Journal of Sports Sciences

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rjsp20

The influence of reliability and variability of objectively measured physical activity on associations with lower body muscle strength in young children

Eivind Aadland & Ada Kristine Ofrim Nilsen

To cite this article: Eivind Aadland & Ada Kristine Ofrim Nilsen (2023): The influence of reliability and variability of objectively measured physical activity on associations with lower body muscle strength in young children, Journal of Sports Sciences, DOI: 10.1080/02640414.2023.2204587

To link to this article: https://doi.org/10.1080/02640414.2023.2204587

Published online: 20 Apr 2023.

JOURNAL OF SPORTS SCIENCES
https://doi.org/10.1080/02640414.2023.2204587

Routledge
Taylor & Francis Group

PHYSICAL ACTIVITY, HEALTH AND EXERCISE

The influence of reliability and variability of objectively measured physical activity on associations with lower body muscle strength in young children

Eivind Aadland and Ada Kristine Ofrim Nilsen

Western Norway University of Applied Sciences, Faculty of Education, Arts and Sports, Department of Sport, Food and Natural Sciences, Sogndal, Norway

ABSTRACT

It is not known how extended or multiple monitoring periods affect associations between accelerometer-measured physical activity and outcomes. The aim of this study was to examine how accelerometer monitoring length influenced cross-sectional and prospective associations for physical activity with lower body muscle strength in young children. 176 Norwegian 2–6-year-old children had 3 valid 14-day periods of accelerometer monitoring (ActiGraph GT3×+) between September 2015 and May 2016 (baseline) as well as baseline and 4-year follow-up measurements of standing long jump. We analysed physical activity using a descriptor with 4 intensities using 6 different monitoring lengths both within and across monitoring periods (1 day, 3 days, 1 week, 2 weeks, 3 weeks, 6 weeks) and determined associations with lower body muscle strength using multivariate pattern analysis. We found that the strength of cross-sectional associations with lower body muscle strength improved for longer monitoring periods (explained variances = 7.7%, 9.8%, 11.8%, 15.8%, 18.4% and 22.9% for 1 day, 3 days, 1 week, 2 weeks, 3 weeks and 6 weeks of measurement). Longitudinal associations were weaker and less clear. Our findings suggest that multiple extended physical activity monitoring periods improve the validity of the study findings regarding associations between physical activity and relevant outcomes.

ARTICLE HISTORY
Received 21 October 2022
Accepted 11 April 2023

KEYWORDS
Reliability; measurement error; regression dilution bias; accelerometry, muscle fitness

Introduction

Physical activity (PA) levels vary over time (i.e., days, weeks, seasons), and many studies have sought to determine the optimal or sufficient wear time to characterize habitual PA as measured by accelerometry (Aadland & Johannessen, 2015; Aadland & Ylvisåker, 2015; Aadland et al., 2017, 2020; Addy et al., 2014; Baranowski et al., 2008; Basterfield et al., 2011; Bisson et al., 2018; Coleman & Epstein, 1998; Hart et al., 2011; Hinkley et al., 2012; Hislop et al., 2014; Janz et al., 1995; Jerome et al., 2009; Kang et al., 2014; Levin et al., 1999; Matthews et al., 2002, 2012; Mattocks et al., 2007; Murray et al., 2004; Ojiambo et al., 2011; Penpraze et al., 2006; Rich et al., 2013; Treuth et al., 2003; Trost et al., 2000, 2005; Wickel & Welk, 2010). The number of days or periods of monitoring that should be included to obtain reliable estimates of habitual PA levels is therefore an important aspect of accelerometer measurements, but the criteria applied to define what constitutes a reliable PA measurement vary extensively (Cain et al., 2013). By increasing the number of days or periods of measurement, random errors in measurements will decrease and the likelihood of type II errors due to regression dilution bias (i.e., attenuation of associations) can be minimized (Hutcheon et al., 2010). Variability in PA measurements will inherently also lead to variability in findings, for example, for associations between PA and an outcome, which might increase the risk of chance findings and type I errors.

In both adults (Coleman & Epstein, 1998; Hart et al., 2011; Jerome et al., 2009; Matthews et al., 2002; Trost et al., 2005) and children (Addy et al., 2014; Basterfield et al., 2011; Bisson et al., 2018; Hinkley et al., 2012; Hislop et al., 2014; Janz et al., 1995; Kang et al., 2014; Murray et al., 2004; Ojiambo et al., 2011; Penpraze et al., 2006; Rich et al., 2013; Treuth et al., 2003; Trost et al., 2000), most evidence suggest that a reasonable reliability (i.e., intra-class correlation (ICC)) of approximately 0.70–0.80 are achieved with 3–7 days of monitoring. However, most previous estimates are derived from the Spearman Brown prophecy formula applied to measurements conducted over a single 7-day period. Previous research has suggested that this approach is inadequate because it leads to an overestimation of reliability (Aadland et al., 2017, 2020; Baranowski et al., 2008; Matthews et al., 2012; Wickel & Welk, 2010). In comparison, studies that have determined the reliability over the course of multiple weeks (Aadland & Johannessen, 2015; Aadland & Ylvisåker, 2015; Aadland et al., 2020) or multiple measurement periods (Aadland et al., 2017, 2020; Levin et al., 1999; Mattocks et al., 2007; Wickel & Welk, 2010) have shown considerable variability over time. Studies including multiple monitoring periods over several seasons have resulted in reliability estimates of approximately 0.50 for 1-week monitoring in children (Aadland et al., 2017, 2020; Mattocks et al., 2007; Wickel & Welk, 2010). Using a sample of 873 children providing one 14-day monitoring period and 221 children providing three separate 14-day monitoring periods, Aadland et al. (Aadland et al., 2020) estimated that an ICC of 0.80 could be achieved with 5.6 and 8.1 monitoring days (means for the PA variables) using a day-by-day approach and the Spearman-Brown formula within 1 and

CONTACT Eivind Aadland eivind.aadland@hvl.no Department of Sport, Food and Natural Sciences, Faculty of Education, Arts and Sports, Western Norway University of Applied Sciences, Campus Sogndal, Box 133, 6851, Sogndal, Norway

across 3 monitoring periods, respectively. However, when reliability was measured across multiple periods, ICC increased marginally from 0.40 for 1 day to 0.51 for 7 days and 0.52 for 14 days of monitoring. Thus, reliability appears to be minimally affected by the number of monitoring days and levels off after 5–6 monitoring days. This finding is consistent with rather similar reliability of different numbers of monitoring hours per day and days per week when reliability is measured across multiple weeks (Aadland & Johannessen, 2015; Aadland & Ylvisåker, 2015; Aadland et al., 2017) but is contrary to the belief that a longer monitoring period better captures habitual PA that varies over time.

While previous studies have concluded that a single measurement period does not adequately characterize habitual PA (Aadland et al., 2020; Levin et al., 1999; Mattocks et al., 2007; Wickel & Welk, 2010), to the best of our knowledge, no studies have examined to which extent multiple monitoring periods affect associations with relevant outcomes through improved precision of PA estimates. Physical fitness and motor skills are highly related because tasks for both constructs are underpinned by neuromuscular control and because constructs and content of measures overlap (Cattuzzo et al., 2016; Utesch et al., 2019). Lower body muscle strength is a relevant measure of physical fitness in young children and has previously been associated with PA both cross-sectionally and longitudinally (Aadland et al., 2022; Leppänen et al., 2016, 2017). The aim of this study was to examine how accelerometer monitoring length and variability of PA influenced cross-sectional and prospective associations for PA with lower body muscle strength in young children both within and across 3 separate 14-day PA monitoring periods.

Materials and methods

Design and participants

This study has a cross-sectional and longitudinal design and was based on data from the *Sogn og Fjordane Preschool Physical Activity Study (PRESPAS)* and the *PRESPAS follow-up study* conducted 2015–2016 and 2016–2019, respectively. PRESPAS was a cross-sectional study conducted in the county of Sogn og Fjordane, a rural area in Western Norway, between September 2015 and June 2016 and involved a total of 1308 children aged 2.7–6.5 years (born in 2010–2012) from 68 preschools (response rate 68%) from 14 municipalities (Nilsen, Anderssen, Ylvisåker, et al., 2019). From the PRESPAS sample, we invited a subsample of 376 children from 20 preschools (all preschools from 3 municipalities participating in PRESPAS) to participate in 3 repeated measurements at baseline (2015–2016) and repeated measurements up to 2019 (*PRESPAS follow-up study)*.

This study included children who provided PA data (explanatory variables) at 3 timepoints at baseline (September – October 2015, January – February 2016 and April – May 2016) and data on standing long jump (outcome variable) at baseline and follow-up in 2019.

Parents of all participating children received oral and written information about the PRESPAS studies and provided written consent prior to testing for both the baseline and the follow-up measurements. Preschools and primary schools received information and agreed to participate in the study. We explained the procedures to the children according to their age and level of understanding. The Norwegian Centre for Research Data (NSD) approved the studies (reference numbers: 39061 (PRESPAS) and 48,016 (PRESPAS follow-up)).

Procedures

Physical activity measurements

PA was measured using the ActiGraph GT3×+ accelerometer (ActiGraph, LLC, Pensacola, Florida, USA) (John & Freedson, 2012). Children wore an elastic belt with the accelerometer on their right hip and were instructed to wear the monitor at all times except during water-based activities and while sleeping (at night) for 14 consecutive days each monitoring period. Accelerometers were initialized at a sampling rate of 30 Hz. Files were analysed restricted to hours 06:00 to 23:59. Data were analysed using a 1-second epoch to capture low- and high-intensity PA (Aadland, Andersen, Anderssen, et al., 2019). Periods of ≥20 min of zero counts were defined as non-wear time (Esliger et al., 2005). We used count-based data from the vertical axis to create a PA descriptor of 4 intensity variables of minutes/day spent in intensities ≤100 (sedentary time, SED), 101–2295 (light intensity PA, LPA), 2296–4011 (moderate intensity PA, MPA) and ≥4012 cpm (vigorous intensity PA, VPA) using the previously established and validated Evenson et al. cut points (Evenson et al., 2008; Trost et al., 2011). For descriptive purposes, we additionally reported the number of valid days, wear time and total PA (average cpm).

We applied a wear time requirement of ≥480 min to define a valid day. As reliability is marginally affected by wear hours per day (≥6 to ≥12 hours/day; Aadland & Johannessen, 2015; Levin et al., 1999; Mattocks et al., 2007), we did not analyse the sensitivity to this wear criterion herein. Children that provided ≥4 valid days of data for one of the 2 weeks of each monitoring period were included in the analyses (Aadland et al., 2017, 2020). We analysed PA data using 6 different monitoring lengths to investigate the effect of the number of monitoring days on associations with lower body muscle strength both within a period (1 day, 3 days, 1 week and 2 weeks) and across multiple periods (mean of the first weeks (3 weeks in total) and both weeks (6 weeks in total) of each monitoring period). The data for 1 and 3 monitoring day(s) were derived from the first day(s) of the first week of monitoring. If the first day(s) were invalid, we manually derived data from the first valid day(s). If the first week was invalid (i.e., <4 valid days), we used data for the second week to derive data for 1 day, 3 days and 1 week. PA data was processed using a custom-made script in MATLAB (MathWorks, Massachusetts, USA).

Anthropometrics, lower body muscle strength and demographics

We assessed children's body mass, height and lower body muscle strength during preschool/school hours at all timepoints. Body mass and height were measured 3 times at baseline (once for each PA monitoring period), whereas lower body muscle strength was measured once at baseline. Body mass, height and lower body muscle strength were measured once at follow-up. Body mass was measured to the nearest 0.1 kg using an electronic scale (Seca 899, SECA GmbH, Hamburg, Germany), and height was measured to the nearest 0.1 cm with a portable stadiometer (Seca 217, SECA GmbH, Hamburg, Germany). Body mass index (BMI, kg/m^2) was calculated, and children were classified as normal weight (including underweight), overweight, or obese based on the criteria proposed by Cole et al. (Cole et al., 2000). Lower body muscle strength was measured using the standing long jump test from the Assessing FITness in PREschoolers (PREFIT) battery (Ortega et al., 2015). PREFIT is an adaptation of the ALPHA-Fitness test battery and has demonstrated good reliability in children (Artero et al., 2011; Cadenas-Sanchez et al., 2016; Ortega et al., 2008, 2015). Children were instructed to jump as far as possible from a standing position, with a two-footed take-off and landing. Children's performance was measured to the nearest 1 cm using the best of 2 consecutive attempts after a familiarisation attempt. Parental education (highest education level of mother or father) was assessed by a questionnaire completed by each child's mother and/or father.

Statistical analyses

Children's characteristics and PA were described as frequencies, means and standard deviations (SD). The multivariate PA intensity signatures associated with standing long jump at baseline and the change in standing long jump from baseline to 4-year follow-up were determined using multivariate pattern analysis (Wold et al., 1984). This analysis allows for including any number of PA intensity variables as explanatory variables irrespective of multicollinearity among variables, as shown in other applications with accelerometer data (Aadland et al., 2018, 2021; Nilsen, Anderssen, Loftesnes, et al., 2019). Partial least squares (PLS) regression analyses were used to determine the association patterns between the outcomes and the PA intensities by decomposing the explanatory variables into orthogonal linear combinations (PLS components), while simultaneously maximizing the covariance with the outcome variable (Wold et al., 1984). Models were cross-validated using Monte Carlo resampling (Kvalheim et al., 2018) with 1000 repetitions by repeatedly and randomly keeping 50% of the participants as an external validation set. For each model, we used target projection (Kvalheim & Karstang, 1989; Rajalahti & Kvalheim, 2011) followed by reporting of multivariate correlation coefficients with 95% confidence intervals (CIs) to show the importance of each PA intensity variable in the multivariate space (Aadland, Andersen, Resaland, et al., 2019; Rajalahti, Arneberg, Berven, et al., 2009; Rajalahti, Arneberg, Kroksveen, et al., 2009). Multivariate correlation coefficients can be interpreted equal to bivariate correlations, though they are derived from the multivariate space. We also reported the total explained variance (R^2) of the models to show the joint association for all variables with the outcome. For adjustment, we obtained residuals from linear regression models using standing long jump (adjusted for sex, age and BMI in the cross-sectional analyses, and additionally for baseline values in the prospective analyses) and PA variables (adjusted for accelerometer wear time, sex, age and BMI) as outcomes, prior to performing the multivariate pattern analyses. Multivariate pattern analyses were performed using Sirius version 11.5 (Pattern Recognition Systems AS, Bergen, Norway), and linear regression analyses were performed using IBM SPSS v. 28 (IBM SPSS Statistics for Windows, Armonk, NY; IBM Corp., USA). We omitted reliability analyses since such results have been reported previously in this dataset (Aadland et al., 2020).

Results

Of the 376 participating children, 268 children provided valid PA data for all 3 timepoints at baseline, whereas 239 children provided data for lower body muscle strength at both baseline and follow-up. A total of 176 children (87 boys and 89 girls) provided data for PA at baseline and standing long jump at baseline and follow-up and were included in the analysis (Table 1).

Children's time spent in PA and SED are shown in Table 2. The mean number of monitoring days varied from 1 day for the 1-day descriptor to 34 days for the 6-week descriptor. Levels of PA (SED) were highest (lowest) for monitoring period 3 and lowest (highest) for monitoring period 2. The variation in PA and SED (SD) decreased with increased monitoring length for all monitoring periods.

We found statistically significant associations for all PA descriptors with standing long jump at baseline (Figures 1–2). Associations strengthened when the monitoring length increased from 1 day to 2 weeks for all monitoring periods, with minimum-maximum R^2 of 5.9–11.9, 6.1–17.8 and 11.1–17.7% for 1-day to 2-week monitoring lengths for monitoring periods 1, 2 and 3, respectively (Figure 1). Despite this consistent pattern, there were clear differences in R^2 between PA descriptors both across monitoring periods and across monitoring lengths. Association patterns were marginally influenced by monitoring length but differed between monitoring periods. While associations for VPA were positive and strongest associated with standing long jump for all descriptors, associations for MPA were weaker and non-significant for monitoring period 3 compared to monitoring periods 1 and 2, associations for LPA were non-significant for monitoring periods 1 and 2 but significantly negative for monitoring period 3 (except for the 1-day descriptor), and associations for SED were negative for monitoring periods 1 and 2 (except for the 3-day descriptor for

Table 1. Children's characteristics at each timepoint. Numbers are mean (SD) or %.

	2015–2016	2019
Age (years)	4.8 (0.9)	8.6 (0.9)
Body mass (kg)	19.3 (3.0)	30.7 (6.0)
Height (cm)	109.0 (7.4)	134.1 (7.0)
Body mass index (kg/m^2)	16.2 (1.3)	16.9 (2.3)
Overweight/Obese (%)		
Under or normal weight	82.4	84.7
Overweight	16.5	11.9
Obese	1.1	3.4
Standing long jump (cm)	82.6 (23.8)	123.8 (20.2)
Parental education level (%)		
Upper secondary school	11.5	9.2
University < 4 years	28.5	29.6
University ≥ 4 years	60.0	61.3

n = 165 and 142 for parental education at baseline and follow-up, respectively.

Table 2. Mean (SD) levels of physical activity and sedentary time for the different descriptors.

	1 day	3 days	1 week	2 weeks	3 weeks	6 weeks
			Monitoring period 1			
Wear days (n)	1 (0)	2.8 (0.5)	6.4 (0.8)	12.4 (1.7)	-	-
Wear time (min/day)	728 (81)	714 (56)	705 (43)	701 (39)	-	-
Total PA (cpm)	721 (231)	696 (198)	706 (177)	709 (166)	-	-
SED (min/day)	500 (71)	497 (53)	487 (43)	484 (42)	-	-
LPA (min/day)	152 (30)	146 (23)	147 (20)	146 (20)	-	-
MPA (min/day)	38 (11)	36 (9)	36 (8)	36 (7)	-	-
VPA (min/day)	37 (16)	35 (13)	35 (12)	35 (11)	-	-
			Monitoring period 2			
Wear days (n)	1 (0)	2.6 (0.6)	6.1 (0.9)	11.8 (2.1)	-	-
Wear time (min/day)	701 (85)	700 (63)	699 (47)	699 (43)	-	-
Total PA (cpm)	636 (185)	608 (151)	621 (141)	621 (133)	-	-
SED (min/day)	494 (75)	499 (59)	496 (48)	496 (42)	-	-
LPA (min/day)	142 (31)	138 (25)	139 (20)	139 (19)	-	-
MPA (min/day)	35 (11)	33 (8)	34 (7)	34 (7)	-	-
VPA (min/day)	31 (13)	30 (10)	30 (9)	30 (9)	-	-
			Monitoring period 3			
Wear days (n)	1 (0)	2.2 (0.6)	5.6 (0.9)	9.9 (2.3)	-	-
Wear time (min/day)	669 (90)	690 (65)	685 (49)	685 (44)	-	-
Total PA (cpm)	859 (274)	823 (236)	806 (199)	817 (187)	-	-
SED (min/day)	447 (70)	463 (54)	461 (46)	458 (42)	-	-
LPA (min/day)	144 (31)	150 (26)	149 (21)	150 (20)	-	-
MPA (min/day)	37 (11)	37 (10)	37 (8)	38 (8)	-	-
VPA (min/day)	41 (17)	39 (14)	38 (11)	39 (11)	-	-
			Across monitoring periods			
Wear days (n)	-	-	-	-	18.1 (1.8)	34.1 (4.2)
Wear time (min/day)	-	-	-	-	696 (36)	695 (34)
Total PA (cpm)	-	-	-	-	711 (143)	716 (138)
SED (min/day)	-	-	-	-	481 (38)	479 (36)
LPA (min/day)	-	-	-	-	145 (17)	145 (17)
MPA (min/day)	-	-	-	-	36 (6)	36 (6)
VPA (min/day)	-	-	-	-	34 (9)	35 (9)

PA = physical activity; SED = sedentary time; LPA = light physical activity; MPA = moderate physical activity; VPA = vigorous physical activity; MVPA = moderate- to vigorous physical activity.

monitoring period 2) but non-significant for monitoring period 3. When comparing mean estimates for different monitoring lengths across the monitoring periods with estimates derived from the PA descriptors including all monitoring periods, we found that R^2 improved from 7.7% to 15.8% from 1 day to 2 weeks of monitoring and further to 18.4% and 22.9% for the 3- and 6-week descriptors, respectively (Figure 2). Association patterns were generally similar for all these estimates.

Prospective associations with standing long jump differed between monitoring periods. While we found significant associations for all descriptors for monitoring period 2, there were no significant associations for monitoring periods 1 and 3. For monitoring period 2, R^2 improved from 2.0% to 4.1% from 1 day to 2 weeks of monitoring but decreased for the PA descriptors including all monitoring periods (R^2 = 1.9% and 2.3% for the 3- and 6-week descriptors, respectively) (Figure 3).

Discussion

The present study aimed to determine how the number of monitoring days and monitoring periods affected associations between PA and lower body muscle strength in young children. We found that the strength of cross-sectional associations improved from an R^2 of 5.9% to 22.9% when using 1 day as compared to 6 weeks of monitoring. A 14-day period led to stronger associations than a 7-day period, and 3 monitoring periods led to stronger associations than 1 monitoring period. Similar patterns were evident for longitudinal associations, although these findings were less clear.

As noise in exposure (x) variables leads to attenuation of regression coefficients (regression dilution bias), and noise in outcome (y) variables increases standard errors (Hutcheon et al., 2010), measurement error may lead to low, and potentially non-significant, effect sizes (i.e., type 2 errors). Variability of measurements may also lead to chance findings (i.e., type 1 errors). The low reliability/high variability of measurements inherently means that data are not reproducible (De Vet et al., 2011), which in turn challenges the replicability of study findings. Our study is well designed to examine the influence of reliability and variability of PA on findings regarding associations between PA and an outcome and shows substantial variability in associations across the measurement periods for both the cross-sectional analysis (1 day of measurement resulting in R^2 of 5.9–11.1%; 3 days of measurement resulting in R^2 of 6.8–14.4%; 1 week of measurement resulting in R^2 of 10.4–13.8%; 2 weeks of measurement resulting in R^2 of 11.9–17.8%) and the longitudinal analysis (associations were significant for monitoring period number 2 but not for monitoring periods 1 and 3). These differences can clearly lead to different study conclusions (i.e., some associations being statistically significant and others not), also depending on data reduction algorithms for accelerometry data, sample size, etc. We are not aware of similar studies allowing for a direct comparison with our findings. Yet, our findings suggest that the variability in associations between PA and various outcomes in children reported in the literature (Poitras et al., 2016; Veldman et al., 2021; Wiersma et al., 2020) partly results from measurement errors in PA estimates.

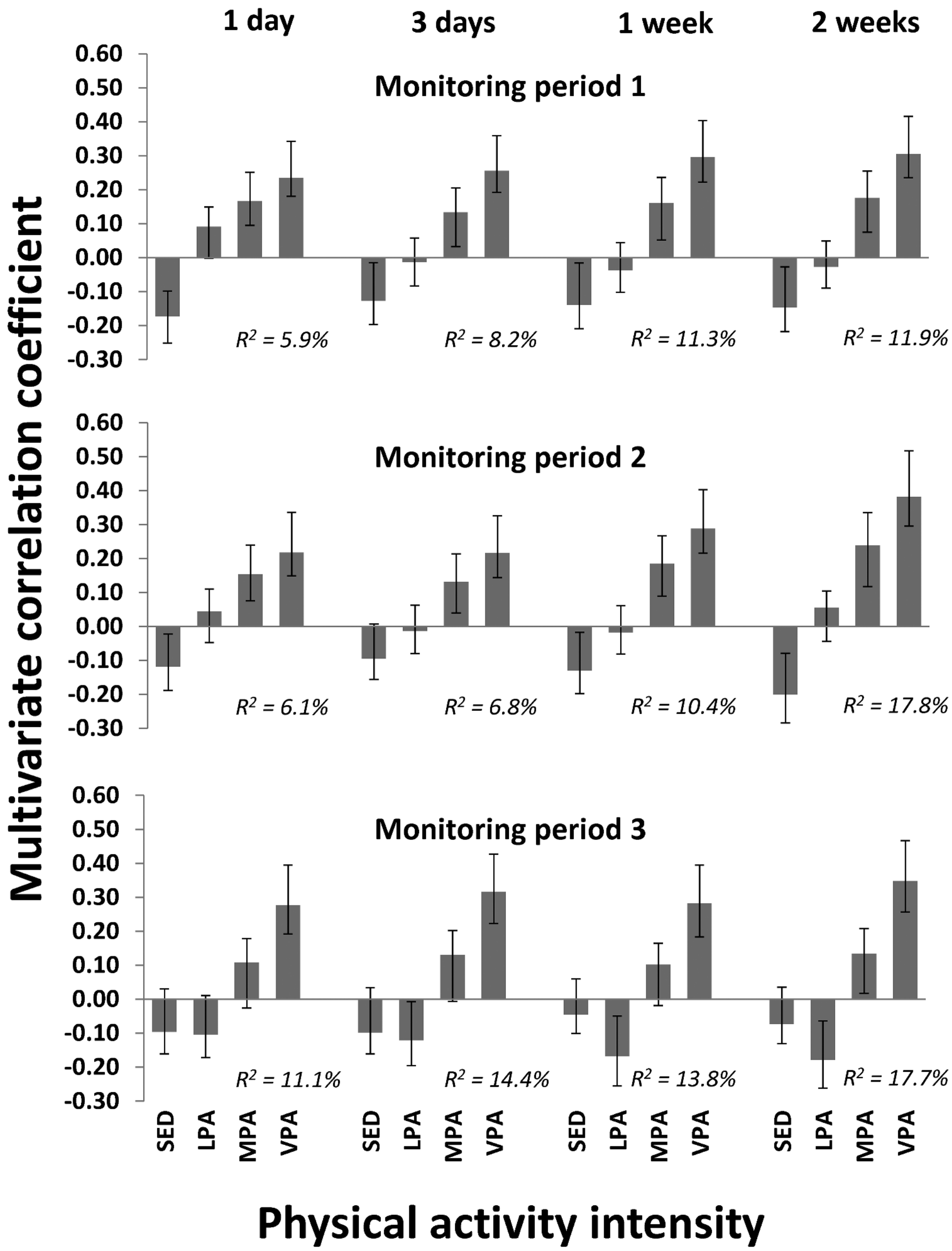

Figure 1. Cross-sectional associations between physical activity and lower body muscle strength for each monitoring period using descriptors derived using different monitoring length. Results are reported as multivariate correlation coefficients from a joint model including the 4 physical activity intensity variables adjusted for sex, age, body mass index and wear time. Multivariate correlation coefficients can be interpreted equivalent to bivariate correlations, although they are derived from the multivariate model. Physical activity assessments were conducted for 14 consecutive days during September – October (monitoring period 1), January – February (monitoring period 2) and April – May (monitoring period 3). SED = sedentary time; LPA = light physical activity, MPA = moderate physical activity, VPA = vigorous physical activity. R^2 = explained variance.

Beyond variability in associations across measurement periods, we found increased strengths of associations from 1 day to 6 weeks of measurement in the cross-sectional analysis (R^2 of 7.7%, 9.8%, 11.8%, 15.8%, 18.4% and 22.9% for 1 day, 3 days, 1 week, 2 weeks, 3 weeks and 6 weeks of measurement). This finding is in line with the literature showing improved reliability with increasing number of monitoring days and monitoring periods. While most previous studies have estimated reliability using a single 7-day monitoring period and concluded that 3–7 monitoring days are sufficient to achieve a reliability of 0.70–0.80 in children aged 2–15 years (Addy et al., 2014; Basterfield et al., 2011; Hinkley et al., 2012; Hislop et al., 2014; Janz et al., 1995; Kang et al., 2014; McGraw & Wong, 1996; Murray et al., 2004; Ojiambo et al., 2011; Penpraze et al., 2006; Rich et al., 2013; Treuth et al., 2003; Trost et al., 2000), studies including 2–4 monitoring periods over different seasons have found reliability estimates of 0.29–0.67 and concluded that longer and/or several monitoring periods are needed (Aadland et al., 2017, 2020; Mattocks et al., 2007; Wickel & Welk, 2010). Interestingly, Aadland et al. (Aadland et al., 2020) found that reliability increased marginally from 0.40 for 1 day to 0.51 for 7 days and 0.52 for 14 days of monitoring (mean of the PA intensities) in 3–5-year-old children. Thus, reliability appears to be minimally affected by the number of monitoring days and levels off

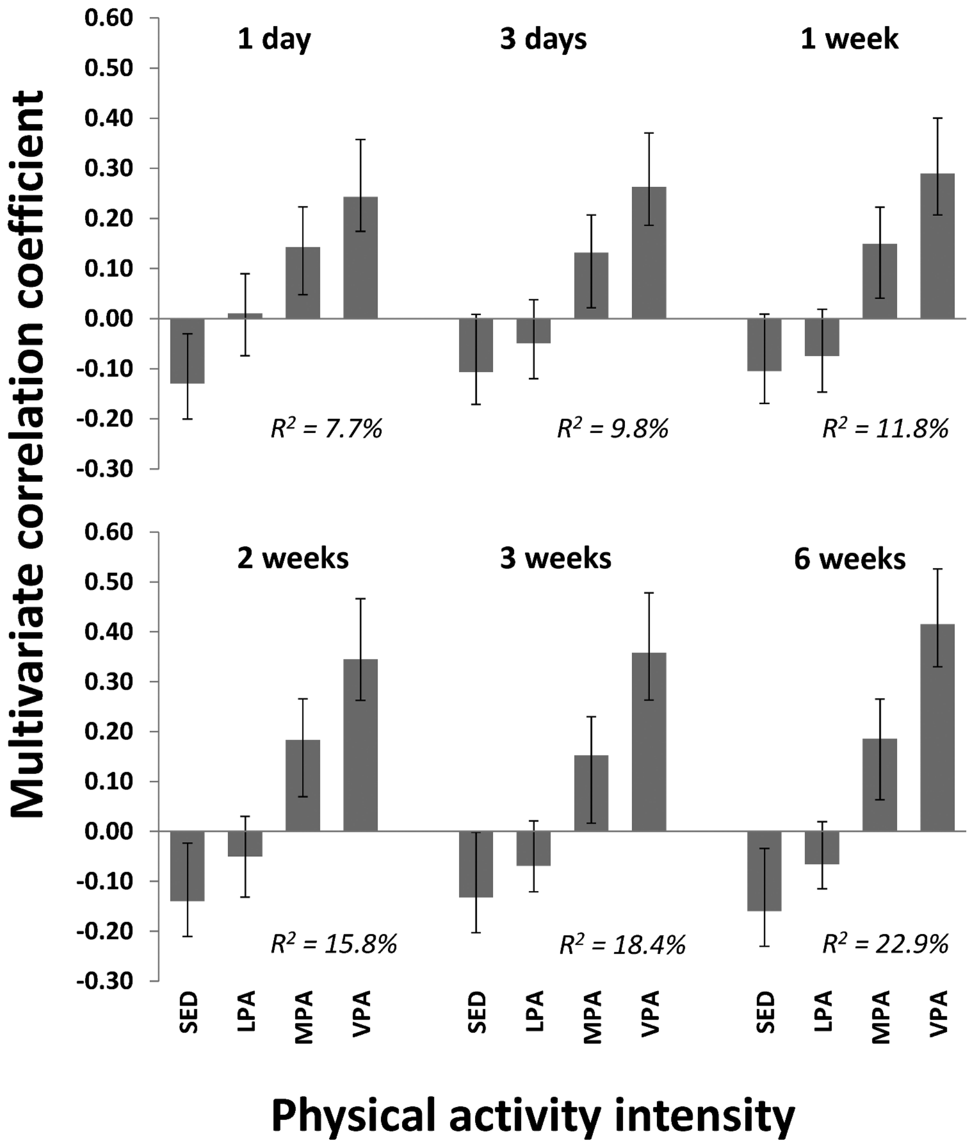

Figure 2. Cross-sectional associations between physical activity and lower body muscle strength across monitoring periods using descriptors derived using different monitoring length. Results are reported as multivariate correlation coefficients from a joint model including the 4 physical activity intensity variables adjusted for sex, age, body mass index and wear time. Multivariate correlation coefficients can be interpreted equivalent to bivariate correlations, although they are derived from the multivariate model. Results for data from 1 day to 2 weeks of monitoring are the means of estimates from each monitoring period modelled separately, while data on 3 and 6 weeks of monitoring are estimates for means across the first and both weeks of monitoring across all monitoring periods, respectively. SED = sedentary time; LPA = light physical activity, MPA = moderate physical activity, VPA = vigorous physical activity. R^2 = explained variance.

after 5–6 monitoring days. This finding contrasts with the finding of improved associations when comparing associations with lower body muscle strength for 1 week (R^2 = 11.8%) and 2 weeks (R^2 = 15.8%) of PA monitoring in the present study. Aadland et al (Aadland et al., 2020) showed that different ways to calculate reliability affected how variance components were estimated This will affect ICCs, which are based on variance partitioning (De Vet et al., 2011; McGraw & Wong, 1996). We also showed in this study that variation for all PA variables decreased when data were averaged over longer periods. However, we have no explanation for these apparently contradictory findings but find it reasonable, consistent with the present findings, that an extended monitoring period is favourable to better capture habitual PA levels that vary over time.

Moreover, consistent with previous studies that have concluded that a single measurement period does not adequately characterize habitual PA (Aadland et al., 2020; Levin et al., 1999; Mattocks et al., 2007; Wickel & Welk, 2010), we found that multiple monitoring periods (R^2 for 3 times 1 week = 18.4% and R^2 for 3 times 2 weeks = 22.9%) led to stronger cross-sectional associations than one monitoring period (R^2 for 1 week = 11.8% and R^2 for 2 weeks = 15.8%). Thus, including multiple separate monitoring periods across seasons seems to improve the precision of capturing habitual PA levels, likely resulting from capturing variability in PA over a longer time-period, including potential differences in activity patterns across seasons. A better capture of this variability leads to less measurement error in characterizing habitual PA and thus stronger associations with an outcome (Hutcheon et al., 2010). Researchers should be aware of the apparent favourable effect of including multiple monitoring periods. Although there are several examples of studies applying extended (Aadland &

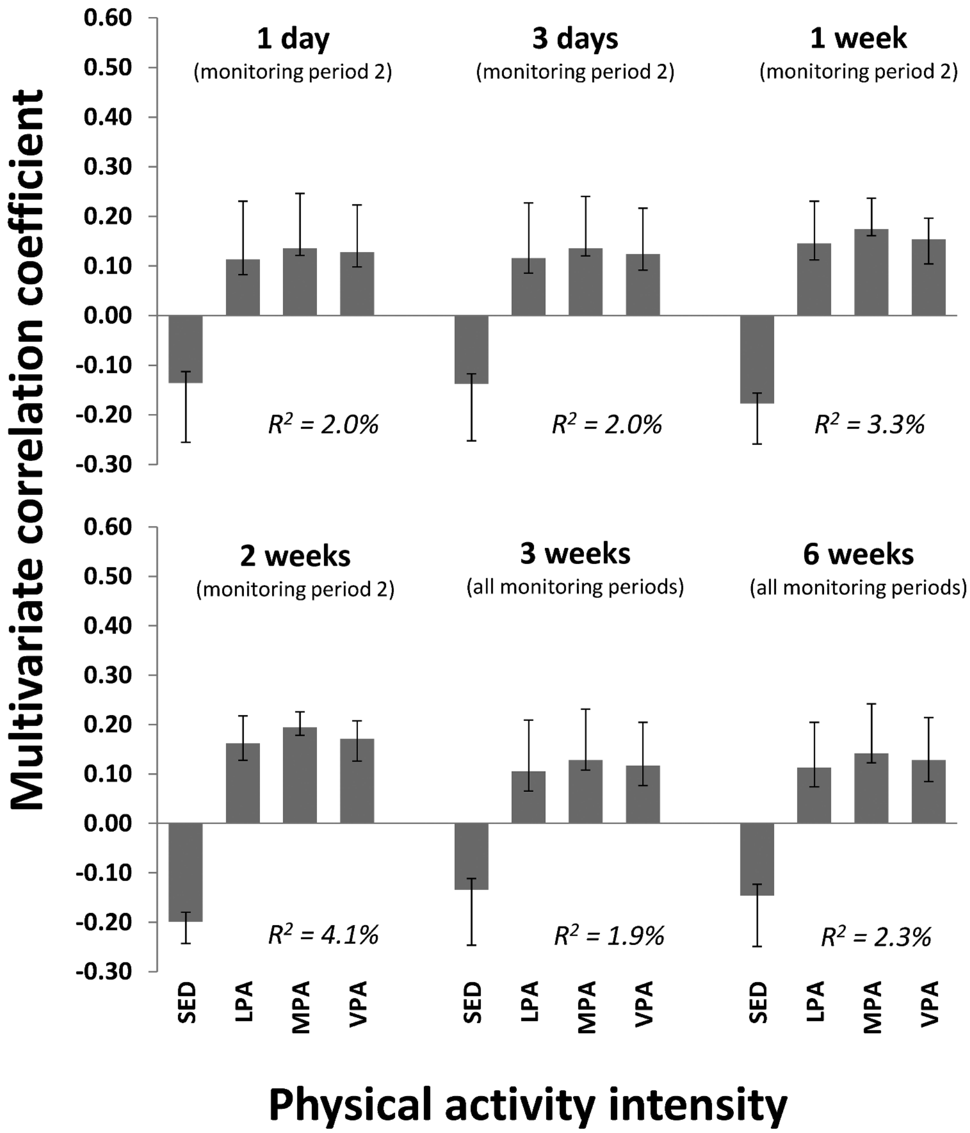

Figure 3. Prospective associations between physical activity and lower body muscle strength for descriptors derived using different monitoring length. Results are reported as multivariate correlation coefficients from a joint model including the 4 physical activity intensity variables adjusted for sex, age, body mass index and wear time. Multivariate correlation coefficients can be interpreted equivalent to bivariate correlations, though they are derived from the multivariate model. Results for data from 1 day to 2 weeks of monitoring are estimates from monitoring period 2 (which was the only monitoring period where we found a significant model), while data on 3 and 6 weeks of monitoring are estimates for means across the first and both weeks of monitoring across all monitoring periods, respectively. SED = sedentary time; LPA = light physical activity, MPA = moderate physical activity, VPA = vigorous physical activity. R^2 = explained variance.

Johannessen, 2015; Aadland & Ylvisåker, 2015; Aadland et al., 2020) and/or multiple (Aadland et al., 2017, 2020; Levin et al., 1999; Matthews et al., 2002; Wickel & Welk, 2010) PA monitoring periods, the improved precision comes at a cost of a substantially increased researcher and participant burden and may not be feasible at a large scale.

The effect of multiple PA monitoring periods appears to be less clear for the longitudinal associations than the cross-sectional associations. The weaker associations for PA data averaged over multiple monitoring periods than for associations for monitoring period 2 is a result of averaging estimates across all monitoring periods where associations for monitoring period 2 were significant and associations for periods 1 and 3 were non-significant. Although the associations for monitoring period 2 are stronger and appear to be "better" than for the associations derived using the average of all monitoring periods, it could be argued that the results for monitoring period 2 are chance findings and that the associations derived when using the average of all monitoring periods are more correct. Thus, our interpretation of our findings is that multiple monitoring periods were important for the validity of both the cross-sectional and the longitudinal associations. Researchers should therefore consider whether PA monitoring protocols extending beyond a single 7-day period might be feasible with regard to both participant and researcher burden.

Strengths and limitations

The main strength of the present study is the inclusion of a relatively large sample of children having 3 separate 14-day PA monitoring periods at baseline and a valid measure of lower body muscle strength at baseline and at 4-year follow-up.

A larger sample would have strengthened the study, but to the best of our knowledge this is the first study to examine how reliability and variability of accelerometer-derived PA data affect associations with an outcome in children. Hopefully, our approach may stimulate further research in this area and inform the development of optimal PA monitoring protocols.

Norway has profound seasonal differences in weather conditions and daylight, which may cause changes in PA levels and types across measurement periods. These characteristics might limit generalizability of our findings compared to areas with less pronounced seasonality. However, our findings on reliability in this sample (Aadland et al., 2020) are consistent with previous studies investigating reliability over multiple seasons in children (Aadland et al., 2017; Mattocks et al., 2007; Wickel & Welk, 2010). Given that reliability of accelerometer-measured PA is similar in children and adults (Aadland & Johannessen, 2015; Aadland & Ylvisåker, 2015; Aadland et al., 2017, 2020; Addy et al., 2014; Baranowski et al., 2008; Basterfield et al., 2011; Bisson et al., 2018; Coleman & Epstein, 1998; Hart et al., 2011; Hinkley et al., 2012; Hislop et al., 2014; Janz et al., 1995; Jerome et al., 2009; Kang et al., 2014; Levin et al., 1999; Matthews et al., 2002, 2012; Mattocks et al., 2007; Murray et al., 2004; Ojiambo et al., 2011; Penpraze et al., 2006; Rich et al., 2013; Treuth et al., 2003; Trost et al., 2000, 2005; Wickel & Welk, 2010), we believe findings generalize to older children and adults. However, future studies should seek to verify our findings in other age groups. Additional limitations are that we did not account for week or weekend days in this study and that children could accrue relatively more of their wear hours during preschool hours or after school (minimum 8 hours per day). However, wear hours per day and requiring inclusion of weekend days have previously been found to marginally affect reliability (Aadland & Johannessen, 2015; Aadland & Ylvisåker, 2015; Aadland et al., 2017).

Conclusion

We conclude that the number of monitoring days and monitoring periods affected associations between PA and lower body muscle strength in young children. A 14-day period led to stronger associations than a 7-day period, and 3 monitoring periods led to stronger associations than 1 monitoring period. Our findings suggest that multiple extended PA monitoring periods may improve the validity of study findings regarding associations between PA and relevant outcomes.

Acknowledgments

We thank all children, parents and staff at the participating preschools for their excellent cooperation during the data collection. We also thank students and colleagues at the Western Norway University of Applied Sciences (formerly Sogn og Fjordane University College) for their assistance during the data collection.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The study was funded by the County Municipality of Sogn og Fjordane and the Western Norway University of Applied Sciences (formerly Sogn og Fjordane University College). None of the funding agencies had any role in the study design, data collection, analyzing or interpreting data or writing the manuscripts.

ORCID

Eivind Aadland http://orcid.org/0000-0001-9654-2653

Authors' contributions

EAA obtained funding for the study. EAA designed the study. AKON performed the data collection. EAA analyzed the data and wrote the manuscript draft. All authors discussed the interpretation of the results and read and approved the final manuscript.

Data availability statement

The datasets used for the current study are available from the corresponding author on reasonable request.

References

Aadland, E., Andersen, L. B., Anderssen, S. A., Resaland, G. K., & Kvalheim, O. M. (2019). Accelerometer epoch setting is decisive for associations between physical activity and metabolic health in children. *Journal of Sports Sciences*, *38*(3), 1–8. https://doi.org/10.1080/02640414.2019.1693320

Aadland, E., Andersen, L. B., Resaland, G. K., & Kvalheim, O. M. (2019). Interpretation of multivariate association patterns between multicollinear physical activity accelerometry data and cardiometabolic health in children—a tutorial. *Metabolites*, *9*(7), 129. https://doi.org/10.3390/metabo9070129

Aadland, E., Andersen, L. B., Skrede, T., Ekelund, U., Anderssen, S. A., Resaland, G. K., & Lucía, A. (2017). Reproducibility of objectively measured physical activity and sedentary time over two seasons in children; Comparing a day-by-day and a week-by-week approach. *PloS One*, *12*(12), e0189304. https://doi.org/10.1371/journal.pone.0189304

Aadland, E., Holmøy, O. K., & Nilsen, A. K. O. (2021). The multivariate physical activity signature associated with body mass index in young children. *Preventive Medicine*, *145*, 106437. https://doi.org/10.1016/j.ypmed.2021.106437

Aadland, E., & Johannessen, K. (2015). Agreement of objectively measured physical activity and sedentary time in preschool children. *Preventive Medicine Reports*, *2*, 635–639. https://doi.org/10.1016/j.pmedr.2015.07.009

Aadland, E., Kvalheim, O. M., Anderssen, S. A., Resaland, G. K., & Andersen, L. B. (2018). The multivariate physical activity signature associated with metabolic health in children. *The International Journal of Behavioral Nutrition and Physical Activity*, *15*(77). https://doi.org/10.1186/s12966-018-0707-z

Aadland, E., Nilsen, A. K. O., Ylvisåker, E., Johannessen, K., & Anderssen, S. A. (2020). Reproducibility of objectively measured physical activity: Reconsideration needed. *Journal of Sports Sciences*, *38*(10), 1132–1139. https://doi.org/10.1080/02640414.2020.1743054

Aadland, E., & Ylvisåker, E. (2015). Reliability of objectively measured sedentary time and physical activity in adults. *PloS One*, *10*(7), 1–13. https://doi.org/10.1371/journal.pone.0133296

Aadland, E., Ylvisåker, E., Johannessen, K., & Nilsen, A. K. O. (2022). Prospective association patterns for the physical activity intensity spectrum with body mass index and lower body muscle strength in

Norwegian children aged 3–9 years. *Journal of Physical Activity & Health*, *20*(1), 20–27. https://doi.org/10.1123/jpah.2022-0157

Addy, C. L., Trilk, J. L., Dowda, M., Byun, W., & Pate, R. R. (2014). Assessing preschool children's physical activity: how many days of accelerometry measurement. *Pediatric Exercise Science*, *26*(1), 103–109. https://doi.org/10.1123/pes.2013-0021

Artero, E. G., Espana-Romero, V., Castro-Pinero, J., Ortega, F. B., Suni, J., Ruiz, J. R., & Castillo-Garzon, M. J. (2011). Reliability of field-based fitness tests in youth. *International Journal of Sports Medicine*, *32*(3), 159–169. https://doi.org/10.1055/s-0030-1268488

Baranowski, T., Masse, L. C., Ragan, B., & Welk, G. (2008). How many days was that? We're still not sure, but we're asking the question better! *Medicine and Science in Sports and Exercise*, *40*(7), S544–9. https://doi.org/10.1249/MSS.0b013e31817c6651

Basterfield, L., Adamson, A. J., Pearce, M. S., & Reilly, J. J. (2011). Stability of habitual physical activity and sedentary behavior monitoring by accelerometry in 6-to 8-year-olds. *Journal of Physical Activity & Health*, *8*(4), 543–547. https://doi.org/10.1123/jpah.8.4.543

Bisson, M., Tremblay, F., Pronovost, E., Julien, A., & Marc, I. (2018). Accelerometry to measure physical activity in toddlers: Determination of wear time requirements for a reliable estimate of physical activity. *Journal of Sport Sciences*, *37*(3), 298–305. https://doi.org/10.1080/02640414.2018.1499391

Cadenas-Sanchez, C., Martinez-Tellez, B., Sanchez Delgado, G., Mora-Gonzalez, J., Castro-Pinero, J., Lof, M., Ruiz, J. R., & Ortega, F. B. (2016). Assessing physical fitness in preschool children: Feasibility, reliability and practical recommendations for the PREFIT battery. *Journal of Science and Medicine in Sport*, *19*(11), 910–915. https://doi.org/10.1016/j.jsams.2016.02.003

Cain, K. L., Sallis, J. F., Conway, T. L., Van Dyck, D., & Calhoon, L. (2013). Using accelerometers in youth physical activity studies: A review of methods. *Journal of Physical Activity & Health*, *10*(3), 437–450. https://doi.org/10.1123/jpah.10.3.437

Cattuzzo, M. T., Henrique, R. D., AHN, R., Oliveira, I. S., Sousa Moura M, M. B., de Sousa Moura, M., de Araújo, R. C., & Stodden, D. (2016). Motor competence and health related physical fitness in youth: A systematic review. *Journal of Science and Medicine in Sport*, *19*(2), 123–129. https://doi.org/10.1016/j.jsams.2014.12.004

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: International survey. *Bmj*, *320*(7244), 1240–1243. https://doi.org/10.1136/bmj.320.7244.1240

Coleman, K. J., & Epstein, L. H. (1998). Application of generalizability theory to measurement of activity in males who are not regularly active: A preliminary report. *Research Quarterly for Exercise and Sport*, *69*(1), 58–63. https://doi.org/10.1080/02701367.1998.10607667

De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge university press.

Esliger, D. W., Copeland, J. L., Barnes, J. D., & Tremblay, M. S. (2005). Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *JPhysAct Health*, *2*(3), 366. https://doi.org/10.1123/jpah.2.3.366

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Hart, T. L., Swartz, A. M., Cashin, S. E., & Strath, S. J. (2011). How many days of monitoring predict physical activity and sedentary behaviour in older adults? *The International Journal of Behavioral Nutrition and Physical Activity*, *8*(1), 8. https://doi.org/10.1186/1479-5868-8-62

Hinkley, T., O'Connell, E., Okely, A. D., Crawford, D., Hesketh, K., & Salmon, J. (2012). Assessing volume of accelerometry data for reliability in preschool children. *Medicine and Science in Sports and Exercise*, *44*(12), 2436–2441. https://doi.org/10.1249/MSS.0b013e3182661478

Hislop, J., Law, J., Rush, R., Grainger, A., Bulley, C., Reilly, J. J., & Mercer, T. (2014). An investigation into the minimum accelerometry wear time for reliable estimates of habitual physical activity and definition of a standard measurement day in pre-school children. *Physiological Measurement*, *35*(11), 2213–2228. https://doi.org/10.1088/0967-3334/35/11/2213

Hutcheon, J. A., Chiolero, A., & Hanley, J. A. (2010). Random measurement error and regression dilution bias. *British Medical Journal*, *340*(jun23 2), 340. https://doi.org/10.1136/bmj.c2289

Janz, K. F., Witt, J., & Mahoney, L. T. (1995). The stability of children's physical-activity as measured by accelerometry and self-report. *Medicine and Science in Sports and Exercise*, *27*(9), 1326–1332. https://doi.org/10.1249/00005768-199509000-00014

Jerome, G. J., Young, D. R., Laferriere, D., Chen, C. H., & Vollmer, W. M. (2009). Reliability of RT3 Accelerometers among Overweight and Obese Adults. *Medicine and Science in Sports and Exercise*, *41*(1), 110–114. https://doi.org/10.1249/MSS.0b013e3181846cd8

John, D., & Freedson, P. (2012). ActiGraph and Actical physical activity monitors: A peek under the hood. *Medicine and Science in Sports and Exercise*, *44*(1 Suppl 1), S86–9. https://doi.org/10.1249/MSS.0b013e3182399f5e

Kang, M., Bjornson, K., Barreira, T. V., Ragan, B. G., & Song, K. (2014). The minimum number of days required to establish reliable physical activity estimates in children aged 2–15 years. *Physiological Measurement*, *35*(11), 2229–2237. https://doi.org/10.1088/0967-3334/35/11/2229

Kvalheim, O. M., Arneberg, R., Grung, B., & Rajalahti, T. (2018). Determination of optimum number of components in partial least squares regression from distributions of the root-mean-squared error obtained by Monte Carlo resampling. *J Chemometrics*, *32*(4), e2993. https://doi.org/10.1002/cem.2993

Kvalheim, O. M., & Karstang, T. V. (1989). Interpretation of latent-variable regression-models. *Chemometrics and Intelligent Laboratory Systems*, *7*(1–2), 39–51. https://doi.org/10.1016/0169-7439(89)80110-8

Leppänen, M., Henriksson, P., Nyström, C. D., Henriksson, H., Ortega, F. B., Pomeroy, J., RUIZ, J. R., CADENAS-SANCHEZ, C., & Löf, M. (2017). Longitudinal physical activity, body composition, and physical fitness in preschoolers. *Medicine & Science in Sports and Exercise*, *49*(10), 49. https://doi.org/10.1249/MSS.0000000000001313

Leppänen, M., Nyström, C. D., Henriksson, P., Pomeroy, J., Ruiz, J., Ortega, F. & Cadenas-Sánchez CLöf, M. (2016). Physical activity intensity, sedentary behavior, body composition and physical fitness in 4-year-old children: Results from the ministop trial. *International Journal of Obesity*, *40*(7), 1126–1133. https://doi.org/10.1038/ijo.2016.54

Levin, S., Jacobs, D. R., Ainsworth, B. E., Richardson, M. T., & Leon, A. S. (1999). Intra-individual variation and estimates of usual physical activity. *Annals of Epidemiology*, *9*(8), 481–488. https://doi.org/10.1016/S1047-2797(99)00022-8

Matthews, C. E., Ainsworth, B. E., Thompson, R. W., & Bassett, D. R. (2002). Sources of variance in daily physical activity levels as measured by an accelerometer. *Medicine and Science in Sports and Exercise*, *34*(8), 1376–1381. https://doi.org/10.1097/00005768-200208000-00021

Matthews, C. E., Hagstromer, M., Pober, D. M., & Bowles, H. R. (2012). Best practices for using physical activity monitors in population-based research. *Medicine and Science in Sports and Exercise*, *44*, S68–76. https://doi.org/10.1249/MSS.0b013e3182399e5b

Mattocks, C., Leary, S., Ness, A., Deere, K., Saunders, J., Kirkby, J., Blair, S. N., Tilling, K., & Riddoch, C. (2007). Intraindividual variation of objectively measured physical activity in children. *Medicine and Science in Sports and Exercise*, *39*(4), 622–629. https://doi.org/10.1249/mss.0b013e318030631b

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Murray, D. M., Catellier, D. J., Hannan, P. J., Treuth, M. S., Stevens, J., & Schmitz, K. H., et al. (2004). School-level intraclass correlation for physical activity in adolescent girls. *Medicine and Science in Sports and Exercise*, *36*(5), 876–882. https://doi.org/10.1249/01.mss.0000126806.72453.1c

Nilsen, A. K. O., Anderssen, S. A., Loftesnes, J. M., Johannessen, K., Ylvisaaker, E., & Aadland, E. (2019). The multivariate physical activity signature associated with fundamental motor skills in preschoolers. *Journal of Sports Sciences*, *38*(3), 1–9. https://doi.org/10.1080/02640414.2019.1694128

Nilsen, A. K. O., Anderssen, S. A., Ylvisåker, E., Johannessen, K., & Aadland, E. Moderate-to-vigorous physical activity among Norwegian preschoolers varies by sex, age, and season. (2019). *Scandinavian Journal of Medicine &*

Science in Sports, *29*(6), 862–873. In press. https://doi.org/10.1111/sms.13405

Ojiambo, R., Cuthill, R., Budd, H., Konstabel, K., Casajus, J. A., Gonzalez-Aguero, A., Anjila, E., Reilly, J. J., Easton, C., & Pitsiladis, Y. P. (2011). Impact of methodological decisions on accelerometer outcome variables in young children. *International Journal of Obesity*, *35*(S1), S98–103. https://doi.org/10.1038/ijo.2011.40

Ortega, F. B., Artero, E. G., Ruiz, J. R., Vicente-Rodriguez, G., Bergman, P., Hagströmer, M., Ottevaere, C., Nagy, E., Konsta, O., Rey-López, J. P., Polito, A., Dietrich, S., Plada, M., Béghin, L., Manios, Y., Sjöström, M., & Castillo, M. J. (2008). Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *International Journal of Obesity*, *32*(5), S49–57. https://doi.org/10.1038/ijo.2008.183

Ortega, F. B., Cadenas-Sanchez, C., Sanchez Delgado, G., Mora-Gonzalez, J., Martinez-Tellez, B., Artero, E. G., Castro-Piñero, J., Labayen, I., Chillón, P., Löf, M., & Ruiz, J. (2015). Systematic review and proposal of a field-based physical fitness-test battery in preschool children: The PREFIT battery. *Sports Medicine*, *45*(4), 533–555. https://doi.org/10.1007/s40279-014-0281-8

Penpraze, V., Reilly, J. J., MacLean, C. M., Montgomery, C., Kelly, L. A., Paton, J. Y., Aitchison , T., & Grant, S. (2006). Monitoring of physical activity in young children: How much is enough? *Pediatric Exercise Science*, *18*(4), 483–491. https://doi.org/10.1123/pes.18.4.483

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. -P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Gorber, S. C., Kho, M. E., Sampson, M., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, *41*(6 Suppl 3), S197–239. https://doi.org/10.1139/apnm-2015-0663

Rajalahti, T., Arneberg, R., Berven, F. S., Myhr, K. M., Ulvik, R. J., & Kvalheim, O. M. (2009). Biomarker discovery in mass spectral profiles by means of selectivity ratio plot. *Chemometrics and Intelligent Laboratory Systems*, *95*(1), 35–48. https://doi.org/10.1016/j.chemolab.2008.08.004

Rajalahti, T., Arneberg, R., Kroksveen, A. C., Berle, M., Myhr, K. M., & Kvalheim, O. M. (2009). Discriminating variable test and selectivity ratio plot: quantitative tools for interpretation and variable (biomarker) selection in complex spectral or chromatographic profiles. *Analytical Chemistry*, *81*(7), 2581–2590. https://doi.org/10.1021/ac802514y

Rajalahti, T., & Kvalheim, O. M. (2011). Multivariate data analysis in pharmaceutics: A tutorial review. *International Journal of Pharmaceutics*, *417*(1–2), 280–290. https://doi.org/10.1016/j.ijpharm.2011.02.019

Rich, C., Geraci, M., Griffiths, L., Sera, F., Dezateux, C., Cortina-Borja, M., & Dasgupta, K. (2013). Quality control methods in accelerometer data processing: Defining minimum wear time. *PloS One*, *8*(6), e67206. https://doi.org/10.1371/journal.pone.0067206

Treuth, M. S., Sherwood, N. E., Butte, N. F., McClanahan, B., Obarzanek, E., Zhou, A., Ayers, C., Adolph, A., Jordan, J., Jacobs, R. JR., & Rochon, J. (2003). Validity and reliability of activity measures in African-American girls for GEMS. *Medicine and Science in Sports and Exercise*, *35*(3), 532–539. https://doi.org/10.1249/01.mss.0000053702.03884.3f

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine and Science in Sports and Exercise*, *43*(7), 1360–1368. https://doi.org/10.1249/MSS.0b013e318206476e

Trost, S. G., McIver, K. L., & Pate, R. R. (2005). Conducting accelerometer-based activity assessments in field-based research. *Medicine and Science in Sports and Exercise*, *37*(11), S531–43. https://doi.org/10.1249/01.mss.0000185657.86065.98

Trost, S. G., Pate, R. R., Freedson, P. S., Sallis, J. F., & Taylor, W. C. (2000). Using objective physical activity measures with youth: How many days of monitoring are needed? *Medicine and Science in Sports and Exercise*, *32*(2), 426–431. https://doi.org/10.1097/00005768-200002000-00025

Utesch, T., Bardid F, Busch, D., Strauss, B., & Bardid, F. (2019). The relationship between motor competence and physical fitness from early childhood to early adulthood: A meta-analysis. *Sports Medicine*, *49*(4), 541–551. https://doi.org/10.1007/s40279-019-01068-y

Veldman, S. L., MJCA, P., & Altenburg, T. M. (2021). Physical activity and prospective associations with indicators of health and development in children aged< 5 years: A systematic review. *The International Journal of Behavioral Nutrition and Physical Activity*, *18*(1), 1–11. https://doi.org/10.1186/s12966-020-01072-w

Wickel, E. E., & Welk, G. J. (2010). Applying generalizability theory to estimate habitual activity levels. *Medicine and Science in Sports and Exercise*, *42*(8), 1528–1534. https://doi.org/10.1249/MSS.0b013e3181d107c4

Wiersma, R., Haverkamp, B. F., van Beek, J. H., Riemersma, A. M. J., Boezen, H. M., Smidt, N., Corpeleijn, E., & Hartman, E. (2020). Unravelling the association between accelerometer-derived physical activity and adiposity among preschool children: A systematic review and meta-analyses. *Obesity Reviews*, *21*(2), 15. https://doi.org/10.1111/obr.12936

Wold, S., Ruhe, A., Wold, H., & Dunn, W. J. (1984). THE COLLINEARITY PROBLEM in LINEAR-REGRESSION - the PARTIAL LEAST-SQUARES (PLS) APPROACH to GENERALIZED INVERSES. *Siam Journal on Scientific and Statistical Computing*, *5*(3), 735–743. https://doi.org/10.1137/0905052