Using Accelerometers in Youth Physical Activity Studies: A Review of Methods

Kelli L. Cain[1], **James F. Sallis**[1], **Terry L. Conway**[1], **Delfien Van Dyck**[3,4], and **Lynn Calhoon**[2]

[1]Department of Psychology, San Diego State University, San Diego, CA

[2]Graduate School of Public Health, San Diego State University, San Diego, CA

[3]Department of Movement and Sports Sciences, Ghent University, Belgium,

[4]Research Foundation Flanders, Brussels, Belgium

Abstract

BACKGROUND: In 2005, investigators convened by the National Cancer Institute recommended development of standardized protocols for accelerometer use and reporting decision rules in articles. A literature review was conducted to document accelerometer methods and decision rule reporting in youth physical activity articles from 2005–2010.

METHODS: Nine electronic databases identified 273 articles that measured physical activity and/or sedentary behavior using the most-used brand of accelerometer [Actigraph]. Six key methods were summarized by age group (preschool, children, and adolescents) and trends over time were examined.

RESULTS: Studies using accelerometers more than doubled from 2005 −2010. Methods included 2 ActiGraph models, 6 epoch lengths, 6 non-wear definitions, 13 valid day definitions, 8 minimum wearing day thresholds, 12 moderate-intensity physical activity cut points, and 11 sedentary cut points. Child studies showed the most variation in methods and a trend toward more variability in cut points over time. Decision rule reporting improved, but only 54% of papers reported on all methods.

CONCLUSION: The increasing diversity of methods used to process and score accelerometer data for youth precludes comparison of results across studies. Decision rule reporting is inconsistent, and trends indicate declining standardization of methods. A methodological research agenda and consensus process are proposed.

INTRODUCTION

Because of the worldwide increase in overweight and obesity in youth, there is a growing interest in the promotion of physical activity in these age groups. However, to draw meaningful conclusions about physical activity prevalence and outcomes of interventions, valid, reliable, and feasible measures are needed. Parent or self-report measures of physical activity are commonly used and can provide useful information about the context of physical

Corresponding author, Kelli L. Cain, San Diego State University (currently at University California San Diego), 3900 5th Avenue, Suite 310, San Diego, CA 92103, Phone: 619-260-5541, Fax: 619-260-1510, kcain@projects.sdsu.edu (currently kcain@ucsd.edu).

activity, but these methods are limited by biased reporting and low validity.[1,2] Consequently, objective methods, especially accelerometers, to assess physical activity in children and adolescents are becoming increasingly used. Several reviews concluded that accelerometers provide an accurate, reliable, and practical objective measure of physical activity in children and adolescents.[3–6] The feasibility of accelerometer use in large studies is illustrated by their use in a nationally representative sample of the United States population, aged 6 years and above, in the National Health and Nutrition Examination Survey (NHANES).[7]

Nonetheless, there are recognized methodological challenges related to accelerometer use and they were outlined in a special issue of *Medicine and Science in Sports and Exercise* (2005), 'Gaps and Challenges in the Science of Accelerometry'. Ward and colleagues[8] made recommendations for future research and best practices for accelerometer use. An important recommendation was for researchers to state their methodologic 'decision rules' (i.e. choice of epoch length for recording data, criteria for valid days and number of wearing days, etc.) in articles. A clear description of these decision rules can allow comparison among studies and facilitate the analysis process. Moreover, Ward et al.[8] called for development of common practices and guidelines when using and scoring accelerometers with children and adolescents. Other researchers like Rowlands[6] and Trost and colleagues[9] also emphasized the need for standardized protocols for accelerometer use.

The recommendations published by Ward et al.[8] in 2005 formed the basis of the present review, with the goal of documenting the evolution of accelerometer methods in children and adolescent studies from 2005–2010. As a first aim, we examined which methods (accelerometer model, epoch length, non-wear time, valid days, minimum number of valid wear days, and cut points) were reported and whether there was a trend for researchers to apply more standardized accelerometer protocols. Secondly, we investigated whether there has been an increase in reporting the specific 'decision rules' in accelerometer studies since 2005. To achieve these aims, we conducted a detailed review of the accelerometer methods in published physical activity studies in preschoolers (<6 years), children (6–11 years), and adolescents (12–18 years).

METHOD

The present review focused on six key methodological issues related to accelerometer data collection and scoring: accelerometer model, epoch length, definition of non-wear time within a day, definition of a valid day, required number of valid wear days, and the choice of cut points for various intensities of physical activity and sedentary behavior.

Accelerometer model

Numerous accelerometer models are commercially available, but a systematic review concluded that the ActiGraph (Pensacola, FL) had the largest body of evidence to support its use[10] and was the most commonly used accelerometer in physical activity research. Therefore, the review was limited to studies using the ActiGraph. In 2005, a new model of ActiGraph (the GT1M) was developed and its comparability with the older model (7164) is still being examined. Some studies have shown comparability between models in moderate-to-vigorous physical activity (MVPA),[11–13] while others have shown differences in overall

counts and on the lower end of the spectrum (i.e., sedentary and light activity).[11] The newer ActiGraph requires a higher acceleration to record a nonzero count than the older 7164 models[14] so the comparability of data across ActiGraph models remains uncertain.

Epoch length

The epoch is the amount of time over which activity counts are integrated and recorded, and it affects estimates of moderate-to-vigorous physical activity in children.[15–17] Because of the limited memory capacities of earlier accelerometer models, a 1-minute epoch has been the most commonly used, but shorter epochs appear to more accurately record the intermittent, short bursts of physical activity common in young people.[5,9]

Processing data

Accelerometers produce large data sets that have grown more complex with the introduction of 3-axis models. Although software programs for screening and scoring accelerometer data are now available (MeterPlus (www.meterplussoftware.com), NHANES SAS program (http://riskfactor.cancer.gov/tools/nhanes_pam), Kinesoft (www.kinesoft.org)), the specific procedures for processing data remain important concerns and affect outcomes. For example, one issue is how researchers define overall 'wear time,' which is important because non-compliance with data collection protocols is common. There is a sequence of decisions that define 'wear time': (a) definition of non-wear time within a day, (b) definition of a valid day, and (c) number of valid wear days for a record to be considered sufficiently reliable and representative of typical behavior. These data processing decisions have the potential to affect sample sizes and physical activity estimates.[18]

Non-wear time within a day.—Non-wear time is typically identified by summing the number of consecutive '0' counts per minute (cpm), and 'wear time' is then calculated by subtracting non-wear time from total possible time. However, it is important to differentiate between '0' counts that are due to: 1) wearing but being sedentary, which can result in '0' counts, and 2) true not wearing, either per instructions (sleeping, swimming, etc.) or due to poor compliance with the wearing protocol. To complicate this issue, sporadic non-zero counts can occur within strings of '0' counts due either to small movements by the person wearing the accelerometer or by incidental movement during non-wear periods (e.g., jostled while on the nightstand). Some non-wear algorithms allow for a small number of non-zero counts during non-wear time (e.g., up to 2 consecutive minutes <100 cpm).[7]

Definition of a valid day.—A valid day is typically defined by a minimum number of wearing hours but can be determined in other ways such as calculating study-specific ratios (e.g., 70/80 rule where valid is 80% of a time period defined by 70% of the sample having data)[19] or applying a minimum and maximum counts per day threshold (e.g., <10,000 and >20,000,000 excluded).[20] The definition might be different for weekdays and weekends and may also differ by age. Ward and colleagues[8] suggested that adolescents need more daily wear time to be representative of usual activity than children, for example. The more wear time required per day, the greater the reduction in valid days and sample sizes.[21]

Minimum number of valid wear days.—The final decision in defining overall wear time is the number of valid wearing days required to represent 'usual' activity. Recommendations typically range from 3 to 10 days[3,22] and reliability estimates differ by age and sex.[23–24] Ward and colleagues[8] suggested that 7 days might be a reasonable standard for all ages. It is recommended that at least one weekend day is required to obtain reliable estimates.[6,9,25] The number of wear days required also impacts sample sizes. Some studies accept fewer days and use imputation or probability methods to maximize sample sizes,[7,19] while other studies apply a more stringent requirement for wear days and place a greater emphasis on boosting compliance rates.[26–27]

Cut points

Moderate-to-vigorous physical activity (MVPA): A proliferation of published thresholds or cut points to translate accelerometer counts into minutes of physical activity has produced an ongoing debate about which cut points to use in research. The choice of MVPA cut points has a tremendous impact on study comparability, and numerous studies have shown large differences in physical activity prevalence depending on the choice of cut points.[3,5,28–34] Cut points can vary based on the results of studies that have "calibrated" accelerometers by defining counts per minute in relation to physical activity intensity. Calibration studies have been conducted in laboratories and field settings and have used a wide variety of methods.[35] At a 4 MET intensity (i.e., 4 times resting energy expenditure), recommended counts per minute have ranged from 1400 to 3600.[25,36–37] Another source of variation is the selection of MET definitions of MVPA, with common definitions ranging from 3 to 4.6 METs.[25,36,38] To add complexity, some calibration studies recommended age-specific cut points,[25,36,39–40] and others used the same cut points for all ages. [34–35,37–38,41–42]

Sedentary behavior: Sedentary behavior has been identified as a chronic disease risk factor that is independent of physical activity.[43] Low counts per minute from accelerometry can be used to measure sedentary time.[38,41–42] Similar to the issues with MVPA cut points, numerous cut points have been created to classify sedentary behavior. Two studies have shown that the choice of a sedentary cut point can change estimates by more than 4 hours per day in preschoolers.[5,34]

Review Methods

An electronic search of studies published in 2005–2010 was conducted through EBSCOhost and PubMed using the following databases: MEDLINE, Academic Search Premier, CINAHL Plus with Full Text, ERIC, MAS Ultra - School Edition, PsycARTICLES, PsycEXTRA, PsycINFO and SPORTDiscus. The literature search used any of the following keywords to capture accelerometer data collection: 'accelerometer', 'accelerometers', 'accelerometry', 'CSA', 'activity meter', 'MTI', 'actigraph', OR 'activity monitor' in combination with (using the Boolean term AND) any of the following keywords to represent physical activity: 'physical activity', 'MVPA', 'intensity', 'threshold', sedentary', 'moderate', 'vigorous', 'cut points', 'cutpoints', 'cut-points', OR 'METs'.

Inclusion criteria included a) studies reporting accelerometer-based physical activity or sedentary outcomes, b) studies of youth aged 0–18 years, and c) full text English language publications. Exclusion criteria included a) studies not using the ActiGraph accelerometer (Pensacola, FL; www.theactigraph.com), b) studies of adults >18 yrs of age, c) studies not reporting ambulatory movement (e.g., sleep studies), d) methodological papers, review papers, editorials, and comments, and e) studies with less than 1 complete day of accelerometer monitoring (e.g., school recess studies).

The initial literature search yielded 3020 articles. After applying an age criterion of 0–18 years and removing 6 duplicate references, 573 articles remained. These abstracts were reviewed for inclusion before obtaining the full text articles. Exclusion details are outlined in Figure 1. After excluding 300 articles (264 based on reviewing the abstract and 36 after a complete review of the article), 273 were deemed suitable for this literature review.

Articles were abstracted by 2 reviewers and discrepancies were resolved by a third reviewer. The following information was extracted from each paper: study characteristics (e.g., authors, year, study name, study aims, age), sample size, ActiGraph model, epoch length, wear time definitions (i.e., non-wear time within a day, number of valid hours, minimum number of wear days required, number of weekend days required), cut point for moderate physical activity, cut point for sedentary behavior, and other relevant methods. If any of the above information was missing from the articles, the variable was coded as 'missing'. When authors referenced other papers for accelerometer methods, the referenced articles were reviewed and if the methods were not reported, the variable was coded as 'missing'. When methods were described in an unclear way, they were also coded as 'missing'. In some cases, the reference did not match reported cut point values. These were flagged for incorrect reference but the reported value was retained (not coded as 'missing').

Results were summarized for the whole age range and according to age group: preschoolers (< 6 yrs or specific reference to "preschool"), children (6–11 yrs or specific reference to elementary or primary school), and adolescents (12–18 yrs or specific reference to secondary, middle, or high school). When the age range spanned categories, mean age or mid-point age (if mean was not available) was used to determine the age category. Papers were sorted by study name and when more than one paper from the same study was discovered, the methods were compared. The studies describing identical methods or using the same reference for methods were grouped together as one 'study' and included only once in the summary table (Table 1). Incomplete methods were filled in from same-study papers where feasible to obtain the most complete report possible. The most recent reference within the same-study papers was used for the year of publication. Studies that published papers at different stages of the study (e.g., TAAG: Trial for Activity in Adolescent Girls), or analyzed data from different cohorts (e.g., EYHS: European Youth Heart Study), were grouped together by participant characteristics. For example, within the TAAG set of papers, there were 3 'studies' (included 3 times in the summary table): 6th graders, 8th graders, and 6th & 8th graders combined. Groupings were also made within studies when different cut points were used. For example, within the EYHS set of papers, there were 6 'studies': 9 yr olds (Freedson cut points),[25,36] 9 yr olds (EYHS cut points),[44–45] 9 & 15 yr olds (Freedson

cut points), 9 & 15 yr olds (Treuth cut points),[38] 9 & 15 yr olds (EYHS cut points), and 9 & 15 yr olds (counts only).

Papers were coded by publication year and divided into 3 groups: 2005–2006, 2007–2008, and 2009–2010. Because the selection of cut points has created the most controversy in the field and has consistently been shown to impact the comparability of results, we plotted the distribution of the most commonly used thresholds by year to examine whether researchers' choices have become more homogeneous over time (Figure 2 and Figure 3). The frequencies of 'missing' methods were also plotted to examine whether reporting on the key issues improved since 2005 (Figure 4). In addition to examining each method separately, a 'reporting' score was calculated for each article; number of methods reported divided by the total number of relevant methods, out of 6 possible. Reporting scores were plotted by publication year (Figure 5). Methods not relevant to an article were not included in the denominator (e.g., an article reporting only activity counts would not have a 'cut points' variable in the denominator).

RESULTS

The results of the literature review are presented in the Review Table (see Supplemental Appendix for Review Table and references). A total of 273 articles (25 preschool, 140 child and 108 adolescent) were included in this review. The distribution of articles by year from 2005–2010 was 27, 37, 46, 55, 62 and 46, respectively.

After grouping papers from the same studies (as described above), it was determined that the 273 identified articles actually represented 183 studies (21 preschool, 94 child, and 68 adolescent), which were used as the basis for the summary of the key methodological issues (Table 1). The key findings are outlined below.

Year of publication

Thirty-eight of the studies in Table 1 were published in 2005–2006, 60 in 2007–2008, and 85 in 2009–2010. While increases in accelerometer studies were seen in all age groups, the largest increase was seen in child studies with more than a threefold increase from 14 to 47 from 2005–2006 to 2009–2010, respectively.

Accelerometer model

Since 2005, two ActiGraph models were used: the older 7164 model (previously referred to as the 'CSA'), and the newer GT1M. The majority of studies across all age groups used the older models. The GT1M first appeared in articles in 2007, with about 7% of studies reporting using it. This increased to 36% of studies in 2010. About 23% of child studies used the GT1M, while only 9.5% and 11.8% of preschool and adolescent studies, respectively, used this newer model. It is notable that about 20% of preschool and adolescent studies did not report which ActiGraph model was used.

Epoch length

Six different epochs were used, ranging from 2s to 60s (seconds). The majority of the studies in children and adolescents used a 60s epoch (66.0% and 63.2% respectively). There was a trend toward using shorter epochs in preschoolers. In preschool studies, an equal number used 15s and 60s (42.9% each). In child studies, the second most common epoch was 10s (9.6%), and in adolescent studies the second most common was 30s (13.2%). Across the child studies, 6 epochs were used, while preschool and adolescent studies used 4 epochs. Most of the studies not reporting epoch length refer to analyzing 'counts per minute' so it could be assumed they used a 60 second epoch. Adolescent studies were the most likely <u>not</u> to report epoch length (16.2%).

Non-wear time within a day

Six definitions of 'non-wear time' were reported, ranging from 10–180 minutes of consecutive zero counts. Three definitions (10, 20 and 60 minutes) were used in preschool studies with similar occurrence. In child and adolescent studies, 10- and 20-minute strings of zeros were the most common. Child studies showed the most variation in determining non-wear time (5 definitions), and adolescent studies used 4 definitions. Non-wear time within a day was the most likely method <u>not</u> to be reported, with 66.7%, 48.9%, and 42.6% of preschool, children, and adolescent studies, respectively, not reporting this.

Definition of a valid day

Determination of a 'valid day' was highly variable, with 13 definitions used. Most studies (110 of 183) used a number of valid hours but a few (16 of 183) used counts or ratios of time (e.g., 80% of average wearing time) to define a valid day. Preschool studies used 7 definitions with 6 and 10 valid hours being the most common (14.3% each). Child studies showed the most variability with 12 different methods reported and a range from <6 to 12 valid hours. About 8% of child studies used a different definition for weekdays and weekends, and using ratios or counts criteria was more common in child studies (about 6% each) than in other age groups. Eight and 10 valid hours were the most common (10.6% and 21.3% respectively). Adolescent studies showed the least variability with 5 definitions. The range of valid hours was 8 to 12, with about 6% using a ratio or counts criteria. The most commonly used were 8 and 10 valid hours (13.2% and 44.1% respectively). About 24% of preschool and child studies, and 27% of adolescent studies, did <u>not</u> report a 'valid day' definition.

Minimum number of wear days

Eight thresholds were used to define the 'minimum number of wear days' for inclusion in analyses. Three and 4 days were the most common in all age groups, used in 56% preschool studies, 68% child studies, and 70% adolescent studies. Preschool studies used 6 thresholds, child studies used all 8 thresholds, and adolescent studies used 5 thresholds. The minimum number of wearing days was not reported in 23.8% of preschool, 17.0% of children, and 17.6% of adolescent studies. As with other methods, child studies showed the most variability with at least one study using every definition between 1 and 10 or more days.

Weekends: The majority of studies in all age groups did not require a weekend day for inclusion in analyses (61.9% preschool, 51.1% child, and 66.2% of adolescent studies), but 30.9% of child studies required one weekend day compared to about 14% of preschool and adolescent studies. Very few studies (<3%) required 2 weekend days.

Cut points

Moderate-to-Vigorous Physical Activity (MVPA): The MVPA cut point showed the most variability, with 12 sets of cut points producing 14 values that ranged from 400 to 3600 cpm (lower threshold for moderate-intensity) across all age groups. In preschoolers, 5 cut points were used that ranged from 614–3200 cpm. The Sirard age-specific cut points (2460–3564 cpm; 38.1% of studies)[39] were the most common, followed by Pate (1680 cpm; 19%), [46] and Puyau (3200 cpm; 19%).[42] These 3 most commonly used cut points accounted for 76.1% of preschool studies. Child studies showed the most variability with 11 cut points that ranged from 400 to 3600 cpm. The 4 most common cut points that accounted for a majority of studies were Freedson youth (614–3500 cpm -- all MET definitions combined; 32.6% of studies),[25,36] Puyau (3200 cpm; 13%), EYHS (2000 cpm; 10.9%)[44–45] and Treuth (1160 & 3000 cpm -- all MET definitions combined; 9.8%).[38] Nine percent of child studies did not use cut points but used 'counts' instead, and 8% either reported results based on more than one cut point or made up their own. In adolescent studies, 8 cut points were used and, as with child studies, the Freedson youth cut points were most common (614–3500 cpm -- all MET definitions combined; 38.5% of studies), followed by Puyau (3200 cpm) and EYHS (2000 cpm) at 10.8% each, and Treuth (1160 and 3000 cpm -- all MET definitions combined; 9.2%). MVPA cut points were not reported in 3.3% of children and 6.2% of adolescent studies. All preschool studies reported the MVPA cut point.

Sedentary behavior: Sedentary behavior was measured in 95% of preschool, 40.4% of child, and 47.1% of adolescent studies included in this review. Across all age groups, 11 sedentary cut points were reported ranging from 9 to 1259 cpm (upper limits). In preschoolers, 4 sedentary cut points were used with the most common being Sirard (1204–1259 cpm; 35% of studies),[39] Reilly (1099 cpm; 30%),[47] and Pate (151 cpm; 15%).[46] Child studies used 9 sedentary cut points with the most common being Treuth/Evenson (100 cpm; 26.3% of studies),[38,41] and Puyau (799 cpm)[42] and EYHS (499 cpm)[44–45] at 21.1% each. Five sedentary cut points were used in adolescent studies with the most common being Treuth/Evenson (100 cpm; 46.9% of studies), Puyau (799 cpm; 25%) and EYHS (499 cpm; 12.5%). Similar to other methods, there was more variability in child studies than in preschool or adolescent studies. Sedentary cut points were the least likely <u>not</u> to be reported (5% of preschool, 2.6% of child, and 3.1% of adolescent studies).

Trends from 2005–2010

MVPA cut points: MVPA cut points were tracked by year to identify any trends. The graphed results for children and adolescents are shown in Figures 2 and 3, respectively. The trends in preschool cut points were not graphed because there have been no changes since 2008 when the Sirard[39] and Pate[46] values replaced the Freedson youth,[25,36] Freedson adult, [48] and Puyau[42] values as the only preschool cut points in use.

In child studies, there has been an increase in the number of cut points used, from 5 in 2005–2006, to 11 in 2009–2010. With the introduction of new cut points, the percent of studies using the Freedson youth,[25,36] Freedson adult,[48] and Puyau[42] cut points decreased, while the percent using Treuth,[38] EYHS,[44–45] and Mattocks[37] increased. However, in 2009–2010, the Freedson youth cut points still were the most commonly used (28.9% of studies), followed by the 'other' category (17.8%), and the EYHS cut points (15.6%).

In adolescent studies, there has not been a clear increase in the number of cut points used since 2005, but the distributions changed as new cut points were introduced and others became less common. The percent of studies using the Freedson adult[48] and Puyau[42] decreased while the percent using Treuth[38] and Freedson youth[25,36] increased. The Freedson youth cut points were the most commonly used in 2009–2010 (41.4% of studies), followed by Treuth (20.7%), and the EYHS cut points (17.2%).

Methods Reporting: Figure 4 displays the percent of papers that did NOT report each key variable (i.e., 'missing' information). Four variables (accelerometer model, non-wear definition, valid hours, and cut points) showed improvements in reporting, identified by a decreasing percent of non-reporting, while epoch length showed an increase in non-reporting and 'valid days' stayed about the same. The non-wear definition (i.e., minutes of consecutive zero counts) remained the most likely variable not to be reported, but it showed the sharpest decline in non-reporting from 65.6% in 2005–2006 to 30.6% in 2009–2010. Non-reporting of cut points declined from 10.9% to 4.6% making cut points the least likely to be missing in 2009–2010. There was an increase in studies not reporting on epoch length, making epoch length among the most common variables not to be reported in recent years along with valid hours and minimum number of days (about 13% each).

Figure 5 displays the percent of articles with a 'complete reporting' score, defined as reporting on 100% of relevant methods, and the percent of articles with a 'poor reporting' score, defined as reporting on 50% or fewer methods. In 2005–2006, 21.9% of papers reported on all methods, which more than doubled to 53.7% in 2009–2010. 'Poor reporting' declined from 18.8% in 2005–2006 to 11.1% in 2009–2010. The mean "percent of key methods reported" scores increased from 74.6% to 85.3% between 2005–2006 and 2009–2010, respectively.

DISCUSSION

Objectively measured physical activity using accelerometers provides many advantages over self-reports, particularly among youth,[2] but there are numerous challenges in applying the technology. As accelerometers are used more in research and in national prevalence studies that provide a basis for public health planning,[7] it becomes more important to derive results that justify high confidence. The results of the present review of 183 studies indicated a continuing lack of consensus among researchers on the methods used to collect, process, and score accelerometer data from children and adolescents.

In 2005, investigators convened by the National Cancer Institute recommended the development of more standardized methods for accelerometer studies and consistent

reporting of "decision rules".[8] Although there has been a large increase in the number of studies using accelerometers with children and adolescents from 2005 to 2010, methods have become more diverse, so results are less comparable. The variability in methods was greater for children than for preschoolers or adolescents. Though reporting of decision rules has improved, in 2009–2010 slightly more than half of published studies reported all major methodological decision rules.

The present review was limited to ActiGraph accelerometers, which are used most commonly. Two *models of ActiGraphs* were used, and though there is growing evidence among adults of differences in sensitivity,[13–14] it is unclear how model differences affect interpretation of data from children. Many studies did not report the model, and it is not clear how data from the two models should be interpreted differently.

Technological advances in accelerometry have allowed researchers to collect even richer data and address previous device limitations, such as limited memory capacity, but these advances may introduce additional challenges. New models of ActiGraphs allow for the collection of data of multiple axes and store raw data with the epoch being defined post-data collection. These new capabilities create more powerful tools, but they produce more complex data for which methodological guidelines are lacking. Thus, advancing technology creates more opportunities for methodological discrepancies and difficulties interpreting results.

The *length of epochs* at which accelerometer data are summarized is particularly important for children, whose activity levels change frequently. Estimated time in vigorous physical activity is inversely related to epoch length, resulting in an underestimation when longer epochs are used.[53] Unfortunately, epoch length was one of the most common not-reported decision rules, and non-reporting increased over time. As the technology has allowed, there was a strong trend toward using epochs shorter than 60s, increasing from 0% of studies in 2005 to 45% in 2010. However, it has been suggested that physical activity outcomes using shorter (5, 15, and 30 seconds) vs longer (60 second) epochs are not comparable in some age groups.[54]

Different decision rules for defining *non-wearing time* further reduce comparability across studies, affecting number of valid days, sample sizes, and particularly, estimates of sedentary time.[55] Colley et al.[21] found the percentage of their sample having at least 4 valid days of data ranged from 38% to 84% when comparing non-wear definitions of 10 minutes of consecutive zero counts to 60 minutes. In this review, criteria ranged from 10 to 180 minutes and the majority of studies used 10 or 20 minutes. Esliger[56] reported that 20 minutes of zeroes was most appropriate for youth and though it would seem unlikely for a child to be absolutely still (no movement recorded) for more than 10 or 20 consecutive minutes, there are few data to inform a criterion for non-wearing time for children. Non-wear times were the least reported decision rules in studies, though there was substantial improvement in reporting over time.

Criterion for defining a *valid day* of accelerometry ranged from 6 to 12 hours of wear time. A minimal daily wearing time of 10 hours, which was seen in 28% of studies, has been

recommended for children and adolescents.[57–58] Some researchers are developing their own definitions as the number in use increased from 6 in 2005 to 12 in 2010; thus, results are becoming less comparable. One quarter of studies did not report on this decision rule. A best practice recommendation is to define valid hours before data collection and screen data immediately for compliance so there is the possibility to ask the participant to re-wear if needed.

Criteria used for defining number of sufficient *days of wearing time* directly affect sample size and reliability of estimates of habitual physical activity.[3] Criteria ranged from 1 to 10 or more days with a fair number (18%) of studies not reporting this decision rule and no improvements seen over time. There appears to be a rough consensus for a 7-day monitoring protocol to achieve a minimum of 4 valid days of monitoring including one weekend day. [3,8–9] However, many researchers have not adopted this guideline. Some researchers use imputational or probability methods to deal with missing data,[7,19] and others focus on boosting compliance via phone/texting, incentives, visual reminders (e.g., stickers), wear time logs, and asking participants to re-wear the accelerometer to meet the criterion.[9,59–60]

The diversity of *cut points* used to define MVPA may have the most severe consequences for clear interpretation of findings and ability to compare across studies. The number of cut points used has increased since 2005. Trost refers to the plethora of choices as "the cut point conundrum".[61] Wide variation in calibration study methods has led to cut points that range from 9 to 1259 cpm (as the upper threshold) for sedentary behavior and 400 to 3600 cpm (as the lower threshold) for moderate intensity activity. To highlight the potential for confusion, the most commonly used set of cut points has a lower MVPA value (705 cpm for 7 year olds)[25,36] than one of the most commonly used sedentary cut point in children's studies (799 cpm).[42] The choice of cut points can drastically affect the results,[5,28–29,31–34] with physical activity prevalence rates ranging from 100% to 8.7% in one study depending on whether Freedson youth or Puyau cut points were used.[30]

One of the drivers of the wide range of cut points used is the lack of consensus as to what MET value defines MVPA in youth. The US physical activity guidelines define the lower limit of moderate physical activity as 3 METs for all ages,[62] but an influential calibration study conducted for the TAAG study identified 4.6 METs as the limit,[38] and NHANES adopted 4.0 METs for national prevalence estimates.[7] In analyses of TAAG data, the prevalence of meeting guidelines was 88% using 3 METs for adolescents, compared to only <1% using 4.6 METs.[33] The lack of consensus in methods is preventing acceptable interpretation of accelerometer results and undermining the value of using objective measures.

A Methodological Research Agenda

The good news of increasing use of accelerometers for objectively measuring physical activity in youth is counterbalanced by the bad news of the chaos in methodology that is getting worse over time. Though nurturing creativity among investigators searching for effective methods is important, we propose that it is now time to develop a consensus on protocols for collecting, processing, scoring, and reporting accelerometer data for children and adolescents. The current situation is producing data that can neither be interpreted

confidently nor compared across studies, and this is unacceptable given the public health importance of youth physical activity. The recent introduction of a new generation of accelerometers providing raw data makes the need for consensus protocols even more urgent. Current data are not a sufficient basis for selecting the most promising protocols, because few studies have compared multiple approaches to the key decision rules. Thus, a research agenda devoted to developing an evidence base on accelerometer methodology decision rules is needed. There are some recent examples of studies working toward this goal.

Trost and colleagues[63] conducted a validation study in 206 5–15 year olds comparing 5 common cut points using VO_2 as the criterion measure during 12 activities. They found that cut point values 3000 cpm[37–38] defining MVPA 4 METs were not supported. Using receiver operator characteristic (ROC) curve analysis, the Evenson cut points[41] showed the best classification at all intensities (e.g., 2296 cpm for moderate physical activity), followed by Freedson 4 METs.[25,36] Similar to what Treuth[38] and Evenson[41] found, the 100 cpm sedentary cut point showed excellent classification accuracy. In contrast, the Puyau[42] cut point for sedentary time (<800 cpm) showed fair classification accuracy. The Trost et al.[63] study seems to be a strong basis for selecting cut points, but it was not designed to resolve controversies regarding MET values and other decision rules.

We propose a research agenda and process to establish at least an interim consensus on accelerometer methodology and reporting standards for studies of youth. This process should be led by the National Institutes of Health or Institute of Medicine, with guidance from a panel of measurement experts. These are the critical features of a research agenda and consensus process.

- The most critical need is to reach consensus on MET values for MVPA, so that definitive youth physical activity prevalence rates can be established. Relating physical activity levels, based on various MET values, to disease risk biomarkers is one evidence-based approach to selecting MET definitions for MVPA.

- Since the trend is toward using shorter epoch lengths, studies should be conducted to determine MET-based cut points using very short epoch lengths (e.g., 1 sec) that provide data that are comparable to longer epoch lengths.

- To identify the best approach to selecting accelerometer cut points, an expert panel should review methods used to establish existing cut points, then design a definitive study on which to base final recommendations. A definitive calibration study should include youth from a wide age range to determine whether age-specific cut points are needed. This is an active controversy in the field.[5,63] Calibration studies should be conducted using various epoch lengths

- A similar process should be used to design definitive studies to inform decision rules on epoch length, non-wear definition, definition of a valid day, and number of valid days.

- NIH or another funding body should hold a competition to conduct studies based on the expert panel's recommendations. Once the results are ready, the expert

panel and study investigators should jointly make recommendations about methodologic decision rules for youth accelerometer studies.

- The consensus decision rules should be made generally available and communicated directly to relevant journal editors. They could be recommended like CONSORT and STROBE reporting standards. An important goal is to increase comparability in describing MVPA, sedentary behavior, and prevalence rates. One approach would be to require authors to use the consensus protocol and reporting standards as the default, then investigators could use and report additional methods or decision rules for specific purposes.

- Ideally the studies would be designed to create decision rules that would apply to the ActiGraph models used in the 183 studies summarized in the present paper as well as the GT3X+ model that provides raw data. It is essential that investigators know how to score raw data so they are comparable to results obtained from earlier models.

- Studies should concurrently examine other commonly used accelerometers and develop methods for creating comparable physical activity and sedentary behavior estimates across multiple brands of accelerometers.

- Recommendations should be formally reviewed every 5 years.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

1. Sallis JF, Saelens BE. Assessment of physical activity by self-report: Status, limitations, and future directions. Res Q Exerc Sport 2000;71:S1–S14. [PubMed: 10925819]

2. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med 2001;31:439–454. [PubMed: 11394563]

3. Corder K, Ekelund U, Steele RM, Wareham NJ, Brage S. Assessment of physical activity in youth. J Appl Physiol 2008;105:977–987. [PubMed: 18635884]

4. Oliver M, Schofield GM, Kolt GS. Physical activity in preschoolers: understanding prevalence and measurement issues. Sports Med 2007;37:1045–1070. [PubMed: 18027993]

5. Reilly JJ, Penpraze V, Hislop J, Davies G, Grant S, Paton JY. Objective measurement of physical activity and sedentary behaviour: review with new data. Arch Dis Child 2008; 93:614–619. [PubMed: 18305072]

6. Rowlands AV. Accelerometer assessment of physical activity in children: an update. Pediatr Exerc Sci 2007;19:252–266. [PubMed: 18019585]

7. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008;40(1):181–188. [PubMed: 18091006]

8. Ward DS, Evenson KR, Vaughn A, Rodgers AB, Troiano RP. Accelerometer Use in Physical Activity: Best Practices and Research Recommendations. Med Sci Sports Exerc 2005;37(Suppl):S582–S588. [PubMed: 16294121]

9. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005;37:S531–S543. [PubMed: 16294116]

10. de Vries SI, Bakker I, Hopman-Rock M, Hirasing RA, van Mechelen W. Clinimetric review of motion sensors in children and adolescents. J Clin Epidemiol 2006;59:670–680. [PubMed: 16765269]

11. Corder K, Brage S, Ramachandran A, Snehalatha C, Wareham N, Ekelund U. Comparison of two Actigraph models for assessing free-living physical activity in Indian adolescents. J Sports Sci 2007;25:1607–1611. [PubMed: 17852668]

12. John D, Tyo B, Bassett DR. Comparison of four ActiGraph accelerometers during walking and running. Med Sci Sports Exerc 2010;42(2):368–374. [PubMed: 19927022]

13. Kozey SL, Staudenmayer JW, Troiano RP, Freedson PS. Comparison of the Actigraph 7164 and the Actigraph GT1M during self-paced locomotion. Med Sci Sports Exerc 2010;42:971–976. [PubMed: 19997000]

14. Rothney MP, Apker G, Song Y, Chen KY. Comparing the performance of three generations of ActiGraph accelerometers. J Appl Physiol 2008;105(4):1091–1097. [PubMed: 18635874]

15. Baquet G, Stratton G, Van Praagh E, Berthoin S. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. Prev Med 2007;44(2):143–147. [PubMed: 17157370]

16. McClain JJ, Abraham T, Brusseau T, Tudor-Locke C. Epoch Length and Accelerometer Outputs in Children: Comparison to Direct Observation. Med Sci Sports Exerc 2008; 40(12):2080–2087. [PubMed: 18981941]

17. Vale S, Santos R, Silva P, Soares-Miranda L, Mota J. Preschool children physical activity measurement: importance of epoch length choice. Pediatr Exerc Sci 2009;21(4):413–420. [PubMed: 20128361]

18. Masse LC, Fuemmeler BF, Anderson B, et al. Accelerometer Data Reduction: A Comparison of Four Reduction Algorithms on Select Outcome Variables. Med Sci Sports Exerc 2005;37(11Suppl):S544–S554. [PubMed: 16294117]

19. Catellier DJ, Hannan PJ, Murray DM et al. Imputation of missing data when measuring physical activity by accelerometry. Med Sci Sports Exerc 2005;37(11 Suppl):S555–S562. [PubMed: 16294118]

20. Telford A, Salmon J, Jolley D, Crawford D. Reliability and Validity of Physical Activity Questionnaires for Children: The Children's Leisure Activities Study Survey (CLASS). Pediatr Exerc Sci 2004;16:64–78

21. Colley R, Gorber SC, Tremblay MS. Quality control and data reduction procedures for accelerometry-derived measures of physical activity. Health Reports 2010;21(1): Statistics Canada, Catalogue no. 82–003-XPE.

22. Mattocks C, Ness A, Leary 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(1Suppl):S98–S111. [PubMed: 18364528]

23. Fairclough SJ, Butcher ZH, Stratton G. Whole-day and segmented-day physical activity variability of northwest England school children. Prev Med 2007;44(5):421–425. [PubMed: 17320158]

24. 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(2):426–431. [PubMed: 10694127]

25. Trost SG, Pate RR, Sallis JF, et al. Age and gender differences in objectively measured physical activity in youth. Med Sci Sports Exerc 2002;34:350–355. [PubMed: 11828247]

26. Sirard J, Kubik M Fulkerson J, Arcan C Objectively Measured Physical Activity in Urban Alternative High School Students. Med Sci Sports Exerc 2008;40(12):2088–2095. [PubMed: 18981940]

27. Wrotniak BH, Epstein LH, Dorn JM, Jones KE, Kondilis VA. The relationship between motor proficiency and physical activity in children. Pediatrics 2006;118(6):e1758–e1765. [PubMed: 17142498]

28. Anderson CB, Hagströmer M, Yngve A. Validation of the PDPAR as an adolescent diary: effect of accelerometer cut points. Med Sci Sports Exerc 2005;37(7):1224–1230. [PubMed: 16015142]

29. Gidlow CJ, Cochrane T, Davey R, Smith H. In-school and out-of-school physical activity in primary and secondary school children. J Sports Sci 2008;26(13):1411–1419. [PubMed: 18942001]

30. Guinhouya CB, Hubert H, Soubrier S, et al. Moderate-to-vigorous physical activity among children: discrepancies in accelerometry-based cut-off points. Obesity 2006; 14(5):774–777. [PubMed: 16855185]

31. McClain JJ, Sisson SB, Washington TL, Craig CL, Tudor-Locke C. Comparison of Kenz Lifecorder EX and ActiGraph Accelerometers in 10-yr-old Children. Med Sci Sports Exerc 2007;39(4):630–638. [PubMed: 17414800]

32. Mota J, Valente M, Aires L, et al. Accelerometer cut points and youth physical activity prevalence. Euro Phys Edu Rev 2007;13(3):287–299.

33. Pate RR, Stevens J, Pratt C, et al. Objectively measured physical activity in sixth-grade girls. Arch Pediatr Adolesc Med 2006;160(12):1262–1268. [PubMed: 17146024]

34. Van Cauwenberghe E, Labarque V, Trost SG, Bourdeaudhuij I de, Cardon G Calibration and comparison of accelerometer cut points in preschool children. Int J Pediatr Obes 2011;6(2–2):e582–e589. [PubMed: 21121867]

35. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. Med Sci Sports Exerc 2005;37(Suppl):S501–S511. [PubMed: 16294113]

36. Freedson PS, Sirard J, Debold E, et al. Calibration of the Computer Science and Applications, Inc. (CSA) accelerometer. Med Sci Sports Exerc 1997;29(suppl):45. [PubMed: 9000155]

37. Mattocks C, Leary S, Ness A, et al. Calibration of an accelerometer during free-living activities in children. Int J Pediatr Obes 2007;2(4):218–226. [PubMed: 17852552]

38. Treuth MS, Schmitz K, Catellier DJ, et al. Defining accelerometer thresholds for activity intensities in adolescent girls. Med Sci Sports Exerc 2004;36(7):1259–1266. [PubMed: 15235335]

39. Sirard JR, Trost SG, Pfeiffer KA, Dowda M, Pate R. Calibration and evaluation of an objective measure of physical activity in preschool children. J Phys Act Health 2005;3:345–357.

40. Trost SG, Ward DS, Moorehead SM, Watson PD, Riner W, Burke JR. Validity of the Computer Science and Applications (CSA) activity monitor in children. Med Sci Sports Exerc 1998;30:629–633. [PubMed: 9565947]

41. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray R. Calibration of two objective measures of physical activity for children. J Sports Sci 2008;26(14):1557–1565. [PubMed: 18949660]

42. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. Obes Res 2002;10:150–157. [PubMed: 11886937]

43. Owen N, Healy GN, Mathews CE, Dunstan DW. Too much sitting: the population-health science of sedentary behavior. Exerc Sport Sci Rev 2010;38:105–113. [PubMed: 20577058]

44. Brage S, Wedderkopp N, Andersen LB, Froberg K. Influence of step frequency on movement intensity predictions with the CSA accelerometer: a field validation study in children. Pediatr Exerc Sci 2003;15:277–87.

45. Ekelund U, Sardinha LB, Andersson, et al., Associations between objectively assessed physical activity and indicators of body fatness in 9- and 10-yr old children: a population based study from 4 distinct regions in Europe (the European Youth Heart Study). Am J Clin Nutr 2004;80:584–590. [PubMed: 15321796]

46. Pate RR, Almeida MJ, McIver KL, Pfeiffer KA, Dowda M. Validation and calibration of an accelerometer in preschool children. Obesity 2006;14:2000–2006. [PubMed: 17135617]

47. Reilly JJ, Coyle J, Kelly L, Burke G, Grant S, Paton JY. An Objective Method for Measurement of Sedentary Behavior in 3- to 4-Year Olds. Obesity 2003;11:1155–1158.

48. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998;30:777–781. [PubMed: 9588623]

49. Krishnaveni GV, Mills IC, Veena SR, et al. Acelerometers for measuring physical activity behavior in Indian children. Indian Pediatr 2009;46(12):1055–1062. [PubMed: 19430084]

50. Metcalf BS, Voss LD, Wilkin TJ. Accelerometers identify inactive and potentially obese children (EarlyBird 3). Arch Dis Child 2002;87:166–167. [PubMed: 12138075]

51. Rowlands AV, Pilgrim EL, Eston RG. Patterns of Habitual Activity across Weekdays and Weekend Days in 9–11-Year-Old Children. Prev Med 2008;46(4):317–24. [PubMed: 18162187]

52. Dowda M, Pate R, Sallis JF, Freedson PS. Accelerometer (CSA) count cut points for physical activity intensity ranges in youth. Med Sci Sports Exerc 1997;29(suppl):72.

53. Nillson 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.

54. Edwardson CL, Gorely T. Epoch Length and Its Effect on Physical Activity Intensity. Med Sci Sports Exerc 2010;42(5):928–934. [PubMed: 19996997]

55. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011;43(2):357–364. [PubMed: 20581716]

56. Esliger DW, Copeland JL, Barnes JD, Tremblay MS. Standardizing and optimizing the use of accelerometer data for free living physical activity monitoring. J Phys Act Health 2005;3:366–383.

57. Macfarlane D, Lee C, Ho E, Chan K, Chan D. Convergent validity of six methods to assess physical activity in daily life. J Appl Physiol 2006;101:1328–1334. [PubMed: 16825525]

58. Penpraze V, Reilly J, MacLean C, et al. Monitoring of physical activity in young children: how much is enough? Pediatr Exerc Sci 2006;18:483–491.

59. Sallis J, Saelens B, Frank L, et al. Neighborhood built environment and income: Examining multiple health outcomes. Soc Sci Med 2009;68:1285–1293. [PubMed: 19232809]

60. Sirard JR, Slater ME. Compliance with Wearing Physical Activity Accelerometers in High School Students. J Phys Act Health 2009;6(1Suppl):S148–S55. [PubMed: 19998861]

61. Trost SG. State of the Art Reviews: Measurement of Physical Activity in Children and Adolescents. Am J Lifestyle Med 2007;1:299–314.

62. United States Department of Health and Human Services physical activity guidelines for Americans. http://www.health.gov/PAGuidelines/guidelines/chapter3.aspx. Accessed 12/31/08.

63. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. Med Sci Sports Exerc 2011;43(7):1360–1368. [PubMed: 21131873]

Author Manuscript

Articles retrieved for more detailed evaluation:	**n=573**

Studies excluded:	**n=300**
No children, or no methods reported for children:	n=56
Not ActiGraph accelerometer:	n=131
Studies not measuring ambulatory movement:	n=19
Studies with <1 complete wearing days (e.g., recess studies, reviews, methods papers):	n=94

|	
Relevant articles for inclusion:	**n=273**
Preschool (ages 2-5 years):	n=25
Children (ages 6-11 years):	n=140
Adolescents (ages 12-18 years):	n=108

FIGURE 1.
Results of the literature search
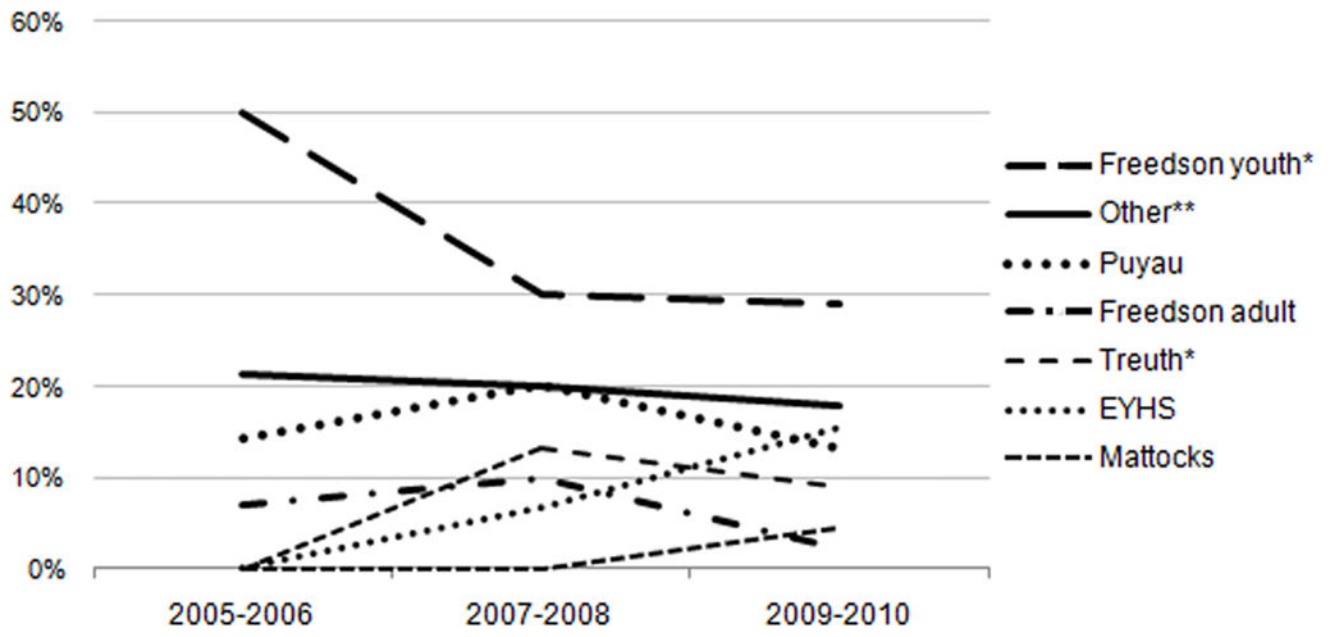
* includes all MET values associated with cut point
** includes Welk 2005, Trost 1998, Metcalf 2002, Krishnaveni 2009, used more than 1, and used own cut point

FIGURE 2.
Trends in accelerometer cut points (MVPA) reported in children's studies from 2005–2010

* includes all MET values associated with cut point
** includes Trost 1998, used more than 1 cut point, and used own cut point
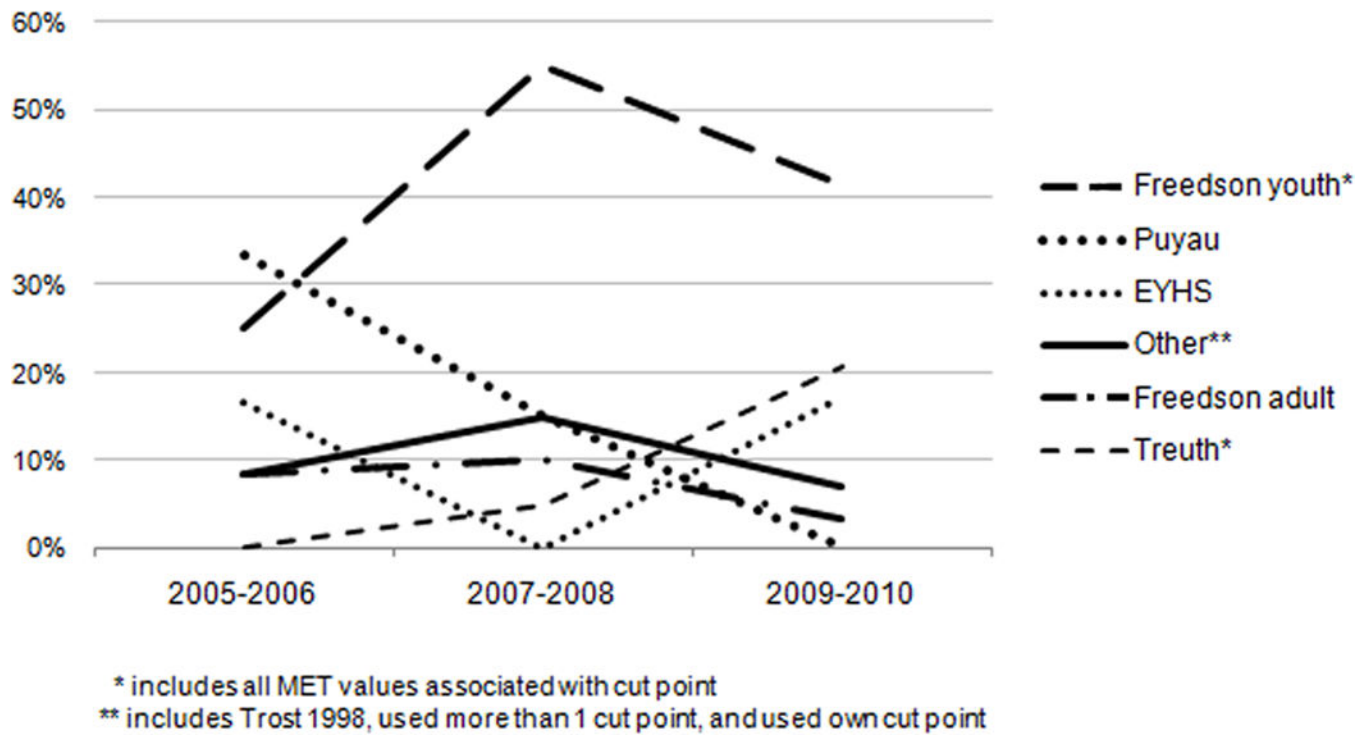
FIGURE 3.
Trends in accelerometer cut points (MVPA) reported in adolescent studies from 2005–2010
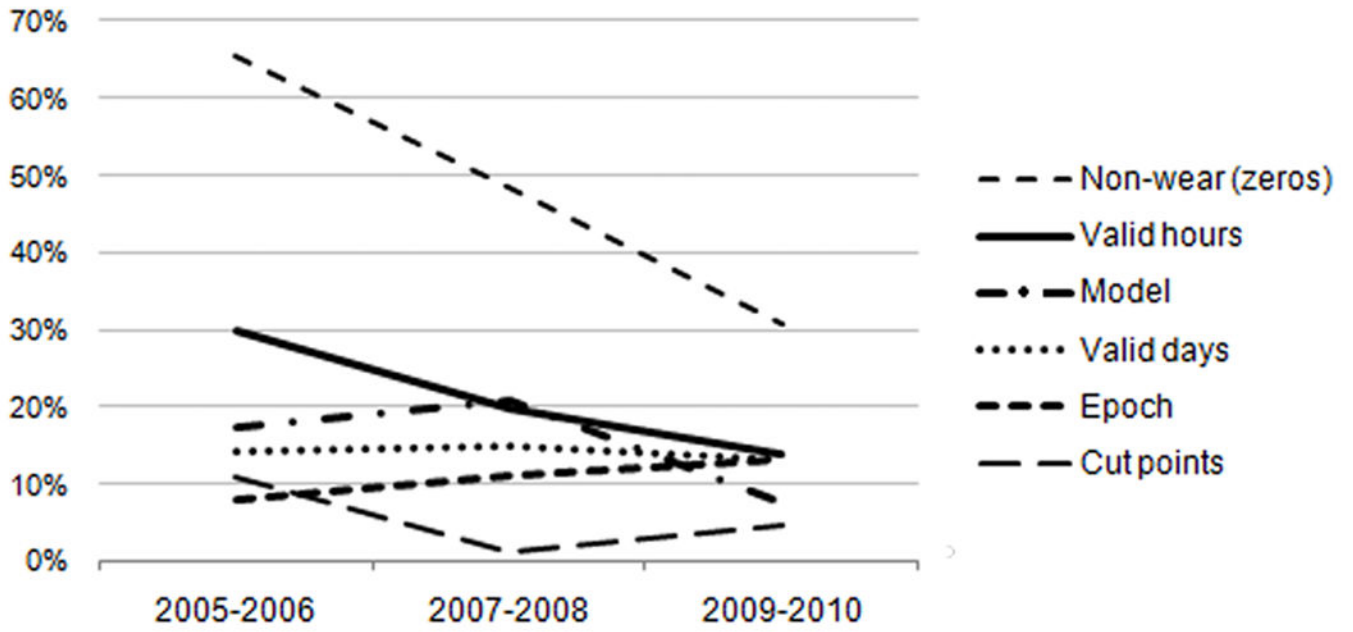
FIGURE 4.
Methods not reported in papers from 2005–2010 (% of papers not reporting)
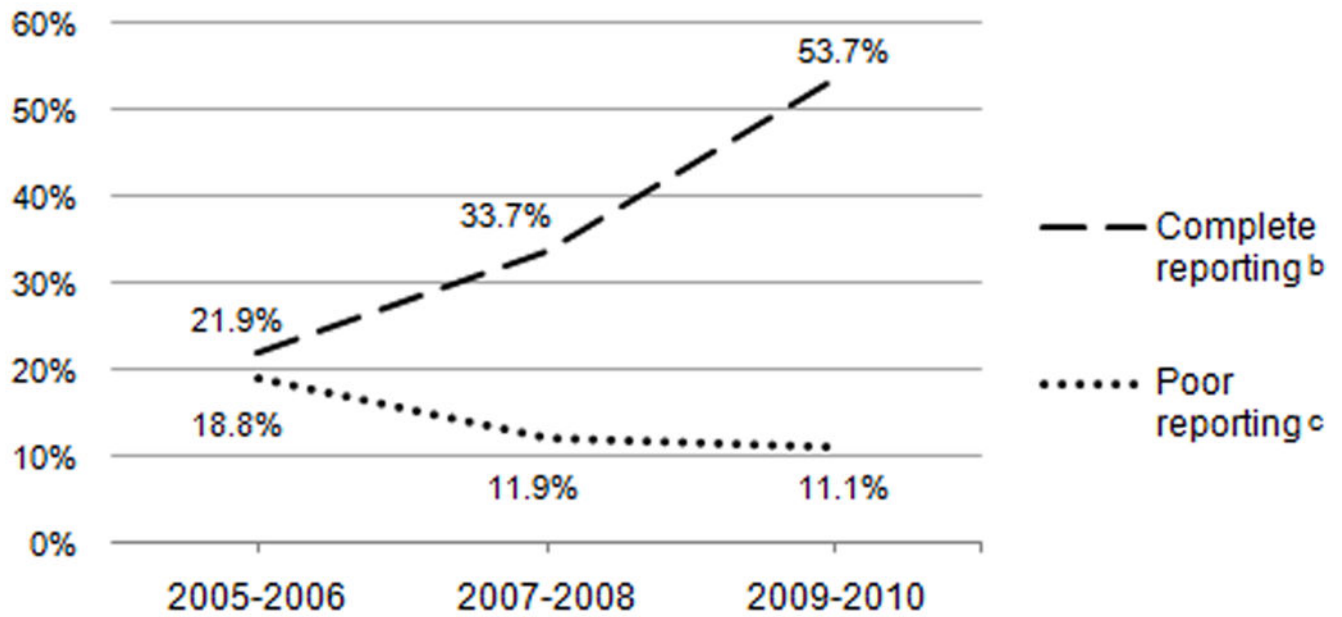
a calculated as # of key methods reported / # total relevant methods
b article reported on all (100%) of key methods
c article reported on 50% or less key methods

FIGURE 5.
Trends in percent of youth accelerometer papers with "complete" versus "poor" reporting scores[a] from 2005–2010

Table 1.

Summary of key methodological issues from a review of 183 youth accelerometer studies from 2005–2010.

	All youth studies (n=183) % (n)	Preschool studies (n=21) % (n)	Children studies (n=94) % (n)	Adolescent studies (n=68) % (n)
Year of publication				
2005–2006	20.8 (38)	30.0 (6)	14.9 (14)	26.5 (18)
2007–2008	32.8 (60)	30.0 (6)	35.1 (33)	30.9 (21)
2009–2010	46.4 (85)	40.0 (9)	50.0 (47)	42.6 (29)
Actigraph model				
7164	67.8 (124)	71.4 (15)	66.0 (62)	69.1 (47)
GT1M	17.5 (32)	9.5 (2)	23.4 (22)	11.8 (8)
Model not specified (e.g., 'Actigraph')	14.8 (27)	19.0 (4)	10.6 (10)	19.1 (13)
Epoch (seconds)				
2	2.2 (4)	0.0 (0)	4.3 (4)	0.0 (0)
5	4.9 (9)	9.5 (2)	6.4 (6)	1.5 (1)
10	5.5 (10)	0.0 (0)	9.6 (9)	1.5 (1)
15	7.7 (14)	42.9 (9)	2.1 (2)	4.4 (3)
30	7.7 (14)	4.8 (1)	4.3 (4)	13.2 (9)
60	62.3 (114)	42.9 (9)	66.0 (62)	63.2 (43)
Missing	9.8 (18)	0.0 (0)	7.4 (7)	16.2 (11)
Non-wear time within a day (consecutive minutes of '0' counts)[a]	(n=181)		(n=92)	
10	17.1 (31)	9.5 (2)	17.4 (16)	19.1 (13)
15	1.7 (3)	0.0 (0)	3.3 (3)	0.0 (0)
20	22.1 (40)	9.5 (2)	20.7 (19)	27.9 (19)
30	5.0 (9)	0.0 (0)	3.3 (3)	8.8 (6)
60	5.0 (9)	14.3 (3)	6.5 (6)	0.0 (0)
180	0.6 (1)	0.0 (0)	0.0 (0)	1.5 (1)
Missing	48.6 (88)	66.7 (14)	48.9 (45)	42.6 (29)
Definition of a valid day (number of valid wear hours)				
<6	2.7 (5)	9.5 (2)	3.2 (3)	0.0 (0)

	All youth studies (n=183) % (n)	Preschool studies (n=21) % (n)	Children studies (n=94) % (n)	Adolescent studies (n=68) % (n)
6	3.8 (7)	14.3 (3)	4.3 (4)	0.0 (0)
6.7	0.5 (1)	4.8 (1)	0.0 (0)	0.0 (0)
8	11.5 (21)	9.5 (2)	10.6 (10)	13.2 (9)
8.1 – 9.9	4.9 (9)	4.8 (1)	8.5 (8)	0.0 (0)
10	29.0 (53)	14.3 (3)	21.3 (20)	44.1 (30)
10.1 – 11.9	0.5 (1)	0.0 (0)	1.1 (1)	0.0 (0)
12	3.3 (6)	0.0 (0)	(2)	5.9 (4)
10 on weekdays, 8 on weekend days	2.2 (4)	0.0 (0)	4.3 (4)	0.0 (0)
12 on weekdays, 10 on weekend days	1.1 (2)	0.0 (0)	2.1 (2)	0.0 (0)
10 on weekdays, 60% mean weekend WT	0.5 (1)	0.0 (0)	1.1 (1)	0.0 (0)
Ratio or % (e.g., 80% of average WT, 70/80 rule [b])	5.5 (10)	4.8 (1)	6.4 (6)	4.4 (3)
Counts criteria (e.g., <10,000 and >20,000,000= invalid day)	3.3 (6)	0.0 (0)	5.3 (5)	1.5 (1)
Accelerometer worn during predefined moments and valid hours not used (e.g., school hours only)	2.7 (5)	9.5 (2)	3.2 (3)	0.0(0)
Unclear or other	2.7 (5)	0.0 (0)	3.2 (3)	2.9 (2)
Varied by purpose of analyses [c]	1.1 (2)	4.8 (1)	0.0 (0)	1.5 (1)
Missing	24.6 (45)	23.8 (5)	23.4 (22)	26.5 (18)
Minimum number of wear days to be included in analyses				
1	1.1 (2)	0.0 (0)	2.1 (2)	0.0 (0)
2	6.6 (12)	4.8 (1)	9.6 (9)	2.9 (2)
3	28.4 (52)	19.0 (4)	29.8 (28)	29.4 (20)
4	26.8 (49)	23.8 (5)	26.6 (25)	27.9 (19)
5	7.1 (13)	0.0 (0)	5.3 (5)	11.8 (8)
6	1.1 (2)	4.8 (1)	1.1 (1)	0.0 (0)
7	4.4 (8)	9.5 (2)	4.3 (4)	2.9 (2)
10 or more	1.1 (2)	4.8 (1)	1.1 (1)	0.0 (0)
Varied by purpose of analyses [c]	5.5 (10)	9.5 (2)	3.2 (3)	7.4 (5)
Missing	18.0 (33)	23.8 (5)	17.0 (16)	17.6 (12)
Number of weekend days to be included in analyses				

	All youth studies (n=183) % (n)	Preschool studies (n=21) % (n)	Children studies (n=94) % (n)	Adolescent studies (n=68) % (n)
0	57.9 (106)	61.9 (13)	51.1 (48)	66.2 (45)
1	23.0 (42)	14.3 (3)	30.9 (29)	14.7 (10)
2	2.7 (5)	4.8 (1)	2.1 (2)	2.9 (2)
Varied by purpose of analyses [c]	0.5 (1)	0.0 (0)	1.1 (1)	0.0 (0)
Missing	15.8 (29)	19.0 (4)	14.9 (14)	16.2 (11)
Cut point for MVPA (counts/min) [d]	*(n=178)*	*(n=21)*	*(n=92)*	*(n=65)*
400 (Krishnaveni)[49]	0.6 (1)	0.0 (0)	1.1 (1)	0.0 (0)
614–2274 (Freedson youth 3METs, 6–18 years [e])[36]	29.2 (52)	9.5(2)	29.3 (27)	35.4 (23)
1160 (Treuth, 3METs)[33]	1.7 (3)	0.0(0)	2.2 (2)	1.5 (1)
1400–3500 (Freedson youth, 4METs, 6–18 years [e])[36]	2.8 (5)	0.0 (0)	3.3 (3)	3.1 (2)
1680 (Pate)[46]	2.2 (4)	19.0(4)	0.0 (0)	0.0 (0)
1952 (Freedson adult, 3 METs)[48]	5.6 (10)	4.8 (1)	4.3 (4)	7.7 (5)
2000 (European Youth Heart Study)[44-45]	9.6 (17)	0.0 (0)	10.9 (10)	10.8 (7)
2172 (Welk)[35]	1.1 (2)	0.0 (0)	2.2 (2)	0.0 (0)
2460–3564 (Sirard, 3–5 years)[39]	4.5 (8)	38.1 (8)	0.0 (0)	0.0 (0)
2500 (Metcalf)[50]	1.1 (2)	0.0 (0)	2.2 (2)	0.0 (0)
3000 (Treuth, 4.6METs)[38]	6.7 (12)	0.0 (0)	7.6 (7)	7.7 (5)
3200 (Puyau)[42]	12.9 (23)	19.0(4)	13.0 (12)	10.8 (7)
3581/3600 (Mattocks, 4METs)[37]	1.7 (3)	0.0 (0)	2.2 (2)	1.5 (1)
Energy expenditure equation (Trost)[40]	2.8 (5)	0.0 (0)	2.2 (2)	4.6 (3)
More than 1, used own cut point	7.3 (13)	4.8 (1)	7.6 (7)	7.7 (5)
Used counts only (no cut points)	6.2 (11)	4.8 (1)	8.7 (8)	3.1 (2)
Missing	3.9 (7)	0.0 (0)	3.3 (3)	6.2 (4)
Cut point for sedentary time (counts/min) [f]	*(n=90)*	*(n=20)*	*(n=38)*	*(n=32)*
9 (Krishnaveni)[49]	1.1 (1)	0.0 (0)	2.6 (1)	0.0 (0)
100 (Treuth, Evenson)[38,41]	28.9 (26)	5.0 (1)	26.3 (10)	46.9 (15)
151 (Pate)[46]	3.3 (3)	15.0 (3)	0.0 (0)	0.0 (0)

	All youth studies (n=183) % (n)	Preschool studies (n=21) % (n)	Children studies (n=94) % (n)	Adolescent studies (n=68) % (n)
199 (Mattocks)[37]	4.4 (4)	0.0 (0)	5.3 (2)	6.3 (2)
299 (Rowlands)[51]	3.3 (3)	0.0 (0)	7.9 (3)	0.0 (0)
499 (European Youth Heart Study)[44-45]	13.3 (12)	0.0 (0)	21.1 (8)	12.5 (4)
799 (Puyau)[42]	17.8 (16)	0.0 (0)	21.1 (8)	25.0 (8)
999 (Metcalf)[50]	2.2 (2)	0.0 (0)	5.3 (2)	0.0 (0)
1099 (Reilly)[47]	8.9 (8)	30.0 (6)	5.3 (2)	0.0 (0)
1204–1259 (Sirard, 3–5 years)[39]	7.8 (7)	35.0 (7)	0.0 (0)	0.0 (0)
Freedson (<1 – 1.9 METs)[36]	3.3 (3)	5.0 (1)	2.6 (1)	3.1 (1)
More than 1, used own cut point	2.2 (2)	5.0 (1)	0.0 (0)	3.1 (1)
Missing	3.3 (3)	5.0 (1)	2.6 (1)	3.1 (1)

WT = wearing time

MET = Metabolic Equivalent Task

[a] Only studies where accelerometer was worn for a whole day were included.

[b] Using the 70/80 rule, a valid day can be defined as the period during which at least 70% of the study population has recorded accelerometer data, and 80% of that observed period constitutes a minimal day for inclusion in data analysis.

[c] In instances where the cohort and stage of study were the same but methods varied, these were noted as 'method varying by analysis or purpose'

[d] Only studies reporting results on moderate or MVPA were included in this summary. Studies only reporting on sedentary or vigorous were not included. Values represent lower threshold.

[e] Includes Dowda,[52] Trost,[25] and Freedson[36] references, and reference without the MET value specifically stated.

[f] Only studies reporting results on sedentary behaviors were included in this summary. Studies using counts only were not included. Values represent upper threshold with lower threshold being "0".