on Friday) and 3 h longer than mid-week at baseline, 6 h longer than a Monday at 7-months and 3 h longer than a Wednesday at 14-months.

Prior compliance predicted hours of monitor wear (Table 3), with the strongest effects for the deployment immediately preceding the one of interest. For example, every valid day of wear at baseline was associated with an extra 11.4 h wear at 7-months and 6.6 h extra wear at 14-months, and every valid day of wear at 7-months was associated with an extra 9.7 h of wear at 14-months.

After controlling for prior compliance, associations of hours of monitor wear with physical activity level, ethnic group and accelerometer measurement start day were largely similar to the model adjusting for all common predictors (Table 2).

Discussion

Compliance to wearing a sports watch style wrist-worn accelerometer during wake and sleep was high, with six valid days of wear being obtained from over 90% of girls at baseline. This reduced with repeated wear occasions, but remained high with 78% of girls providing six valid days of wear

on the third deployment, over a year later. Importantly, a minimum of three days of wear provided measures of overall physical activity and MVPA that were equivalent to 7-day measures. Cross-sectionally, this was obtained from >95% of girls at baseline and at 7-months and from 91% of girls at 14-months. Cumulatively 84% of girls across all three deployment points provided three valid days of wear across all three deployment points, consistent with the compliance reported in children across two deployment points (25). Examining just under 30 000 adult participant with perfect 7-day wear time compliance from the UK Biobank cohort, Doherty et al. (14) also reported that 72 h (three days) of accelerometer wear were needed for a reliable measure, defined as within 10% of a complete 7-day measure (determined using intra-class correlation coefficients).

Employing a waking wear protocol, Howie et al. (36) observed a drop in 10-12 y old children's compliance to a wrist-worn accelerometer from 89% at baseline to 83%, 82% and 79% at the end of three subsequent 8-week conditions (n = 73), when compliance was defined as >10 h wear per day on three weekdays and one weekend day. The decrease in compliance from waking wear protocols appears to come from shorter daily wear durations on the repeated occasions, as compliance was relatively unchanged over the four deployments (96% to 92%) when a less stringent compliance definition of eight hours wear per day on three days was used. Similarly, in a cross-sectional study employing a waking wear protocol, >95% of 9-10 y old children (n = 129) achieved one valid day of wrist accelerometer wear during a single deployment, whether a valid day was defined as 6 or 12 hours (4). However, this diverged considerably when more valid days were required, with less than half the children who met the 6 h wear per day criteria on six days providing 12 h wear per day for six days (64% and 30%, respectively) (4). During a 24 h/day wear protocol, if a monitor is worn on a given day, the daily hours of wear tend to far exceed minimum wear definitions (13, 14, 15, 25). For example, in the current study the average daily hours of wear were >23 h on valid days irrespective of time-point, with around 75% of valid days at each time-

point containing 24 h of wear. Consistent with this, using a 24 h/day wrist accelerometer protocol Price et al. (25) showed that the percentage of 9-10 y old children (N=851) complying to minimum wear criteria of 8 and 16 h over 1-6 days differed by less than 1.8 percentage points. This suggests that classification of valid days in 24 h protocols is less dependent on variations in minimal wear duration, with invalid days resulting mainly from removal of the monitor for the majority of a day and/or participants stopping wear before the measurement period has finished.

Irrespective of deployment point, higher compliance was associated with being more active, being SA, and the first day of full monitor wear not being on a weekend or a Monday. Each of these was associated with up to half a day of extra wear. At 14-months, being in the intervention group was also associated with half a day of extra wear. At baseline, in the model adjusted for clustering at the school level and all other predictors, being in the intervention group was associated with 4 h of extra wear. However, this was not the case in the model adjusted for clustering only and it should be noted that, at the baseline measure, the participants were unaware of which group their school had been randomised to. Considering all deployment points, there was a tendency for Friday to be the optimal first full measurement day (i.e. accelerometer deployment on Thursday). Saturday was the worst first full measurement day at baseline, i.e. accelerometer deployment on Friday (note 7-month and 14-month deployment points did not use a Saturday as a first full measurement day); this suggests that it may be beneficial if the first and last full days of accelerometer wear are in school, where other children may also be wearing an accelerometer and/or teachers are aware of the study. Overall, prior compliance was the strongest predictor of subsequent compliance, with every extra valid day at a given time-point being associated with an extra 10-12 h of compliance for the next deployment point. Studies with multiple deployment points may benefit from identifying participants with poor compliance at the first deployment point and trying to establish any reasons for poor compliance. This may facilitate the

implementation of targeted strategies aimed at improving the compliance of these participants at the next deployment point.

Evidence in the literature for factors associated with compliance is mixed (16). Further, studies have focussed on compliance to wearing accelerometers at the hip for a waking protocol (e.g. 12, 16). As compliance to protocols incorporating wrist-wear and 24/7 wear differs (11, 13, 15), the factors associated with compliance also likely differ. Further, many studies took place before the recent explosion in consumer wearables (37) which have likely 'normalised' the wearing of activity monitors. This limits the extent to which our findings are comparable with those in the literature.

The size of associations with compliance were larger at 14-months than baseline. It is conceivable this was related to the increasing age of the girls, especially as age was negatively associated with compliance at each time-point in the unadjusted model and compliance has been shown to be lower in older children previously (18, 19, 20). However, as the effect of age largely disappeared after controlling for other associates this was more likely an effect of the repeated deployment rather than age.

Although factors associated with compliance were identified, it is notable that overall compliance was generally high. The use of a 24/7 protocol with wrist-worn accelerometers may reduce the impact of demographics and other variables on the quality of the physical activity data obtained. This is supported by the consistently high compliance reported to 24/7 wrist accelerometer wear in a range of different populations and different monitor deployment methods. For example, in a face-to-face physical activity intervention targeting rural adult Australians in South Australia (n= 171), 6+ valid days of GENEActiv wear were provided by 93% of participants at baseline (38), similar to baseline compliance in the current study with adolescent girls. In UK Biobank

accelerometers were mailed out and 80.6% of >100 000 participants, aged 40-69 y, wore the wrist-worn Axivity 24/7 for 150 h (6.25 days) (14), only approximately 10% lower than the studies using a face-to-face protocol. The percentage providing three days of data was >93%, very similar to the current study. Irrespective, poor compliance did lead to some data loss. When designing studies and calculating sample size, researchers should take into account data loss associated with compliance to the accelerometer protocol they are deploying as well as considering dropout from the study.

This study has a number of limitations that should be highlighted. We do not know the degree to which the small financial incentive may have contributed to compliance in this study. This is a limitation, as it has previously been reported that incentives for returning accelerometers and for providing sufficient data leads to greater compliance to waking protocols with hip-worn accelerometers in adolescents (39, 40). Secondly, as stated in the introduction, the definition used for non-wear time impacts on estimates of compliance and on apparent associations with compliance. Therefore, it may be necessary to replicate this type of study as methods for wear-time detection advance. Thirdly, these findings are specific to our population of adolescent girls and may not be generalizable to other age-groups, boys and/or different monitor deployment protocols. Finally, we did not consider whether compliance was affected by weather conditions; future studies should investigate whether variation in season, e.g. weather, seasonal attire and seasonal sports, impacts on compliance.

In conclusion, the high cross-sectional and cumulative compliance to 24/7 wrist accelerometer wear in the evaluation is very encouraging, particularly as adolescent girls can be a challenging population to assess with accelerometry. They have expressed more concerns about the appearance of (hip-worn) accelerometers than boys (40), although have also indicated that they

found the sports watch style GENEActiv wrist-worn monitor less embarrassing (11). Three-quarters of the sample provided six valid days at baseline and 14-months, or five valid days at all three deployment points. A valid estimate of physical activity (equivalent to a 7-day measure) was obtained with three days; achieved by >90% of the sample for any single time-point, and ≥84% of girls for multiple deployment points. High compliance reduces measurement error as a more reliable and complete measure of physical behaviours can be captured and reduces bias as fewer participants are excluded due to failing to meet minimum wear criteria.

References

1. Carson V, Tremblay MS, Chaput J-P, Chastin S. Associations between sleep duration, sedentary time, physical activity, and health indicators among Canadian children and youth using compositional analyses. *Appl Physiol Nutr Metab.* 2016; 41: S294–S302. doi: 10.1139/apnm-2016-0026.

2. Sherar LB, Griew P, Esliger DW et al. International children's accelerometry database (ICAD): design and methods*. BMC Public Health.* 2011; 11: 485. doi: 10.1186/1471-2458-11-485.

3. Troiano R, Berrigan D, Dodd KW et al., Physical activity in the united states measured by accelerometer. *Med Sci Sports Exerc.* 2008; 40(1): 181-188.

4. Fairclough SJ, Noonan R, Rowlands AV, van Hees V, Knowles Z, Boddy LM. Wear compliance and activity in children wearing wrist and hip mounted accelerometers. *Med Sci Sports Exerc.* 2016; 48, 243-253.

5. Rowlands AV, Accelerometer Assessment of Physical Activity in Children: An Update. *Pediatr Exerc Sci.* 2007; 19: 252-266.

6. **E**sliger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA accelerometer. *Med Sci Sports Exerc.* 2011; 43: 1085-1093.

7. Phillips LRS, Parfitt CG, Rowlands AV. Calibration of the GENEA accelerometer for assessment of physical activity intensity in children. *J Sci Med Sport*, 2013; 16: 124-128.

8. van Hees VT, Sabia S, Anderson KN et al. A Novel, Open Access Method to Assess Sleep Duration Using a Wrist-Worn Accelerometer. *PLoS ONE.* 2015; 10 (11): e0142533. doi:10.1371/journal.pone.0142533.

9 Barreira TV, Schuna Jr JM, Mire EF, Katzmarzyk PT, Chaput JP, Leduc G, Tudor_Locke C. Identifying Children's Nocturnal Sleep Using 24-h Waist Accelerometry. *Med Sci Sports Exerc.* 2015; 47: 937–943.

10. Howie EK, Straker L. Rates of attrition, non-compliance and missingness in randomized controlled trials of child physical activity interventions using accelerometers: A brief methodological review. *J Sci Med Sport.* 2016: 19; 830-836.

11. Scott JJ, Rowlands AV, Cliff D, Morgan PJ, Plotnikoff RC, Lubans DR. Testing the validity and feasibility of the GENEActiv accelerometer in free-living adolescents. *J Sci Med Sport.* 2017. doi.org/10.1016/j.jsams.2017.04.017.

12. Toftager M, Kristensen PL, Oliver M et al. Accelerometer data reduction in adolescents: effects on sample retention and bias. *Int J Behav Nutr Phys Act.* 2013; 10:140: doi: 10.1186/1479-5868-10-140.

13. Freedson PS, John D. Comment on "estimating activity and sedentary behavior from an accelerometer on the hip and wrist". *Med Sci Sports Exerc.* 2013. 45(5): 962-963.

14. Doherty A, Jackson D, Hammerla N et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PLoS ONE.* 2017. doi.org/10.1371/journal.pone.0169649.

15. Tudor-Locke C, Barreira TV, Schuna J et al., Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). *Int J Behav Nutr Phys Act.* 2015; 12: 11. doi: 10.1186/s12966-015-0172-x.

16. Rich C, Cortina-Borja M, Dezateux C, et al. Predictors of non-response in a UK-wide cohort study of children's accelerometer-determined physical activity using postal methods. *BMJ Open.* 2013; 3: e002290. doi: 10.1136/bmjopen-2012-002290.

17. van Coevering P, Harnack L, Schmitz K, Fulton JE, Galuska DA, Gao S. Feasibility of using accelerometers to measure physical activity in young adolescents. *Med Sci Sport Exerc.* 2005; 37: 867-871.

18. Mattocks C, Ness A, Learly S et al. Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision. *J Phys Act Health.* 2008; 5: S98-S111.

19. Sirard JR, Slater ME. Compliance with wearing physical activity accelerometers in high school students. *J Phys Act Health.* 2009; 6(suppl): S148-S155.

20. Sutherland R, Campbell E, Lubans D et al., 'Physical Activity 4 Everyone' school-based intervention to prevent decline in adolescent physical activity levels: 12 month (mid-intervention) report on a cluster randomised trial. *Br J Sports Med.* 2016; 50(8): 488-95.

21. Owen CG, Nightingale CM, Rudnicka AR, Cook DG, Ekelund U, Whincup PH. Ethnic and gender differences in physical activity levels among 9-10-year-old children of White European, South Asian, and African-Caribbean origin: the Child Heart Health Study in England (CHASE) study. *Int J Epidemiol.* 2009; 38: 1082-1093.

22. Cohen KE, Morgan PJ, Plotnikoff RC, et al. Physical activity and skills intervention: SCORES cluster randomized controlled trial. *Med Sci Sports Exerc.* 2015; 47:765–74.

23. Lubans DR, Morgan PJ, Okely AD et al. Preventing obesity among adolescent girls: one-year outcomes of the nutrition and enjoyable activity for teen girls (NEAT Girls) cluster randomized controlled trial. *Arch Pediat Adolesc Med.* 2012; 166: 821–7.

24. Grydeland M, Bergh IH, Bjelland M et al. Intervention effects on physical activity: the HEIA study—a cluster randomized controlled trial. *Int J Behav Nutr Phys Act.* 2013;10:17.

25. Price L, Wyatt K, Lloyd J et al. Children's compliance with wrist worn accelerometry within a cluster randomised controlled trial: Findings from The Healthy Lifestyles Programme (HeLP). *Pediatr Exerc Sci.* In press.

26. Edwardson CL, Harrington DM, Yates T et al. A cluster randomised controlled trial to investigate the effectiveness and cost effectiveness of the 'Girls Active' intervention: a study protocol. *BMC Public Health.* 2015; 15(1):526. doi: 10.1186/s12889-015-1886-z.

27. Cole TJ, Freeman JV, Preece MA. Body mass index reference curves for the UK, 1990. *Arch Dis Childhood.* 1995; 73(1): 25-9.

28. Malina RM, Bouchard C, Bar-Or O. *Growth, Maturation and Physical Activity.* Champaign, IL: Human Kinetics; 2004. pp. 277-302.

29. 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. doi: 10.1371/journal.pone.0061691.

30. van Hees VT, Fang Z, Langford J et al. Auto-calibration 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): 738-744.

31. da Silva ICM, van Hees VT, Ramires VV et al. Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. *Int J Epidemiol.* 2014; 43(6):1959-1968.

32. Rowlands AV, Cliff DP, Fairclough SJ et al. Moving forward with backwards compatibility: Translating wrist accelerometer data. *Med Sci Sports Exerc.* 2016; 48: 2142-2149.

33. Hildebrand M, Van Hees VT, Hansen BH, Ekelund U. Age-Group Comparability of Raw Accelerometer Output from Wrist- and Hip-Worn Monitors. *Med Sci Sport Exerc.* 2014; 46: 1816-1824.

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

35. Wellek S. *Testing Statistical Hypotheses of Equivalence.* Boca Raton (FL): Chapman & Hall/CRC; 2003, p. 284.

36. Howie EK, McVeigh JA, Straker LM. Comparison of compliance and intervention outcomes between hip- and Wrist-Worn accelerometers during a randomized crossover trial of an active video games intervention in children. *J Phys Act Health.* 2016; 13: 964-969.

37. Geib RW, Swink PJ, Vorel AJ, Shepard CS, Gurovich AN, Waite GN. (2015). The bioengineering of changing lifestyle and wearable technology: A mini review. *Biomed Sci Instrum.* 2015; 51: 69-76.

38. Mitchell BL, Lewis NR, Smith AE, Rowlands AV, Parfitt G, Dollman J. Rural Environments and Community Health (REACH): a randomised controlled trial protocol for an online walking intervention in rural adults. *BMC Public Health*. 2014; 14: 969. doi: 10.1186/1471-2458-14-969.

39. Sirard JR, Kubik MY, Fulkerson JA, Arcan C. Objectively measured physical activity in urban alternative high school students. *Med Sci Sports Exerc.* 2008; 40: 2088-2095.

40. Audrey S, Bell S, Hughes R, Campbell R. Adolescent perspectives on wearing accelerometers to measure physical activity in population-based trials. *Eur Public Health.* 2013; 23: 475-480.

Acknowledgements

We thank all the pupils and teachers who took part in the Girls Active evaluation study and the Girls Active project staff: Mrs Kyla Harrington (project manager), Ms Tatiana Plekhanova (Research Assistant), and Mrs Harshada Chauhan (research administrator). We thank the individual members of the measurement teams that went into schools (comprising of staff and intern students from Leicester Diabetes Centre and the NIHR BRU and students from Loughborough University) and the Leicester CTU staff (data manager, database builder and the statistical team, in particular Laura Gray). We thank the members of the Girls Active Trial Steering Committee for their advice and support over the course of the study: Angie Page (University of Bristol; chair), Esther van Sluijs (University of Cambridge), Siobhan Creanor (Plymouth University) and our public members Stephanie Dunkley (Leicester City Council) and Wayne Allsop (New College, Leicester). Finally, we thank the Youth Sport Trust (YST): Chris Wright, Dominique Lammie, Paulina Skladanowska and Ali Goodall as well as YST tutors and hub school staff.

Conflicts of interest and sources of funding

The Girls Active evaluation was funded by the NIHR Public Health Research Programme (13/90/30). Professors Davies and Khunti are NIHR Senior Investigators. University of Leicester authors are supported by the NIHR Leicester-Loughborough Biomedical Research Unit 92012-2017), the NIHR Leicester Biomedical Research Centre (2017-2022) and the Collaboration for leadership in Applied Health Research and Care (CLAHRC) East Midlands. The Girls Active evaluation was undertaken in collaboration with the Leicester Clinical Trials Unit, a UKCRC-

registered clinical trials unit in receipt of NIHR CTU support funding. The Youth Sport trust or the aforementioned funders had no involvement in the data analysis, data interpretation, data collection, or writing of this manuscript. There are no other conflicts of interest. The results of the present study do not constitute endorsement by the ACSM. The results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation.

List of Figures

Figure 1. a) Cross-sectional compliance and data loss at baseline, 7-months and 14-months; b) Cumulative compliance and data loss at repeated time-points: Baseline and 7 months; Baseline and 14 months; Baseline, 7 months and 14 months

Figure 2. Equivalence between average acceleration measured for 7-days and for 1-6 days at a) baseline, b) 7-months and c) 14-months

Figure 3. Equivalence between MVPA measured for 7-days and for 1-6 days at a) baseline, b) 7-months and c) 14-months

List of Supplemental Digital Content

SUPPLEMENTARY DIGITAL CONTENT 1. Percentage of girls meeting minimum wear criteria at each deployment point (cross-sectional) and at sets of deployments (cumulative).

SUPPLEMENTARY DIGITAL CONTENT 2. Number of valid days (mean(SD)) at each deployment point by demographic and physical categories.

SUPPLEMENTARY DIGITAL CONTENT 3. Generalised estimating equation for prediction of total number of hours monitor worn at baseline, 7-months and 14-months, adjusted for clustering at school level only (Model 1).

Table 1. Descriptive characteristics of the Girls Active participants included in these analyses

Variable	Mean (SD) value at each deployment point except where stated otherwise		
	Baseline (n=1734)	7-months (n = 1381)	14-months (n=1326)
Age (y)	12.8 (0.8)	13.5 (0.8)	14.0 (0.8)
Years from Age at Peak Height Velocity (APHV)	0.7 (0.8)	1.4 (0.8)	1.8 (0.8)
Height (cm)	155.9 (7.9)	158.9 (7.2)	160.2 (6.9)
Mass (kg)	48.8 (12.4)	52.1 (12.5)	54.1 (12.5)
Body mass index ($kg \cdot m^{-2}$)	19.9 (4.0)	20.5 (4.1)	21.0 (4.2)
Body mass index z-score	0.18 (1.33)	0.25 (1.32)	0.33 (1.27)
Average daily acceleration (m*g*)	36.2 (8.9)	33.9 (7.8)	34.5 (8.5)
MVPA (min/day)	45.5 (20.8)	41.0 (18.4)	42.7 (20.3)
Compliance: Valid days*	6.7 (1.1)	6.4 (1.4)	6.1 (1.7)
Compliance: Hours per valid day*	23.6 (1.2)	23.5 (1.5)	23.3 (1.7)
Median (25^{th} to 75^{th} %ile)	24.0 (24.0, 24.0)	24.0 (23.8, 24.0)	23.0 (24.0, 24.0)
Compliance: Hours of monitor wear†	161.4 (17.3)	157.9 (21.9)	153.0 (26.7)
Compliance: Hours per day	22.9 (3.7)	22.1 (5.0)	21.4 (5.7)

* Valid day: 16 h minimum wear

† Maximum hours = 167.3 h (7 days: GENEActiv max recording time at 100 Hz)

SD = standard deviation; MVPA = moderate-to-vigorous physical activity

Table 2. Generalised estimating equation for prediction of total number of hours monitor worn at baseline, 7-months and 14-months, adjusted for clustering at school level and all other predictors except prior compliance (Model 2).

		Baseline (listwise N= 1529)		7 months (listwise N = 1214)		14 months (listwise N = 1157)	
Continuous variables		Change in compliance (h)	95% CI	Change in compliance (h)	95% CI	Change in compliance (h)	95% CI
Age		-0.23	-1.37, 0.91	**-1.38**	**-2.50, -0.25**	-1.18	-3.02, 0.66
Socio-economic status (SES)		0.00	-0.00, 0.00	0.00	-0.00, 0.00	0.00	-0.00,0.00
BMI z-score		-0.05	-0.59, 0.49	0.56	-0.43, 1.56	0.59	-0.67, 1.84
Physical activity level (average acceleration, m*g*)		**0.21**	**0.06, 0.36**	**0.32**	**0.17, 0.46**	**0.37**	**0.16, 0.58**
Categorical variables							
Randomisation	Control	0	Reference	0	Reference	0	Reference
	Intervention	**3.94**	**1.68, 6.20**	1.63	-0.81, 4.07	**7.22**	**4.17, 10.26**
Biological maturity	Early maturers	0	Reference	0	Reference	0	Reference
	Average maturers	-0.81	-2.76, 1.14	-1.00	-4.93, 2.92	-2.06	-5.65, 1.53
	Late maturers	-1.17	-4.32, 1.98	-0.15	-4.72, 4.41	0.02	-4.27, 4.31
Ethnicity	White European	0	Reference	0	Reference	0	Reference
	South Asian	**1.91**	**0.57, 3.25**	**3.50**	**0.70, 6.30**	**4.01**	**0.72, 7.31**
	Other	1.61	-0.31, 3.54	0.54	-2.40, 3.49	-2.07	-7.68, 3.55
Start day	Friday	0	Reference	0	Reference	0	Reference
	Saturday	**-5.92**	**-10.89, 0.96**	N/A	N/A	N/A	N/A
	Monday	N/A	N/A	**-5.81**	**-7.13, -4.48**	N/A	N/A
	Tuesday	**-2.94**	**-5.25, -0.63**	-0.82	-5.59, 3.95	-3.65	-8.10, 0.80
	Wednesday	**-3.68**	**-6.37, -0.99**	-2.75	-5.72, 0.22	**-2.64**	**-4.65, -0.64**
	Thursday	-1.30	-3.86, 1.27	-1.53	-4.04, 0.97	-2.28	-6.39, 1.83

Significant results denoted in bold ($p < 0.05$)
N.B. Excluded participants whose monitors recorded for <6 days to ensure no faulty devices and/or short measurement periods (Baseline N = 34, 7-months N = 38, 14-months N = 25)

h = hours; CI = confidence interval

Table 3. Generalised estimating equation for prediction of total number of hours monitor worn at 7-months and 14-months, adjusted for clustering at school level and all other predictors including prior compliance (Model 3).

		7 months (listwise N= 1169) **Compliance at baseline**		14 months (listwise N = 1113) **Compliance at baseline**		14 months (listwise N = 974) **Compliance at baseline and 7 months**	
Continuous variables		Change in compliance (h)	95% CI	Change in compliance (h)	95% CI	Change in compliance (h)	95% CI
Age		**-1.67**	**-2.77, -0.58**	-0.74	-2.61, 1.13	-0.40	-2.19, 1.39
Socio-economic status (SES)		-0.00	-0.00, 0.00	0.00	-0.00, 0.00	0.00	-0.00, 0.00
BMI z-score		0.20	-0.84, 1.24	0.53	-0.62, 1.68	0.14	-0.92, 1.19
Physical activity level (average acceleration, m*g*)		**0.27**	**0.11, 0.43**	**0.38**	**0.14, 0.63**	**0.28**	**0.02, 0.55**
Baseline compliance (days)		**11.38**	**6.65, 16.11**	**6.54**	**0.30, 12.78**	2.91	-2.03, 7.86
7-month compliance (days)		N/A	N/A	N/A	N/A	**9.74**	**6.38, 13.11**
Categorical variables							
Randomisation	Control	0	Reference	0	Reference	0	Reference
	Intervention	1.14	-1.11, 3.39	**6.60**	**3.01, 10.18**	**5.42**	**3.79, 7.06**
Biological maturity	Early maturers	0	Reference	0	Reference	0	Reference
	Average maturers	-0.62	-4.63, 3.38	-2.74	-6.28, 0.79	-1.16	-4.12, 1.81
	Late maturers	0.13	-4.91, 5.16	-0.86	-4.92, 3.19	-0.66	-4.34, 3.01
Ethnicity	White European	0	Reference	0	Reference	0	Reference
	South Asian	**3.29**	**0.75, 5.82**	**4.87**	**0.75, 9.00**	**5.66**	**2.01, 9.30**
	Other	-0.63	-3.42, 2.17	-2.32	-7.68, 3.04	-1.36	-6.57, 3.85
Start day	Friday	0	Reference	0	Reference	0	Reference
	Saturday	N/A	N/A	N/A	N/A	N/A	N/A
	Monday	**-7.40**	**-8.70, -6.10**	N/A	N/A	N/A	N/A
	Tuesday	1.87	-0.39, 4.12	-3.58	-8.84, 1.67	-1.37	-4.15, 1.41
	Wednesday	-2.39	-4.81, 0.04	**-2.35**	**-4.40, -0.30**	-1.67	-4.33, 1.00
	Thursday	-0.58	-2.89, 1.73	-2.45	-6.75, 1.86	**-3.58**	**-6.23, -0.93**

Significant results denoted in bold ($p < 0.05$)
N.B. Excluded participants whose monitors recorded for <6 days to ensure no faulty devices and/or short measurement periods (Baseline N = 34, 7-months N = 38, 14-months N = 25)

h = hours; CI = confidence interval

SUPPLEMENTARY DIGITAL CONTENT 1. Percentage of girls meeting minimum wear criteria at each deployment point (cross-sectional) and at sets of deployments (cumulative).

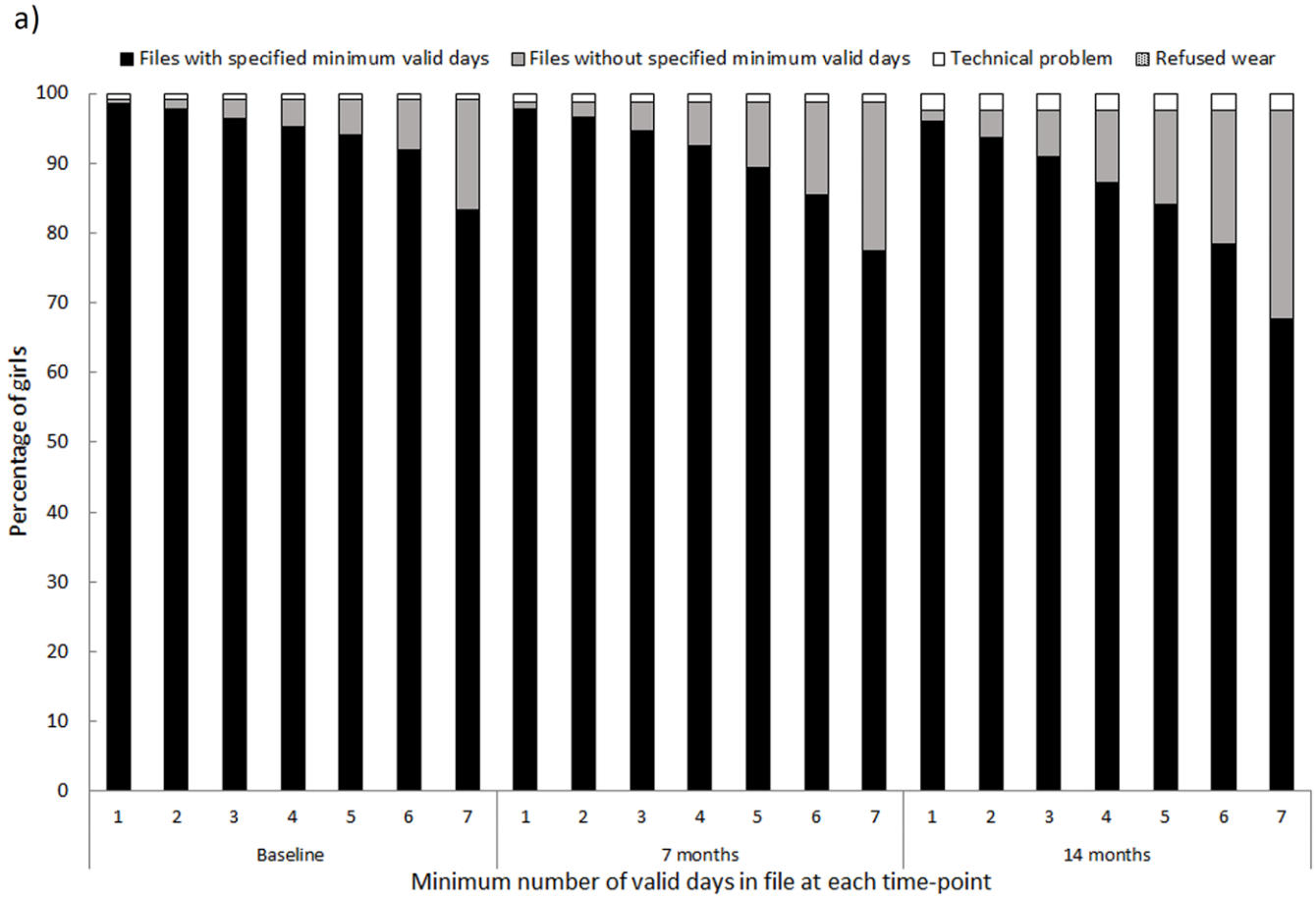
Cross-sectional: % of girls achieving minimum number of days at each deployment point								
Deployment point (months)	1-day	2-days	3-days	4-days	5-days	6-days	7-days	4-days including weekend day
0 (n=1734)	98.5	97.8	96.4	95.2	94.1	92.0	83.3	94.5
7 (n=1381)	97.7	96.6	94.6	92.5	89.4	85.5	77.4	90.3
14 (n=1326)	95.9	93.7	90.9	87.3	84.1	78.4	67.7	84.0
Cumulative: % of girls achieving minimum number of days at each set of deployment points								
0, 7 (n=1381)	95.7	94.4	91.5	89.0	85.7	80.7	67.5	86.8
0, 14 (n=1326)	94.0	91.5	88.1	84.0	80.5	73.8	59.4	80.7
0, 7, 14 (n=1227)	90.9	88.1	84.0	79.1	74.6	67.4	51.6	75.7

SUPPLEMENTARY DIGITAL CONTENT 2. Number of valid days (mean(SD)) at each deployment point by demographic and physical categories.

Variable		Deployment point Baseline		7 months		14 months	
Categorical variables		N(%)	Valid days (mean(SD))	N(%)	Valid days (mean(SD))	N(%)	Valid days (mean(SD))
Experimental group	Control	843(50.2)	6.6 (1.1)	590(44.5)	6.5 (1.4)	583(46.0)	5.9 (1.9)
	Intervention	836(49.8)	6.8 (0.8)	735(55.5)	6.5 (1.3)	684(54.0)	6.4 (1.3)
Biological maturity	Early	265(15.9)	6.8 (0.7)	202(15.3)	6.6 (1.1)	202(16.0)	6.2 (1.7)
	On time	1142(68.3)	6.7 (1.0)	906(68.6)	6.5 (1.4)	870(68.7)	6.2 (1.6)
	Late	264(15.8)	6.7 (1.0)	212(16.1)	6.5 (1.2)	19(15.3)	6.2 (1.6)
Weight status	Normal weight	1182(71.7)	6.7 (0.9)	910(70.0)	6.5 (1.4)	871(69.9)	6.2 (1.7)
	Overweight	322(19.5)	6.7 (0.9)	269(20.7)	6.6 (1.1)	259(20.8)	6.2 (1.6)
	Obese	144(8.7)	6.7 (0.9)	122(9.4)	6.6 (1.1)	116(9.3)	6.3 (1.4)
Ethnicity	White European	1291(77.1)	6.7 (0.9)	985(74.4)	6.5 (1.4)	937(74.1)	6.2 (1.6)
	South Asian	191(11.4)	6.7 (1.2)	164(12.4)	6.6 (1.3)	172(13.6)	6.3 (1.7)
	Other	193(11.5)	6.8 (0.8)	175(13.2)	6.5 (1.2)	155(12.3)	6.0 (1.8)
Start day	Monday	0	N/A	76(5.7)	6.4 (1.3)	0	N/A
	Tuesday	336(20.0)	6.7 (1.0)	149(11.3)	6.5 (1.2)	302(23.8)	6.0 (1.8)
	Wednesday	330(19.7)	6.8 (0.8)	365(27.6)	6.4 (1.6)	290(22.9)	6.2 (1.6)
	Thursday	421(25.1)	6.8 (0.8)	289(21.8)	6.4 (1.4)	335(26.4)	6.2 (1.7)
	Friday	336(20.0)	6.8 (1.0)	446(33.7)	6.6 (1.1)	340(26.8)	6.3 (1.5)
	Saturday	256(15.3)	6.4 (1.1)	0	N/A	0	N/A

N.B. Excluded participants whose monitors recorded for <6 days to ensure no faulty devices and/or short measurement periods (Baseline N = 34, 7-months N = 38, 14-months N = 25)
SD = standard deviation

Figure 1. a) Cross-sectional compliance and data loss at baseline, 7-months and 14-months; b) Cumulative compliance and data loss at repeated time-points: Baseline and 7 months; Baseline and 14 months; Baseline, 7 months and 14 months



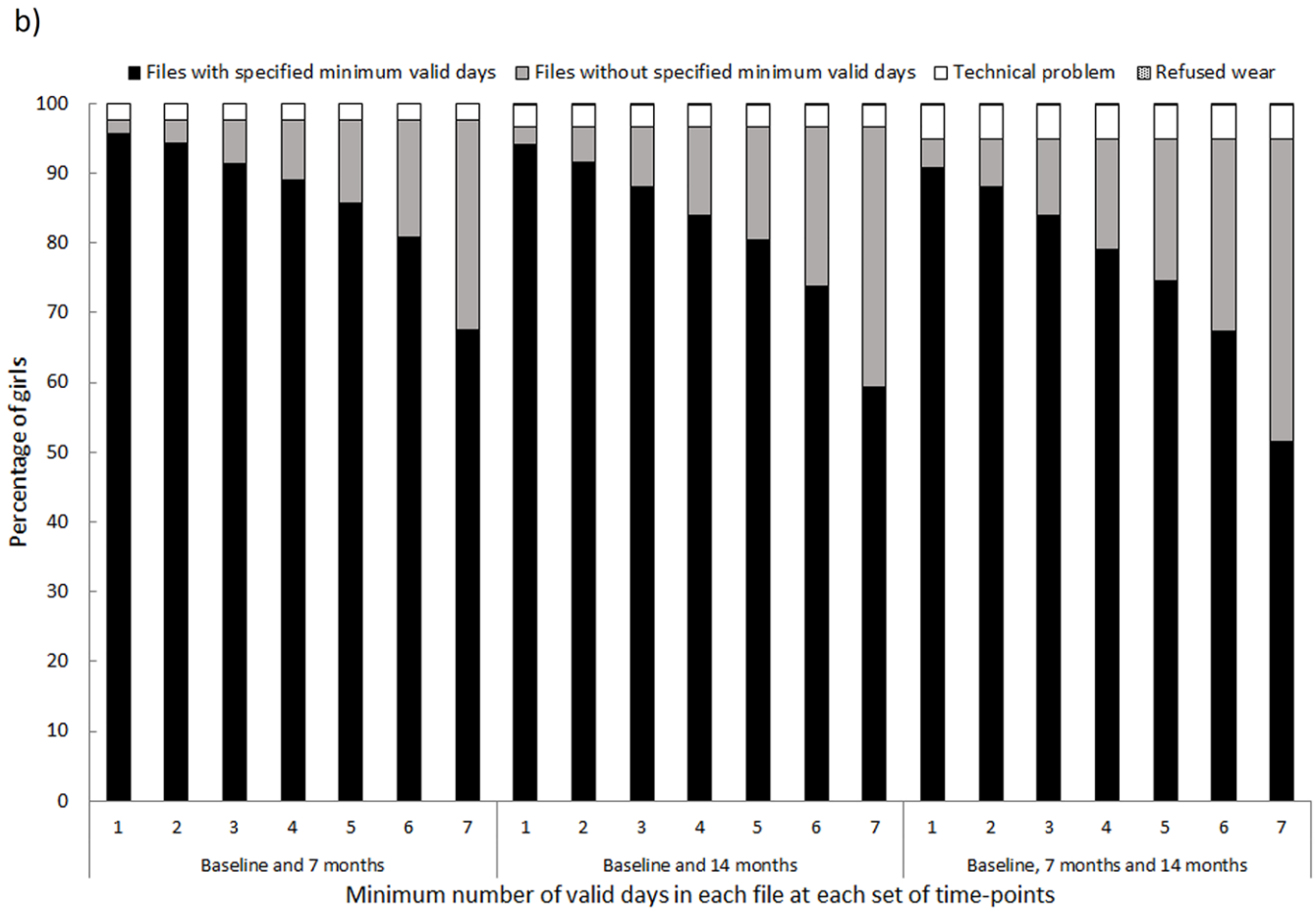
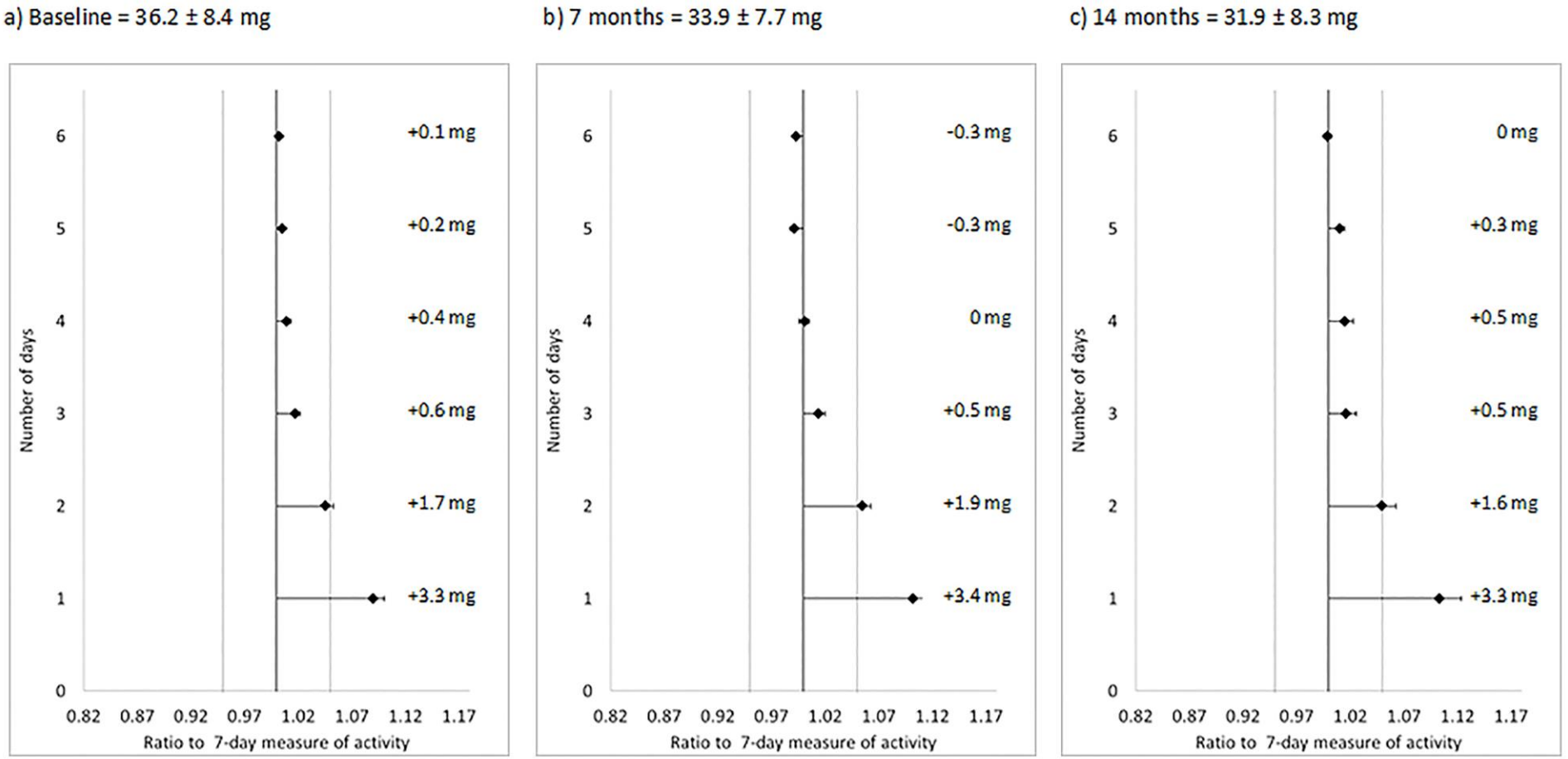
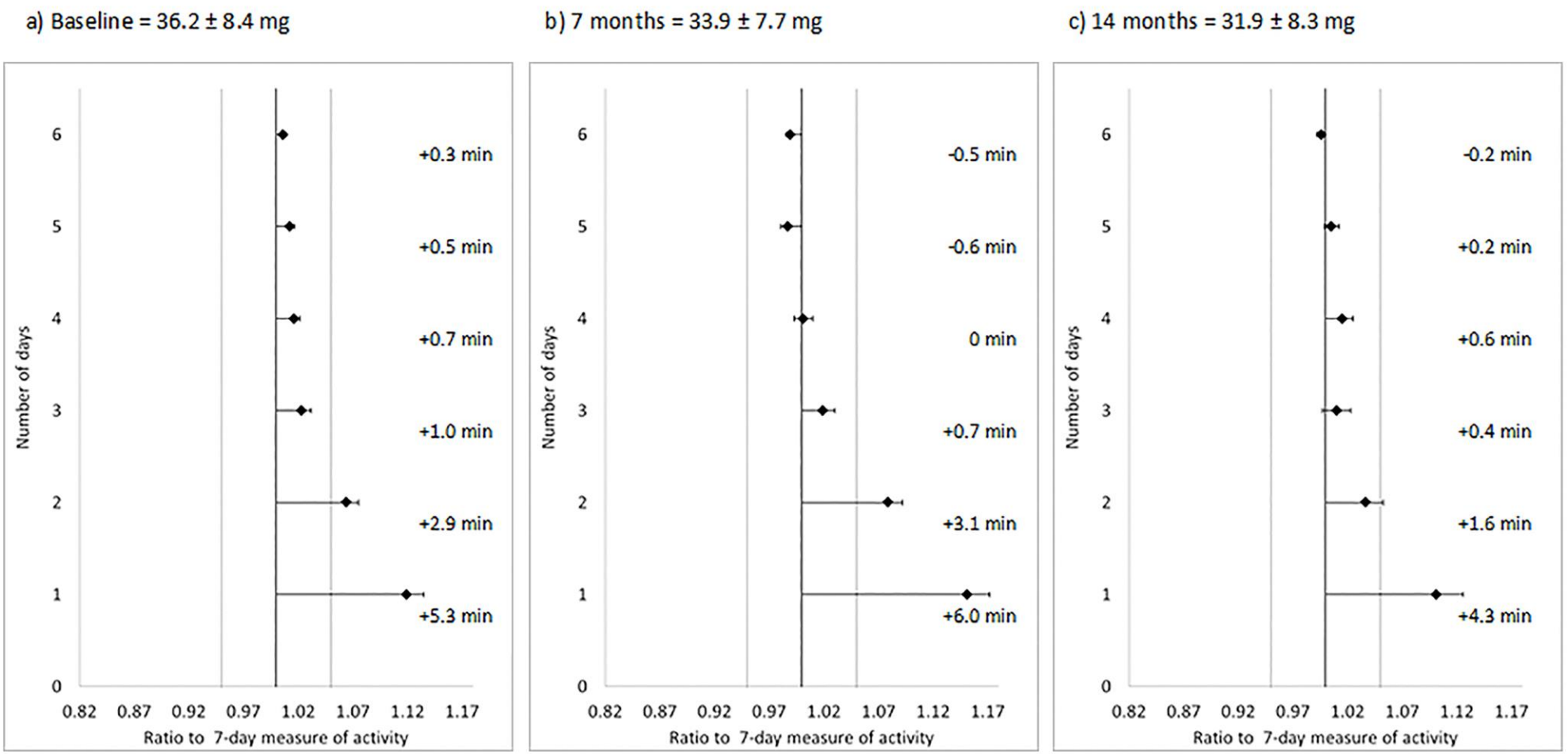


Figure 2. Equivalence between average acceleration measured for 7-days and for 1-6 days at a) baseline, b) 7-months and c) 14-months



Error bars represent the 95% confidence interval of the ratio. Equivalence = 1 (solid line), proposed equivalence zone (i.e. +/-5% of the mean) represented by grey vertical lines. Numbers to the right of each figure represent the mean over- (+) or under- (-) estimation of the 7-day value.

Figure 3. Equivalence between MVPA measured for 7-days and for 1-6 days at a) baseline, b) 7-months and c) 14-months



Error bars represent the 95% confidence interval of the ratio. Equivalence = 1 (solid line), proposed equivalence zone (i.e. +/-5% of the mean) represented by grey vertical lines. Numbers to the right of each figure represent the mean over- (+) or under- (-) estimation of the 7-day value.

View publication stats