1-2008

Use of Accelerometers in a Large Field-Based Study of Children: Protocols, Design Issues, and Effects on Precision

Calum Mattocks

Andy R. Ness

Sam D. Leary

Kate Tilling

Steven N. Blair
University of South Carolina - Columbia, sblair@mailbox.sc.edu

See next page for additional authors

Publication Info

Published in *Journal of Physical Activity and Health*, Volume 5, Issue Supplement 1, 2008, pages S98-S111.

Mattocks, C., Ness, A., Leary, S., Tilling, K., Blair, S. N., Shield, J., ... Riddoch, C. (2008). Use of accelerometers in a large field-based study of children: Protocols, design issues, and effects on precision. *Journal of Physical Activity and Health, 5*(Supplement 1), S98-S111.

Author(s)

Calum Mattocks, Andy R. Ness, Sam D. Leary, Kate Tilling, Steven N. Blair, Julian Shield, Kevin Deere, Joanne Saunders, Joanne Krikby, George Davey Smith, Jonathan Wells, Nicholas Wareham, John J. Reilly, and Chris J. Riddoch

Journal of Physical Activity & Health, 2008, 5, Supp 1, S98-S111

Use of Accelerometers in a Large Field-Based Study of Children: Protocols, Design Issues, and Effects on Precision

Calum Mattocks, Andy Ness, Sam Leary, Kate Tilling, Steven N. Blair, Julian Shield, Kevin Deere, Joanne Saunders, Joanne Kirkby, George Davey Smith, Jonathan Wells, Nicholas Wareham, John Reilly, and Chris Riddoch

Background: Objective methods can improve accuracy of physical activity measurement in field studies but uncertainties remain about their use. ***Methods:*** Children age 11 years from the Avon Longitudinal Study of Parents and Children (ALSPAC), were asked to wear a uni-axial accelerometer (MTI Actigraph) for 7 days. ***Results:*** Of 7159 children who attended for assessment, 5595 (78%) provided valid measures. The reliability coefficient for 3 days of recording was .7 and the power to detect a difference of 0.07 SDs ($P \leq .05$) was > 90%. Measures tended to be higher on the first day of recording (17 counts/min; 95% CI, 10–24) and if children wore the monitor for fewer days, but these differences were small. The children who provided valid measures of activity were different from those who did not, but the differences were modest. ***Conclusion:*** Objective measures of physical activity can be incorporated into large longitudinal studies of children.

Keywords: epidemiology, physical activity, pediatrics, ALSPAC

Background

Physical activity has been defined by Caspersen et al[1] as "any bodily movement produced by skeletal muscles that results in energy expenditure."[(p126)] Evidence accumulated over the last 50 years suggests that regular physical activity in adult life has beneficial effects on health. In particular, regular physical activity is

Mattocks, Tilling, Deere, Saunders, Kirkby, and Smith are with the Dept of Social Medicine, University of Bristol, UK. Ness and Leary are with the Dept of Oral and Dental Science, Bristol Dental School, BS1 2LY, UK. Blair is with the Dept of Exercise Science, University of South Carolina, Columbia, SC 29208. Shield is with the Institute of Child Life and Health, United Bristol Healthcare Trust, Bristol, BS2 8AE, UK. Wells is with the Childhood Nutrition Centre, Institute of Child Health, London, WC1N 1EH, UK. Wareham is with the Medical Research Council Epidemiology Unit, Cambridge, CB2 0QQ, UK. Reilly is with the Division of Developmental Medicine, Yorkhill Hospitals, University of Glasgow, G3 8SJ, UK. Riddoch is with Sport and Exercise Sciences, School for Health, University of Bath, BA2 7AY, UK.

associated with reductions in morbidity from and mortality attributed to coronary heart disease, non–insulin-dependent diabetes and certain cancers.[2] In contrast to the body of evidence on the health benefits in adults, the empirical evidence in children is less extensive.[3,4] This, in part, reflects difficulties in measuring physical activity in children.[5]

Objective methods of assessing physical activity have been developed and offer the opportunity to collect more accurate data in children.[6] Although these methods are increasingly being used in large studies,[7,8] there are still uncertainties about their use. These include the number of hours per day and total number of days of measurement required to characterize usual activity, the likely compliance in large studies (and the resulting potential bias introduced by nonresponse), the potential for instrument reactivity (changes in activity resulting from wearing the instrument), and the potential for bias resulting from differences in number of days of measurement and different start days.

This article describes how we have addressed these uncertainties in the protocols we have adopted for objective measurement of physical activity in a large, prospective study—the Avon Longitudinal Study of Parents and Children (ALSPAC). Some of the protocol decisions were made based on the experience of other groups[7] before the data presented here were analyzed, and some were necessary to fit around the data collection procedures of ALSPAC. For example, the method of dispensing the monitors needed to fit into the ALSPAC clinic schedule and protocols. The longitudinal nature of the study also meant that we were keen to reduce subject attrition in order to maintain statistical power. We recognize that other studies with different designs (eg, intervention studies) and sample sizes might have different requirements, that the methods presented here might not be suitable for some studies, and that the results will not be informative in all circumstances.

Study Design and Methods

Study Population

ALSPAC is a geographically based birth cohort that has been described in detail previously.[9] Briefly, all pregnant women in the former Avon Health Area who had an expected delivery date between April 1, 1991, and December 31, 1992, were asked to take part in the study. A total of 14,541 pregnant women were enrolled, and this resulted in 14,062 live births. Detailed data have been collected by self-completed questionnaires from pregnancy onward. From the age of 7, all children have been invited to regular research clinics.

Ethical Approval and Consent

Ethical approval for the measurement of physical activity was given by the ALSPAC Law and Ethics Committee and the 3 Local Research Ethics Committees. For the physical activity study, verbal consent was given by the child and main carer.

Rationale for Selection of Measurement Instruments

The MTI Actigraph (Manufacturing Technology Incorporated [MTI], Fort Walton Beach, FL) is a uni-axial accelerometer that allows volumes and patterns of physical

activity to be measured with substantially increased precision in comparison to self-report methods.[6] The Actigraph has been calibrated in both children and adolescents against heart-rate telemetry,[10] indirect calorimetry,[11] observational techniques,[12] and energy expenditure measured by doubly labeled water.[13]

Data Collection During Pregnancy

At 12 weeks gestation, the mother and her partner were asked about their height and weight (prepregnancy weight for the mother). Body mass index for mother and partner was calculated by dividing weight (kg) by height squared (m^2). At 32 weeks gestation, the mother was asked to record her highest education level and her partner's highest education level (5 categories from basic high school to degree; the 2 lowest categories were combined, giving 4 categories). The mother also recorded the occupation of both herself and her partner, and this information was used to allocate them to social-class groups (classes I to V, with III split into nonmanual and manual, giving 6 categories; the 2 lowest and the 2 highest were combined for the analysis presented here, giving 4 categories) using the 1991 Office of Population Censuses and Surveys classification.[14] Gestational age was estimated using the date of the last menstrual period as reported by the mother at enrollment and the date of delivery. Infant gender and birth weight were recorded at delivery and abstracted from obstetric records or birth notifications.

Data Collection at Age 11

Height and weight of the children at age 11 were measured with a Harpenden Stadiometer (Holtain Ltd, Crosswell, UK) and a Tanita body-fat analyzer and weighing scale (Model TBF 305, Tanita UK Ltd Middlesex, UK). All children who attended the ALSPAC study clinic at age 11 were asked to wear an MTI Actigraph AM7164 2.2 accelerometer for 7 days.

Physical Activity Measurement Protocol

Actigraphs were normally initialized (using Actigraph Reader Interface Unit RIU-41A with RIU software version 2.26B, MTI Health Services, Fort Walton Beach, FL) to start recording at 5 AM on the day following each child's clinic visit. An epoch time of 1 minute was used. Each child was asked to begin wearing the Actigraph on the right hip on the morning following the clinic visit. Children were asked to wear the Actigraphs during waking hours and to take it off only for showering, bathing, or any water sports. Children were asked to post the Actigraph back. A daily time sheet was provided for the child to record the times they put on and took off the Actigraph and the reason for doing so. They were also asked to record any times (in minutes) that they swam or cycled each day. Data from the returned Actigraphs were downloaded using the Actigraph Reader Interface Unit and software. The raw data were then imported using customized software into a Microsoft Access 2000 database. The software produced a series of derived variables describing levels and patterns of physical activity (see Table 1). These variables were chosen because they provide data that will answer some current questions regarding children's physical activity. For example, are current recommendations of 60 minutes per day of MVPA sufficient for maintaining and improving health,

and can these be accumulated in short bouts? Cut points for moderate and vigorous physical activity (≥3600 and ≥6200 counts/min, respectively) were derived from a calibration study of 246 children in which Actigraph counts/min were compared with oxygen uptake.[15] The sedentary cut point was similar to that used by Treuth et al,[16] who defined sedentary as <50 counts per 30 seconds. The software used in this study derived categories of physical activity intensity in blocks of 200 counts/min, and sedentary was defined as 0 to 199 counts/min.

Instruments were calibrated with every battery change—about every 6 months. Over the 2-year data-collection period, 267 instruments were used (acquired in batches over the study period), and of these, approximately 15 developed faults during the course of the study. A total of 518 calibrations were carried out according to manufacturers' specifications. Of the 518 calibrations, 394 (77%) required no adjustment.

Data Management

Ten or more minutes of consecutive zeros were regarded as periods in which the monitor was unworn, and these were deleted from each file.[7] If on any 1 day the average counts/min was less than 150 or the average counts/min was more than 3 SDs above the mean,[17] we excluded this day of recording because we considered this level of physical activity to be behaviorally implausible. In our calibration study,[15] the children were asked to walk briskly while wearing an Actigraph (mean walking speed 5.8 kph). Counts/min ranged from 1816 to 7136. The mean plus 3 SDs

Table 1 Descriptive Statistics of Physical Activity Summary Measures Derived From Raw Data

Summary measure	Mean or median	SD or IQR	Range
Number of valid days	5.9	1.2	3–7
Mean total counts	2,758,408	993,916	577,701–7,361,597
Mean total min	4585	1026	1895–6777
h/weekday	13.1	0.9	10.0–18.3
h/weekend day	12.3	1.2	10.0–20.5
Counts/min	604	178	204–1520
MVPA[a,b]	19.7	11.7, 31.0	0.3–125.5
Number of bouts of 5–9 min of MVPA[a]	0.6	0.2, 1.3	0.0–6.4
Number of bouts of 10–19 min of MVPA[a]	0.0	0.0, 0.25	0.0–3.3
Number of bouts of ≥20 min of MVPA[a]	0.0	0.0, 0.0	0.0–1.3
Number of bouts of ≥30 min of sedentary[a,c]	0.7	0.3, 1.2	0.0–4.8

Abbreviations: IQR, interquartile range; MVPA, moderate-to-vigorous physical activity.

[a] Median.

[b] MVPA defined as ≥3600 counts/min.

[c] Sedentary defined as ≤199 counts/min.

before removal of spurious data was 1665 counts/min, and we felt it unlikely that a child could sustain this level of intensity for an entire day. Also, as part of the same calibration study, children were asked to lie still for 5 minutes and then to sit still for 5 minutes while wearing an Actigraph. Although 88% of children managed to lie still enough to accrue no counts and 77% managed to sit still enough to accrue no counts, 6 children did have from 60 to over 100 counts/min for each period of either lying or sitting. We felt it was unlikely that many children could maintain a level of average activity below 150 counts/min over an entire day. A day was considered to be valid if the monitor was worn for at least 600 minutes.

Statistical Analyses

Statistical analyses were carried out using Stata Version 8.0 for Windows (Stata Corporation, College Station, TX). Means and standard deviations were calculated for continuous variables and proportions were calculated for categorical variables. Differences between continuous variables were tested using *t* tests and analysis of variance (ANOVA), and differences between categorical variables were tested with chi squared tests. The number of days of monitoring and the number of minutes per day required to achieve reliabilities of .7, .8, and .9 were calculated using the Spearman–Brown prophecy formula,[18] which uses the intraclass correlation coefficient (ICC) as a measure of reliability. The ICC is defined as the ratio of between individual variance to the sum of the between- and within-individual variance.[18] The ICC for a single day of monitoring was calculated from formula (1), where σ_b^2 is the between individual variance and σ_w^2 is the within individual variance.[19]

$$(1)\quad ICC_s = \sigma_b^2/(\sigma_b^2 + \sigma_w^2)$$

The formula for estimating the number of days of measurement to achieve a specified reliability is shown in equation (2), where N is the number of days required to achieve ICC_t, the desired reliability, and ICC_s is the single day ICC from equation (1).

$$(2)\quad N = [ICC_t/(1 - ICC_t)][1 - ICC_s/ICC_s]$$

Power to detect a difference of 0.07 SDs ($P \leq .05$) in counts/min between any 2 groups was also calculated for various combinations of numbers of days of measurement and hours per day. To test for instrument reactivity, we used a linear regression model that specifically allowed for clustering in the data to examine associations between total activity and day of measurement. The cluster option was used with the regress command in Stata to specify that the observations were independent across groups (ie, individuals) but not within groups. This allows for repeated observations on individuals without violating the assumption of independence in the data.[20]

Results

A total of 11,952 children were invited to come to the 11-year clinic, of whom 7159 (60%) came for assessment and 6622 (93%) agreed to wear an Actigraph (Figure 1). Some of the variables derived by the macro are summarized in Table 1.

Number of Minutes and Days of Measurement

Table 2 shows the reliability coefficients for different combinations of number of days and number of minutes per day. The single-day ICC for 600 minutes of measurement was .45. Single-day ICCs (from equation (1)) for different numbers of minutes per day were similar. Table 3 shows power (to detect a difference of 0.07 SDs in counts/min between any 2 groups, $P \leq .05$ to .07 SDs is equivalent to

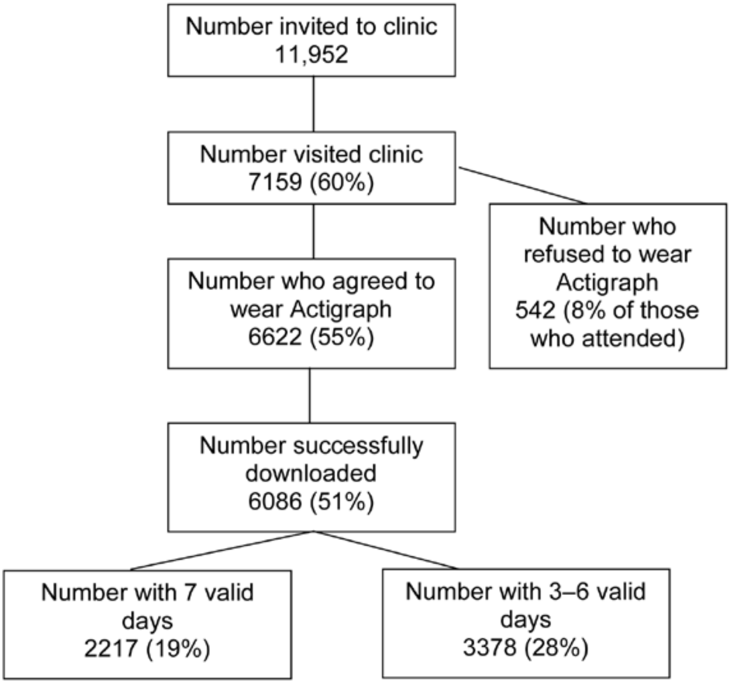

Figure 1 — Flow of physical activity study participants through the ALSPAC clinic

Table 2 Reliability of Different Combinations of Minutes per Day and Days of Measurement

		Days of measurement[b]		
min/day	ICC[a]	*R* = .7	*R* = .8	*R* = .9
600	.45	2.9	4.9	11
540	.44	3.0	5.1	11.5
480	.44	3.0	5.1	11.5
420	.43	3.1	5.3	11.9

Abbreviations: ICC, intraclass correlation coefficient (interindividual variation/total variation).

[a] Based on maximum number of valid days available.

[b] Predicted by Spearman–Brown prophecy formula.

Table 3 Power and Sample Size for Different Combinations of Minutes per Day and Days of Measurement

	Number of days							
	≥3 (any type)		≥3 (≥2 weekday + ≥1 weekend)		≥5 (any type)		≥5 (including ≥1 weekend)	
min/day	N	Power[a] (%)	N	Power (%)	N	Power (%)	N	Power (%)
600	5601	90.8	4980	87.4	4760	89.7	4543	88.3
540	5717	91.7	5284	89.5	5073	91.9	4924	91.1
480	5780	91.9	5448	90.4	5284	92.8	5172	92.3
420	5812	92.3	5529	91.1	5397	93.7	5304	93.3

[a] Power to detect a difference of .07 SDs in counts/min between 2 groups ($P \leq .05$). One SD = 178 counts/min; .07 SD = 13 counts/min.

about 13 counts/min in our sample) and sample size for different combinations of number of days of measurement and number of minutes per day. Weekday mean counts/min was slightly higher than weekend mean counts/min (16 counts/min; 95% CI, 10–22), based on those with at least 3 days recording. Examination of various combinations of days and minutes per day (540, 480, and 420 minutes) revealed little difference in power. An *a priori* decision to specify a valid day as 600 minutes was taken in order to reduce variation in day length and its potential to affect counts/min and sedentary and light activities[21] and to allow direct comparison with the European Youth Heart Study that used the same criteria.[7] Data were considered valid if a child had at least 3 days of at least 600 minutes per day recorded. This combination gave reasonable reliability (Table 2), power >90% (Table 3), and ensured a sufficient sample size for future analyses. Although a weekend day was not specified in order to fulfill validity criteria, 90% of children had at least 1 weekend day of recording.

Final Numbers With Valid Measurement

Applying the above criteria gave us a final sample of 5595 returned Actigraphs that satisfied the validity criteria—2662 boys and 2933 girls (Figure 1). Of the 1027 children who were excluded, 171 were excluded because of broken or malfunctioning instruments and the remainder because the monitor was worn for an insufficient amount of time.

Differences Between Participants and Nonresponders

Children who provided valid recordings differed from children who failed to provide valid recordings in terms of age, weight, body mass index, sex, and pubertal status, but the size of the differences were small (Table 4). More girls than boys returned instruments with valid data (81% of girls versus 76% of boys; $P < .001$). Parental variables were not strongly associated with compliance. Children were more likely to comply if their mothers had a higher level of education, but again, the differences were small.

Table 4a Comparison of Children Who Had Valid Data With Those Who Did Not, Continuous Child Variables

Characteristic	Attended clinic but did not have valid data, N = 1564, mean (SD)	Attended clinic and had valid data, N = 5595, mean (SD)	*P* value
Age (y)	11.81 (0.26)	11.77 (0.23)	< .001
Height (cm)	151.1 (7.3)	150.7 (7.2)	.097
Weight (kg)	44.9 (11.1)	43.5 (9.9)	< .001
Body mass index (kg/m^2)	19.5 (3.8)	19.0 (3.3)	< .001
Birth weight (g)	3445 (537)	3433 (523)	.43

Table 4b Comparison of Children Who Had Valid Data With Those Who Did Not, Categorical Child Variables

	Attended clinic but did not have valid data, N = 1564, percentages	Attended clinic and had valid data, N = 5595, percentages	*P* value
Gender (male)	53.7	49.1	< .001
Pubertal stage (% above Tanner stage 1)	73.6 (70.3, 76.9)	78.2 (76.7, 79.7)	.009
Parity[a]			
0	43.8	43.4	.66
1	32.0	32.3	
2	13.2	12.4	
≥3	11.0	11.9	

[a] Parity was recorded at 18 weeks of gestation by self-report questionnaire and is defined as the number of previous pregnancies resulting in live or still births.

Table 4c Comparison of Children Who Had Valid Data With Those Who Did Not, Continuous Parental Variables

	Attended clinic but did not have valid data, N = 1564, mean (SD)	Attended clinic and had valid data, N = 5595, mean (SD)	*P* value
Maternal height[a] (cm)	164.3 (6.8)	164.2 (6.6)	.50
Maternal BMI[a] (kg/m^2)	23.0 (3.9)	22.9 (3.7)	.34
Paternal height[a] (cm)	176.0 (7.1)	176.4 (6.8)	.06
Paternal BMI[a] (kg/m^2)	25.1 (3.3)	25.1 (3.3)	.96
Maternal age	29.0 (4.8)	29.0 (4.6)	.78

Abbreviation: BMI, body mass index.

[a] Maternal and partner height and body-mass-index data were from self-report questionnaire at 12 weeks of gestation.

Instrument Reactivity, Number of Days of Measurement, and Start Day

The mean difference between total activity on day 1 and the mean of total activity on the remaining days was 17 counts/min higher on day 1 (95% CI, 10–24) or about 0.1 SD. Linear regression, allowing for multiple measurements per child, indicated that day 1 of measurement tended to show slightly higher activity levels than subsequent days (*P* for trend <.001). This remained unchanged after adjustment for gender. There was a difference between the activity levels of children with different numbers of valid days of measurement, and this increased slightly after adjustment for confounding factors (Table 5). There were also differences in total

Table 4d Comparison of Children Who Had Valid Data With Those Who Did Not, Categorical Parental Variables

	Attended clinic but did not have valid data, N = 1564	Attended clinic and had valid data, N = 5595	*P* value
Social class[a]			
1	27.8	29.7	.19
2	26.2	27.2	
3	29.5	26.6	
4	16.6	16.5	
Maternal education[a]			
1	14.5	16.4	.04
2	26.6	26.7	
3	34.4	35.8	
4	24.5	21.1	
Paternal education[a]			
1	21.9	22.0	.66
2	27.6	28.6	
3	21.2	21.9	
4	29.2	27.5	

[a] Socioeconomic variables and parental education from self-report questionnaire at 32 weeks of gestation (coded as 1 = highest, 4 = lowest).

Table 5 Mean Counts/Min by Number of Days of Measurement and by Start Day

Number of valid days	Mean counts/min[a,c]	Frequency	Mean counts/min[b,c]	Frequency
3	630	314	618	130
4	622	527	636	269
5	619	951	626	488
6	603	1586	606	828
7	590	2217	595	1291
Total	604	5595	608	3006
	$\beta = -11$ (–14, –7), $P < .001$		$\beta = -11$ (–16, –6), $P < .001$	
Start day	**Mean counts/min[a,c]**	**Frequency**	**Mean counts/min[b,d]**	**Frequency**
Monday	558	112	601	62
Tuesday	602	486	599	255
Wednesday	628	997	627	533
Thursday	608	924	620	479
Friday	605	955	602	526
Saturday	607	979	604	549
Sunday	582	1142	592	549
Total	604	5595	608	3006

[a] Unadjusted.
[b] Adjusted for age, gender, pubertal status, body mass index, and maternal and paternal education level and social class.
[c] $P = .001$.
[d] $P = .007$.

activity levels depending which day measurement started on (Table 5). There was a small difference in activity depending on whether children started on a weekday or weekend day (mean difference 17 counts/min; 95% CI, 7–29, $P < .001$). This represented about 0.1 of a SD. A total of 3474 children (62%) started on a weekday, and 2121 (38%) started on a weekend day.

Discussion

The study provides an overview of the accelerometry processing and compliance patterns in a large study of children. We have shown that 3 days of measurement resulted in good reliability and power and that systematic differences in counts/min between numbers of days of measurement were small. Furthermore, we have described the characteristics of children who completed the protocol and those that did not and have shown that these differences were modest. There was some potential for bias depending on which day of the week children started wearing the Actigraph.

Compliance With the Study Protocol

Of the children who attended the 11-year clinic, 78% provided valid data. This is consistent with results from previous studies. Riddoch et al measured physical activity in 9- and 15-year-old children in 4 European countries. Using similar procedures to this study to exclude nonvalid data, 75% of children who took part in the study had valid data for at least 3 days for at least 10 hours per day.[7] In a feasibility study of accelerometry in children from grades 6 through 8, Van Coevering et al found that 234 of 282 children (83%) provided 3 valid days of recording.[8]

Number of Days of Measurement and Start Day

The use of 3 days of physical activity measurement gave good reliability ($R = .7$; Table 2). Although it has been previously suggested that a minimum of 4 days of measurement is required to give a .8 reliability coefficient,[22] we feel that a reliability coefficient of .7 is acceptable and justified on the basis that it maximizes power and reduces the number of participants excluded for future analyses. The small difference in counts/min between weekdays and weekend days and similar power suggests that including a weekend is not necessary in our sample, although this might not be the case for smaller studies. Similar single-day ICCs for different minimum acceptable day lengths gave similar estimates for the number of days required to achieve prespecified reliabilities. Despite this, 600 minutes was chosen as the minimum day length as it (1) minimizes the possible effects of varying day length on physical activity outcomes and (2) allows comparison with other studies. Use of the Spearman–Brown formula does have some limitations as it relies on compound symmetry (ie, the variances for each day are equal and correlations between pairs of days are equal.)[23] Correlations for counts/min between pairs of days in this study ranged from .39 to .48, and we feel that this range in the data is unlikely to constitute a problem.

Children with fewer days of recording tended to have higher total physical activity (counts/min), but these differences were modest; the biggest difference

(40 counts/min or about 0.2 SDs) was between 3 days of measurement (the minimum) and 7 days of measurement (the maximum). We found that there were differences in total counts/min depending on the start day. Children who started on a Monday had the lowest counts/min, and children who started on a Wednesday had the highest counts/min. Saturday was the most popular day to attend the clinic, and counts/min was also lower than average when children started on this day. This is difficult to explain and might represent a chance finding, although it might be that children with a Monday start day were a different group from the rest because there was no clinic on Sundays, and a Monday start day had to be specifically requested. There was also a difference in activity depending on whether children started on a weekday or weekend day, although the difference was small (17 counts/min or 0.1 SDs).

Instrument Reactivity

Reactivity (the tendency of the instrument to modify normal behavior) has been highlighted as a potential problem in the measurement of physical activity.[19] Although a difference in total activity between the first day and the mean of all subsequent days was found, this was small at 17 counts/min (about 0.1 SDs) and it is unlikely that this would introduce bias into this study.

Differences Between Participants and Nonresponders

There were small differences between children with and without valid Actigraph data. Boys were less likely to provide valid data, and children with valid data had lower body mass index, although the difference was small. This is in contrast to Van Coevering et al who found that overweight children were more likely to provide 7 days of complete data.[8] Children with valid data tended to be younger and lighter, and more children with valid data were in later stages of pubertal development compared with those who did not provide valid data, but again, these differences were small. There were also some differences in terms of maternal education—a marker of socioeconomic position. Those who provided valid data tended to have mothers with higher educational levels. The difference was small, however, and unlikely to introduce bias because the association between socioeconomic position and physical activity in children in this population is weak and inverse (ie, lower socioeconomic position is associated with slightly higher total physical activity).[24] These results suggest subject characteristics that might be targeted to maximize compliance.

Limitations of the Study

Because of the large volume of data collected, it was not possible to examine each Actigraph file individually to check for errors. This might have resulted in spurious patterns of data being accepted as valid. However, we feel the stringent exclusion criteria for dealing with outliers during the data cleaning kept implausible data to a minimum. Most studies, including ours, use a 1-minute measurement epoch. There has been considerable debate on the most appropriate epoch length to use.[19] The concern is that using a 1-minute epoch might mask the short bursts of vigorous activity that are typical in children[25] by averaging them over 1 minute.

Nilsson et al[26] found that vigorous and very vigorous activity was underestimated when a 1-minute epoch was compared with a 5-second epoch. This might result in misclassification of children in these categories of intensity. This would not, however, affect the main outcome measure of this study, which was volume of physical activity assessed by counts/min.

Although we feel that the results of our study will be informative for other researchers when deciding on study methodology, we recognize that our analyses and the decisions we made might not be appropriate for all studies. For example, our relatively large sample maximized power and gave reasonable reliability with 3 days of recording. A duration of 3 days is shorter than has previously been recommended,[22] although other large studies have used 3 days.[7] Studies with smaller sample sizes might wish to measure a greater number of days and to process data by hand to improve measurement precision.

Conclusions

We have demonstrated the feasibility of incorporating an objective measure of physical activity into a large, ongoing birth cohort. The use of an objective measure of physical activity in this large field-based study of children will allow us to describe the determinants of physical activity at age 11 and to examine the health consequences of physical activity in children more accurately than has previously been possible. Furthermore, we have reported on a number of design issues and how they affect the validity of accelerometer data, which will be of use to researchers who wish to use accelerometry in large studies of children.

Acknowledgments

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. We would also like to thank Sue Ingle for additional support with statistical analyses. The UK Medical Research Council, the Wellcome Trust, and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors, and Calum Mattocks and Sam Leary will serve as guarantors for the contents of this article. This research was specifically funded by a grant from the US National Heart, Lung and Blood Institute (R01 HL071248-01A).

References

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep.* 1985;100:126-131.
2. Department of Health. *At Least Five a Week: Evidence on the Impact of Physical Activity and its Relationship to Health*. A Report from the Chief Medical Officer. London, UK: Dept of Health; 2004.
3. Twisk JW. Physical activity guidelines for children and adolescents: a critical review. *Sports Med.* 2001;31:617-627.
4. Strong WB, Malina RM, Blimkie CJR, et al. Evidence based physical activity for school-age youth. *J Pediatr.* 2005;146:732-737.

5. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Med.* 2001;31:439-454.
6. Welk GJ. Use of accelerometry-based activity monitors to assess physical activity. In Welk GJ. *Physical Activity Assessments for Health-Related Research.* Champaign, IL: Human Kinetics; 2002: 125-141.
7. Riddoch CJ, Andersen LB, Wedderkopp N, et al. Physical activity levels and patterns of 9 and 15 yr-old European children. *Med Sci Sports Exerc*. 2004; 36:86-92.
8. 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.
9. Golding J, Pembrey M, Jones R. ALSPAC—the Avon Longitudinal Study of Parents and Children: I. study methodology. *Paediatr Perinat Epidemiol.* 2001;15:74-87.
10. Janz KF. Validation of the CSA accelerometer for assessing children's physical activity. *Med Sci Sports Exerc.* 1994;26:369-375.
11. Melanson EL, Freedson PS. Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. *Med Sci Sports Exerc.* 1995;27:934-940.
12. Fairweather SC, Reilly JJ, Grant S, Whittaker A, Paton JY. Using the Computer Science and Applications (CSA) activity monitor in preschool children. *Pediatr Exerc Sci.* 1999;11:413-420.
13. Ekelund U, Sjostrom M, Yngve A, et al. Physical activity assessed by activity monitor and doubly labeled water in children. *Med Sci Sports Exerc.* 2001;33:275-281.
14. Her Majesty's Stationary Office. *Standard Occupational Classification.* London, UK: Office of Population and Census Surveys; 1991.
15. Mattocks C, Ness A, Leary S, et al. Calibration of an accelerometer during free-living activities in children. *Int J Pediatr Obes.* DOI: 10.1080/17477160701408809.
16. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. *Med Sci Sports Exerc.* 2004;36:1259-1266.
17. Tremblay MS, Barnes JD, Copeland JL, Esliger DW. Conquering childhood inactivity: is the answer in the past? *Med Sci Sports Exer.* 2005;37:1187-1194.
18. Streiner DL, Norman GJ. *Health Measurement Scales: A Practical Guide to Their Development and Use.* 2nd ed. Oxford, UK: Oxford University Press; 1995.
19. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exer.* 2005;37(suppl 11):S531-S543.
20. Stata Corporation. *Stata Base Reference Manual.* Vol 3. Reference N-R. Release 8. College Station, TX: Stata Press; 2003.
21. Masse LC, Fuemmeler BF, Anderson CB, et al. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. *Med Sci Sports Exerc.* 2005;37(suppl 11):S544-S554.
22. 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.
23. Baranowski T, Smith M, Thompson WO, Baranowski J, Hebert D, de Moor C. Intraindividual variability and reliability in a 7-day exercise record. *Med Sci Sports Exerc.* 1999;31:1619-1622.
24. Riddoch C, Mattocks C, Leary S, et al. Physical activity levels and patterns of 11-year old children: findings from the Avon Longitudinal Study of Parents and Children. *Arch Dis Child.* In press.
25. Bailey RC, Olson J, Pepper SL, Porszasz J, Bartow TJ, Cooper DM. The level and tempo of children's physical activities: an observational study. *Med Sci Sports Exerc.* 1995;27:1033-1041.
26. Nilsson A, Ekelund U, Yngve A, Sjöström M: Assessing physical activity among children with accelerometers using different time sampling intervals and placements. *Pediatr Exerc Sci.* 2002;14:87-96.